



UNIVERSIDAD NACIONAL DE LA PLATA

FACULTAD DE CIENCIAS EXACTAS

DEPARTAMENTO DE FÍSICA

Trabajo de Tesis Doctoral:

**Teorías efectivas en antiferromagnetos cuánticos para modelar
materiales magnéticos no convencionales**

Tesista: Santiago Daniel Acevedo

Director/a: Carlos Alberto Lamas

Codirector/a: Aníbal Iucci

Año: 2022

Agradecimientos: A mi familia, amigos, amigas y docentes.

La gran mayoría de los resultados presentes en esta tesis doctoral forman parte de cinco publicaciones científicas que detallamos a continuación.

- 2021 **Effective field theory approach for the $S = 3/2$ bilayer honeycomb antiferromagnet**
S. ACEVEDO, C.A. LAMAS, AND P. PUJOL
Physical Review B
<https://doi.org/10.1103/PhysRevB.104.214412>
- 2021 **Phase diagram study of a two-dimensional frustrated antiferromagnet via unsupervised machine learning**
S. ACEVEDO, M. ARLEGO AND C.A. LAMAS
Physical Review B
<https://doi.org/10.1103/PhysRevB.103.134422>
- 2021 **Exploring neural network training strategies to determine phase transitions in frustrated magnetic models**
I. CORTE, S. ACEVEDO, M. ARLEGO AND C.A. LAMAS
Computational materials science
<https://doi.org/10.1016/j.commatsci.2021.110702>
- 2020 **Current jumps in flat band ladders with Dzyaloshinskii-Moriya interactions**
S. ACEVEDO, P. PUJOL AND C.A. LAMAS
Physical Review B
<https://doi.org/10.1103/PhysRevB.102.195139>
- 2019 **Magnon crystals and magnetic phases in a Kagomé-stripe antiferromagnet**
S. ACEVEDO, C.A. LAMAS, M. ARLEGO, P. PUJOL
Physical Review B
<https://doi.org/10.1103/PhysRevB.100.195145>

La estructura del texto es la siguiente. El capítulo 1 comienza con una introducción general a algunos aspectos centrales en teorías de campos en materia condensada. Luego continúa desarrollando el formalismo de integrales de camino a partir de estados coherentes de espín, y finaliza dando un ejemplo concreto; la teoría de campos efectiva para una cadena de Heisenberg antiferromagnética. El capítulo 2 se centra en el estudio analítico y numérico de antiferromagnetos cuánticos frustrados y agrupa tres de nuestros trabajos, ordenados cronológicamente y a la vez por nivel de complejidad, en tres geometrías diferentes; la cadena Kagomé, la cadena diente de sierra y la bicapa hexagonal (panal de abejas). El capítulo 3 comienza con una introducción a las redes neuronales artificiales y continúa desarrollando nuestros dos trabajos relacionados con la aplicación de técnicas de aprendizaje automático al estudio de sistemas magnéticos en el límite de alta anisotropía de Ising, con el mismo orden que en el capítulo anterior. Para finalizar, el capítulo 4 presenta las conclusiones globales, y el capítulo 5, el apéndice.

Tabla de contenidos

1	Teorías de campos en sistemas magnéticos	6
1.1	Introducción	6
1.1.1	Modelo XY: Defectos topológicos y la transición BKT	7
1.1.2	Grupo de Renormalización	9
1.2	Estados coherentes de espín	13
1.2.1	Bosones de Schwinger	13
1.2.2	Definición y propiedades	14
1.3	Integrales funcionales en sistemas de espines	17
1.4	Teoría de campos efectiva para la cadena de espín antiferromagnética	19
1.4.1	Transformación de dualidad	22
2	Antiferromagnetos cuánticos frustrados en baja dimensión	25
2.1	Cadena Kagomé	25
2.1.1	Introducción: Cristal de magnones	25
2.1.2	Hamiltoniano perturbativo efectivo	28
2.1.3	Primeras predicciones	36
2.1.4	Plateaux fraccionarios	38
2.1.5	Conclusiones	46
2.2	Cadena diente de sierra	47
2.2.1	Introducción	47
2.2.2	Cristal de magnones	49
2.2.3	Teoría de campos efectiva	50
2.2.4	Resultados numéricos (DMRG)	52
2.2.5	Conclusiones	60
2.3	Bicapa hexagonal	61
2.3.1	Introducción	62
2.3.2	Teoría de campos efectiva	63
2.3.3	Orden por desorden	66
2.3.4	Fase de dímeros	69
2.3.5	Conclusiones	70
3	Aprendizaje automático de fases magnéticas	72
3.1	Introducción a las redes neuronales artificiales	72
3.1.1	Redes neuronales completamente conectadas	72
3.1.2	Sobreajuste y técnicas de regularización	76
3.1.3	Redes neuronales convolucionales	77
3.2	Modelos	78
3.3	Aprendizaje supervisado en antiferromagnetos de Ising	79
3.3.1	Transferencia de aprendizaje	82

3.3.2	Aprendizaje por confusión	85
3.3.3	Conclusiones	87
3.3.4	Tecnicismos	88
3.4	Aprendizaje no supervisado: AutoEncoders	90
3.4.1	Entrenamientos a baja temperatura	91
3.4.2	Diagrama de fases	98
3.4.3	Entrenamientos a alta temperatura	99
3.4.4	Conclusiones	103
3.4.5	Tecnicismos	104
3.4.6	Análisis de componentes principales (PCA)	105
4	Conclusiones generales	108
5	Apéndice	111
5.1	Density Matrix Renormalization Group (DMRG)	111
5.1.1	Líneas centrales del algoritmo	111
5.1.2	Estados de productos de matrices	112
5.2	Simulación Monte Carlo	114
5.3	Maximum likelihood estimation	119

1 Teorías de campos en sistemas magnéticos

1.1 Introducción

Comenzamos presentando una breve descripción del modelo XY clásico en dos dimensiones espaciales para introducir la transición de fase topológica que aparecerá con frecuencia a lo largo nuestro estudio de sistemas magnéticos en la formación de plateaux de magnetización.

En primer lugar, mencionamos el teorema de Mermin y Wagner, debido a que nuestros sistemas magnéticos de estudio corresponden a lo que se conoce como sistemas en baja dimensionalidad. Esto quiere decir que nuestros arreglos de objetos interactuantes se disponen en geometrías con dimensión espacial menor o igual a dos.

Teorema de Mermin y Wagner

En general el teorema de Mermin y Wagner[1] afirma que no puede haber ruptura espontánea de una simetría continua en sistemas con interacciones de corto alcance en dimensiones $d \leq 2$. Esta dimensión límite $d = 2$ es conocida como la dimensión crítica inferior. Para el caso particular de sistemas de espines, usando el Hamiltoniano de Heisenberg

$$H = \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j, \quad (1.1)$$

la simetría continua es la rotación colectiva de todos los espines, o simetría completa de $SU(2)$. Aquí, interacciones de corto alcance significa

$$\frac{1}{N} \sum_{ij} |J_{ij}| |\mathbf{x}_i - \mathbf{x}_j|^2 < \infty, \quad (1.2)$$

donde N es el número total de espines, y la ausencia de ruptura espontánea de simetría significa

$$\lim_{N \rightarrow \infty} m(N) = 0, \quad (1.3)$$

con m el parámetro de orden del sistema.

1.1.1 Modelo XY: Defectos topológicos y la transición BKT

El modelo XY estándar consiste en espines clásicos de dos componentes con interacciones a primeros vecinos. El Hamiltoniano del sistema es

$$H = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j), \quad (1.4)$$

donde $\langle i, j \rangle$ indica que la suma está tomada sobre los primeros vecinos, $J > 0$ (ferromagneto) y el ángulo $\theta_i \in [0, 2\pi)$ marca la orientación del espín del sitio i . En Hamiltonianos clásicos es usual escribir $K = \beta J > 0$, de manera tal que

$$-\beta H = K \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j). \quad (1.5)$$

Vamos a centrarnos en el caso dos-dimensional¹, donde el sistema no puede tener orden de largo alcance a causa del teorema de Mermin y Wagner. Sin embargo, puede probarse[2] que a alta temperatura el modelo tiene correlaciones de dos puntos que decaen exponencialmente

$$\langle \mathbf{S}_1 \cdot \mathbf{S}_r \rangle \sim e^{-\frac{r}{\xi}}, \quad (1.6)$$

siendo ξ la longitud de correlación, mientras que a baja temperatura las correlaciones decaen algebraicamente

$$\langle \mathbf{S}_1 \cdot \mathbf{S}_r \rangle \sim r^{-\alpha}, \quad \alpha > 0. \quad (1.7)$$

Un decaimiento del tipo ley de potencias implica autosimilaridad (no hay longitud de correlación) y corresponde usualmente a un punto crítico. Se dice que esta fase de baja temperatura del modelo XY posee orden de cuasi-largo alcance (*quasi-long-range order*), para diferenciarlo del verdadero orden de largo alcance que está acompañado de una magnetización finita. Los distintos decaimientos asintóticos de las correlaciones a altas y bajas temperaturas sugiere la presencia de una transición de fase entre ambos regímenes. Berezinskii, Kosterlitz y Thouless propusieron que el desorden está causado por defectos topológicos que no pueden considerarse como simples deformaciones del estado fundamental. La temperatura a la cual ocurre la transición, T_{KT} puede encontrarse numéricamente, o con técnicas del grupo de renormalización. A continuación mostraremos un enfoque simplificado para motivar la transición con argumentos de energía y entropía.

El argumento general es el siguiente. Cuando se calcula el cambio en la energía por agregar un vórtice aislado de vorticidad (o carga topológica) $q = 1$ al sistema, el resultado es proporcional al logaritmo del número N de partículas del sistema. Pero cuando tenemos en cuenta el número de sitios en los que puede ubicarse el centro del vórtice, hay una contribución entrópica que también es proporcional a $\log N$. Luego, al calcular la energía libre $F = E - TS$, el resultado es que para bajas temperaturas la creación de vórtices libres es energéticamente desfavorable y para altas temperaturas los vórtices proliferan, destruyendo el orden de cuasi-largo alcance.

Incluso en el régimen de alta temperatura, dos espines que están cerca entre sí van a interactuar, por lo que expandimos el coseno en (1.4) y, en el límite continuo, el Hamiltoniano

¹A lo largo de este documento utilizaremos la notación ‘dos-dimensional’ (o ‘uno-dimensional’) para referirnos a modelos y/o geometrías en dos (o una) dimensiones.

puede escribirse

$$H = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j) \approx \frac{J}{2} \iint d^2r (\partial_x \theta)^2 + (\partial_y \theta)^2 = \frac{J}{2} \iint d^2r (\nabla \theta)^2. \quad (1.8)$$

Del lado derecho de la ecuación anterior estamos obteniendo una teoría efectiva clásica de campos para el sistema. Aquí remarcamos que si minimizamos la energía tomando la derivada funcional $\delta H / \delta \theta = 0$, obtenemos la ecuación de movimiento $\nabla^2 \theta = 0$, que tiene soluciones regulares y también soluciones singulares.

Consideremos la siguiente configuración de espines

$$\theta_v(\mathbf{r}) = \varphi(\mathbf{r}) + \theta_0 \quad (1.9)$$

donde $\tan \varphi(\mathbf{r}) = y/x$, siendo $\mathbf{r} = (x, y)$, y θ_0 algún ángulo constante de referencia. Esta configuración es una solución singular de la ecuación de Laplace y corresponde a un único vórtice aislado, ubicado en $\mathbf{r} = 0$, con vorticidad $q = 1$, definida por la ecuación

$$2\pi q = \oint_C \nabla \theta(\mathbf{r}) d\mathbf{r}, \quad (1.10)$$

donde tenemos que extender el dominio de $\theta(\mathbf{r})$ a $(-\infty, +\infty)$. El gradiente de θ_v es $\nabla \theta_v = \frac{1}{r} \mathbf{e}_\varphi$, con $\mathbf{e}_\varphi = (-\sin \varphi, \cos \varphi)$ y entonces, la energía de un único vórtice aislado con vorticidad 1 es

$$H_v \approx \frac{J}{2} \int_a^R \int_0^{2\pi} r dr d\varphi \left(\frac{1}{r}\right)^2 = \pi J \log \frac{R}{a}, \quad (1.11)$$

donde a es el tamaño del núcleo del vórtice, que se introduce para regularizar la integral. En la ecuación anterior vemos que la energía crece logarítmicamente con el tamaño del sistema, y no estamos escribiendo la contribución energética del núcleo que corresponde a integrar en distancias dentro de un disco de radio a . Por otra parte, la entropía de una configuración con un único vórtice también crece logarítmicamente con el tamaño, dado que hay aproximadamente R^2/a^2 posiciones donde puede ubicarse el centro del vórtice:

$$S_v = 2k_B \log \frac{R}{a}. \quad (1.12)$$

Por lo tanto, la energía libre de Helmholtz para una configuración con un único vórtice aislado de vorticidad 1 es

$$F = U - TS = (\pi J - 2k_B T) \log \frac{R}{a}. \quad (1.13)$$

A baja temperatura entonces el costo energético de crear un vórtice aislado es demasiado alto, pero a $K_c = 2/\pi$, la proliferación de vórtices se vuelve favorable a causa del término entrópico en la energía libre. Los primeros vórtices en aparecer son aquellos con $q \pm 1$, ya que vorticidades mayores implican un mayor costo energético. El modelo XY puede mapearse a un gas de Coulomb en donde los vórtices interactuantes corresponden a partículas con carga[2]. La fase de bajas temperaturas puede interpretarse entonces como un gas neutro de dipolos fuertemente ligados (un par vórtice-antivórtice muy cercanos entre sí), mientras que la fase de altas temperaturas corresponde a un gas neutro de cargas aisladas (plasma).

1.1.2 Grupo de Renormalización

El grupo de renormalización puede ser presentado como un marco teórico o un conjunto de ideas que tienen en común el objetivo de reexpresar los parámetros que definen un dado problema en términos de otros, quizás más simples, dejando invariantes los aspectos físicos del problema de interés. Los contextos para aplicar las ideas del grupo de renormalización son muy variados, sin embargo en el marco de esta tesis el escenario es el de los fenómenos críticos y las teorías de campos en física estadística. Estas teorías trabajan en el límite continuo, introduciendo campos $\phi(x)$ que pueden corresponder por ejemplo a la magnetización, mientras que la información de la naturaleza discreta la red cristalina se introduce con la presencia de un valor de corte Λ para los valores de momentos k permitidos en la transformada de Fourier $\phi_{\mathbf{k}}$. Esto se debe a que variaciones en los campos en distancias menores a la distancia característica de la red cristalina no tienen significado físico. La transformación del grupo de renormalización, en lo que se denomina ‘el enfoque de Wilson’, consiste en integrar grados de libertad microscópicos, dejando invariante la física de grandes distancias. Esta integración puede hacerse tanto en espacio real como en espacio de momentos, siendo el segundo el enfoque más popular por comodidad a la hora de calcular. La transformación es analítica y preserva la función de partición del sistema. A partir de ella pueden construirse diagramas de fase, calculando las llamadas dimensiones de escala (*scaling dimensions*) y funciones de escala (*scaling functions*), objetos que definen las llamadas clases de universalidad, y permiten determinar los exponentes críticos de sistemas de muchos cuerpos, entre otras cantidades de interés.

Bibliografía sobre el grupo de renormalización y teorías de campos en física estadística puede encontrarse por ejemplo en las referencias [2, 3, 4] y en las clases de David Tong². A continuación, para mostrar algunas ideas centrales del grupo de renormalización, mostraremos la renormalización de la llamada teoría Gaussiana para un campo escalar.

Renormalización de la teoría Gaussiana

Sea ϕ un campo escalar en d dimensiones. Consideramos a ϕ definido en una red cristalina hipercúbica con distancia a entre sitios vecinos, y definimos para ϕ la energía libre de Landau-Ginzburg en el límite continuo, dada por

$$\mathcal{F}[\phi] = \int d^d x \left(\frac{1}{2} (\nabla \phi)^2 + \frac{\mu_0^2}{2} \phi^2 + \frac{g_0}{4!} \phi^4 \right). \quad (1.14)$$

Muchos resultados y muchas definiciones en teorías de campos en física estadística provienen de teorías de campos en física de partículas. Así podemos llamar entonces al primer término en la ecuación anterior el término cinético, al segundo el término de masa y al tercero el término de interacción. En física de partículas μ_0 es la masa de una dada partícula y g_0 la constante de acoplamiento entre partículas. Desde el punto de vista de física estadística podemos considerar estos parámetros simplemente como parámetros efectivos de la teoría, y pueden tener diferentes interpretaciones en diferentes contextos físicos, como superconductividad, superfluidez, magnetismo, etc[2]. Bajo la convención de unidades donde \mathcal{F} es adimensional, el coeficiente del término cinético es a menudo fijado a 1/2, permitiendo fijar las dimensiones del campo ϕ . La

²<http://www.damtp.cam.ac.uk/user/tong/teaching.html>

función de partición del sistema está dada por

$$\mathcal{Z} = \int \mathcal{D}\phi e^{-\mathcal{F}[\phi]} \quad (1.15)$$

donde $\int \mathcal{D}\phi$ es una integral funcional, que corresponde a el límite continuo de $\Pi_{\mathbf{x}} \int_{-\infty}^{\infty} d\phi(\mathbf{x})$.

Sea $\Lambda \propto 1/a$, el valor de corte tal que $\phi_{\mathbf{k}} = 0$ si $|\mathbf{k}| > \Lambda$. Para $g_0 = 0$ la energía libre se denomina Gaussiana, libre, o no interactuante, y puede calcularse exactamente de manera analítica pasando al espacio de momentos, escribiendo

$$\phi(\mathbf{x}) = \int \frac{d^d k}{(2\pi)^d} e^{i\mathbf{k}\cdot\mathbf{x}} \phi_{\mathbf{k}}, \quad (1.16)$$

donde $\phi_{-\mathbf{k}} = \phi_{\mathbf{k}}^*$, ya que $\phi(\mathbf{x})$ es real. Usando la expresión anterior podemos reescribir a la energía libre como

$$\mathcal{F}[\phi] = \frac{1}{2} \int \frac{d^d k}{(2\pi)^d} (k^2 + \mu_0^2) \phi_{\mathbf{k}} \phi_{-\mathbf{k}} = \frac{1}{2} \int \frac{d^d k}{(2\pi)^d} (k^2 + \mu_0^2) |\phi_{\mathbf{k}}|^2, \quad (1.17)$$

donde $k = |\mathbf{k}|$. Aquí hacemos una distinción arbitraria entre lo que se suelen llamar modos rápidos y modos lentos. Para ello definimos un nuevo valor de corte

$$\Lambda' = \Lambda/s < \Lambda, \quad (1.18)$$

con $s > 1$, y descomponemos a nuestro campo según

$$\phi_{\mathbf{k}} = \phi_{\mathbf{k}}^+ + \phi_{\mathbf{k}}^-, \quad (1.19)$$

donde los modos rápidos $\phi_{\mathbf{k}}^+$ son aquellos con momento mayor que Λ' y los modos lentos son aquellos con momento menor que Λ' , es decir

$$\phi_{\mathbf{k}}^+ = \begin{cases} \phi_{\mathbf{k}} & \text{si } \Lambda' < k < \Lambda \\ 0 & \text{en caso contrario} \end{cases} \quad \phi_{\mathbf{k}}^- = \begin{cases} \phi_{\mathbf{k}} & \text{si } k < \Lambda' \\ 0 & k > \Lambda'. \end{cases} \quad (1.20)$$

La teoría Gaussiana tiene la particularidad de que no acopla modos lentos con modos rápidos, $\mathcal{F}[\phi] = \mathcal{F}[\phi^+] + \mathcal{F}[\phi^-]$. Por otra parte, la integral de los modos rápidos puede calcularse cerradamente gracias a que es una productoria de integrales Gaussianas, y da una contribución analítica a la función de partición que descartaremos. Luego de integrar (en inglés, *integrate out*) los modos rápidos, la nueva función de partición es

$$\mathcal{Z} = \int \mathcal{D}\phi^- e^{-\mathcal{F}[\phi^-]} \quad (1.21)$$

y la nueva teoría tiene un nuevo valor de corte Λ' . Para poder comparar teorías, hacemos coincidir el valor de Λ' con Λ reescalando los momentos, definiendo

$$\mathbf{k}' = s\mathbf{k}, \quad (1.22)$$

con s definido en (1.18). En el espacio real, la transformación anterior corresponde a $\mathbf{x}' = \mathbf{x}/s$, por lo que la interpretación del reescalo en momentos puede tomarse como mirar al sistema

desde un punto más lejano, donde todas las longitudes se reducen en un factor s . Al hacer el reescalo anterior, la energía libre toma la forma

$$\mathcal{F}[\phi^-] = \frac{1}{2} \int \frac{d^d k'}{(2\pi)^d} s^{-d} \left(\left(\frac{k'}{s} \right)^2 + \mu_0^2 \right) \phi_{\mathbf{k}'/s}^- \phi_{-\mathbf{k}'/s}^-. \quad (1.23)$$

Finalmente, si tomamos el reescalo

$$\phi'_{\mathbf{k}'} = s^{-\omega} \phi_{\mathbf{k}}^- \quad (1.24)$$

para el campo, podemos tomar ω de manera tal que nuestra energía libre tome la forma inicial, donde solo cambia el valor del coeficiente del término cuadrático en ϕ . En particular, fijando $\omega = \frac{d+2}{2}$, se obtiene

$$\mathcal{Z} = \int \mathcal{D}\phi' \exp \left(\frac{1}{2} \int \frac{d^d k}{(2\pi)^d} \left(k^2 + \mu^2(s) \right) \phi'_{\mathbf{k}} \phi'_{-\mathbf{k}} \right), \quad (1.25)$$

donde definimos el parámetro renormalizado

$$\mu^2(s) = s^2 \mu_0^2. \quad (1.26)$$

La transformación (1.26) nos muestra que frente a una transformación del grupo de renormalización la masa μ_0 aumenta, salvo en dos casos particulares denominados puntos fijos de la transformación. En efecto, la masa queda invariante si partimos de $\mu_0 = \infty$ o $\mu_0 = 0$. El primer punto fijo corresponde a un estado con temperatura infinita, mientras que el segundo se denomina punto fijo Gaussiano.

Puede mostrarse que la longitud de correlación ξ en la teoría Gaussiana está relacionada con el coeficiente μ_0^2 según $\xi^2 \propto 1/\mu_0^2$. Entonces, en el punto fijo Gaussiano la longitud de correlación diverge, describiendo a un sistema crítico, mientras que en el punto fijo de temperatura infinita la longitud de correlación se anula. Por otra parte, usando (1.26), luego de una transformación del grupo de renormalización la longitud de correlación disminuye en un factor s ,

$$\xi \rightarrow \frac{\xi}{s}. \quad (1.27)$$

Es importante notar que si partimos de cualquier μ_0 finito, sucesivas transformaciones del grupo de renormalización alejan a la teoría del punto crítico, haciendo que fluya hacia el punto fijo de alta temperatura. Por lo tanto, los puntos fijos $\mu_0 = 0$ y $\mu_0 = \infty$ se denominan inestable y estable, respectivamente.

El resultado (1.27) obtenido para la teoría Gaussiana es completamente general, puesto que todas las longitudes deben achicarse en un factor s luego de una transformación del grupo de renormalización. A continuación, introducimos los conceptos de dimensión de escala, relevancia, irrelevancia y marginalidad en el contexto del grupo de renormalización.

La relevancia del escaleo

En general, supongamos tener un término en la energía libre de la forma

$$\mathcal{F}_{\mathcal{O}}[\phi] = \int d^d x g_{\mathcal{O}} \mathcal{O}(\mathbf{x}), \quad (1.28)$$

donde \mathcal{O} es alguna función de ϕ y sus derivadas, y supongamos que frente a una transformación de escala

$$\mathbf{x} \rightarrow \mathbf{x}' = \mathbf{x}/s, \quad (1.29)$$

con $s > 1$, el operador³ $\mathcal{O}(\mathbf{x})$ tiene una dimensión de escala $\Delta_{\mathcal{O}}$ bien definida, dada por

$$\mathcal{O}'(\mathbf{x}') = s^{\Delta_{\mathcal{O}}} \mathcal{O}(\mathbf{x}). \quad (1.30)$$

Luego de una transformación del grupo de renormalización, si escribimos

$$\mathcal{F}_{\mathcal{O}'}[\phi'] = \int d^d x' g'_{\mathcal{O}'} \mathcal{O}'(\mathbf{x}') = \int \frac{d^d x}{s^d} g'_{\mathcal{O}'} s^{\Delta_{\mathcal{O}}} \mathcal{O}(\mathbf{x}), \quad (1.31)$$

podemos identificar

$$g'_{\mathcal{O}'} = g_{\mathcal{O}} s^{d-\Delta_{\mathcal{O}}}, \quad (1.32)$$

donde hay tres casos posibles para el comportamiento de $g_{\mathcal{O}}$ luego de una transformación del grupo de renormalización. Si $\Delta_{\mathcal{O}} < d$, \mathcal{O} se dice relevante, pues $g_{\mathcal{O}}$ se incrementa. Si $\Delta_{\mathcal{O}} > d$, \mathcal{O} se dice irrelevante, pues $g_{\mathcal{O}}$ se achica. Si $\Delta_{\mathcal{O}} = d$, \mathcal{O} se dice marginal, pues $g_{\mathcal{O}}$ no cambia.

Utilizando la teoría Gaussiana podemos estudiar la relevancia de términos adicionales en la teoría frente a una transformación del grupo de renormalización. Consideremos la presencia de un término de fuente en (1.14) de la forma

$$\mathcal{F}_J = -J \int d^d x \phi(x) = -J \int d^d x \int \frac{d^d k}{(2\pi)^d} e^{ikx} \phi_{\mathbf{k}} = -J \phi_{\mathbf{0}}, \quad (1.33)$$

donde $\phi_{\mathbf{0}} = \phi_{\mathbf{k}=\mathbf{0}}$. Si la teoría describe un ferromagneto de Ising, J corresponde al producto βh , con β la inversa de la temperatura y h el campo magnético aplicado. Si integramos los grados de libertad microscópicos de la teoría y luego reescalamos los momentos según (1.22) y el campo según (1.24), obtenemos que

$$\mathcal{F}_J = -J s^{\omega} \phi_{\mathbf{0}} = -J s^{1+d/2} \phi_{\mathbf{0}}, \quad (1.34)$$

donde vemos que el campo magnético es una variable relevante en cualquier dimensión, puesto que luego de una transformación del grupo de renormalización el nuevo coeficiente $J' = J s^{1+d/2}$ es mayor que el anterior. Por otra parte, podemos considerar la presencia de un término de la forma

$$\mathcal{F}_L = \frac{L}{2} \int d^d x (\nabla^2 \phi)^2 = \frac{L}{2} \int \frac{d^d k}{(2\pi)^d} k^4 \phi_{\mathbf{k}} \phi_{-\mathbf{k}}. \quad (1.35)$$

En este caso, luego de una transformación del grupo de renormalización obtenemos

$$\mathcal{F}_L = \frac{L s^{-d-4+2\omega}}{2} \int d^d x (\nabla^2 \phi)^2 = \frac{L s^{-2}}{2} \int \frac{d^d k}{(2\pi)^d} k^4 \phi_{\mathbf{k}} \phi_{-\mathbf{k}}, \quad (1.36)$$

por lo que \mathcal{F}_L es un término irrelevante en el sentido del grupo de renormalización en cualquier

³Nuevamente, terminología tomada de física de partículas.

dimensión espacial.

1.2 Estados coherentes de espín

Para construir nuestras teorías de campos efectivas en sistemas de espines usamos los denominados estados coherentes de espín, que serán definidos en breve. Sin embargo, para hacerlo siguiendo el libro de Auerbach[1] necesitamos introducir primero los bosones de Schwinger, que definen un mapeo entre sistemas de espines y sistemas bosónicos.

1.2.1 Bosones de Schwinger

Existen varias representaciones de operadores de espín en términos de operadores de creación y destrucción de partículas, tanto con estadística fermiónica como bosónica. Dentro del primer grupo, la transformación más popular es la de Jordan-Wigner ([5] capítulo 6), generalmente aplicada a sistemas de espín en una dimensión. En el segundo grupo las dos transformaciones más populares son la Holstein-Primakov([1] capítulo 7), muchas veces usada para cálculos semi-clásicos de ondas de espín y la de Schwinger, típicamente empleada en cálculos de campo medio y algunos esquemas de teorías de campos, como en nuestro caso presente. Dado un operador de espín \mathbf{S} , los bosones de Schwinger están definidos por la siguiente ecuación

$$\begin{aligned} S^+ &= a^\dagger b \\ S^- &= b^\dagger a \\ S^z &= \frac{a^\dagger a - b^\dagger b}{2}, \end{aligned} \tag{1.37}$$

donde los operadores a y b satisfacen el álgebra bosónica, y entre ellos cumplen $[a, b] = 0$. En esta representación, los autoestados de S^z son de la forma

$$|n_a, n_b\rangle = \frac{(a^\dagger)^{n_a}}{\sqrt{(n_a)!}} \frac{(b^\dagger)^{n_b}}{\sqrt{(n_b)!}} |0\rangle, \tag{1.38}$$

donde de los infinitos estados posibles en la ecuación anterior, sólo $2S + 1$ son físicos, con S el número cuántico de espín. En esta base, se tiene

$$\begin{aligned} S^+ |n_a, n_b\rangle &= \sqrt{(n_a + 1)n_b} |n_a + 1, n_b - 1\rangle \\ S^- |n_a, n_b\rangle &= \sqrt{n_a(n_b + 1)} |n_a - 1, n_b + 1\rangle, \end{aligned} \tag{1.39}$$

donde puede observarse que en el subespacio físico

$$n_a + n_b = 2S, \tag{1.40}$$

ya que como máximo puede aplicarse el operador S^+ (o el S^-) $2S$ veces. Además, según (1.37) y (1.38), se tiene

$$n_a - n_b = 2m. \tag{1.41}$$

Así resulta $n_a = S + m$ y $n_b = S - m$. Por lo tanto, podemos escribir los estados $|S, m\rangle$ en la nueva representación según

$$|S, m\rangle = \frac{(a^\dagger)^{S+m}}{\sqrt{(S+m)!}} \frac{(b^\dagger)^{S-m}}{\sqrt{(S-m)!}} |0\rangle. \quad (1.42)$$

Esta ecuación es importante en teoría de grupos y momento angular, pues muestra que un objeto complejo con momento angular S esta compuesto por $2S$ momentos angulares $1/2$. Como ejemplo, para espines $1/2$ se los estados se escriben

$$\begin{aligned} |+\rangle &= a^\dagger |0\rangle \\ |-\rangle &= b^\dagger |0\rangle. \end{aligned} \quad (1.43)$$

Rotaciones sobre los bosones de Schwinger

Definimos el operador unitario de rotación como

$$R(\theta, \phi) = e^{i\phi S^z} e^{i\theta S^y}, \quad (1.44)$$

el cual satisface

$$R(\theta, \phi) S^z R^\dagger(\theta, \phi) = \mathbf{S} \cdot \boldsymbol{\Omega}, \quad (1.45)$$

con $\boldsymbol{\Omega} = (\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta)$. Los bosones de Schwinger transforman ante rotaciones como vectores de $SU(2)$:

$$\begin{pmatrix} a'^\dagger \\ b'^\dagger \end{pmatrix} = \begin{pmatrix} u & v \\ -v^* & u^* \end{pmatrix} \begin{pmatrix} a^\dagger \\ b^\dagger \end{pmatrix}, \quad (1.46)$$

donde $|u|^2 + |v|^2 = 1$. Para conocer u y v en términos de θ y ϕ , usamos la ecuación (1.45)

$$\mathbf{S} \cdot \boldsymbol{\Omega} = R(\theta, \phi) S^z R^\dagger(\theta, \phi) = R(\theta, \phi) \frac{a^\dagger a - b^\dagger b}{2} R^\dagger(\theta, \phi) = \frac{a'^\dagger a' - b'^\dagger b'}{2}. \quad (1.47)$$

Luego de un poco de álgebra usando (1.37), se llega a la solución

$$\begin{cases} u = \cos \frac{\theta}{2} e^{i\frac{\phi}{2}} \\ v = \sin \frac{\theta}{2} e^{-i\frac{\phi}{2}}. \end{cases} \quad (1.48)$$

1.2.2 Definición y propiedades

Por definición, un estado coherente de espín $|\boldsymbol{\Omega}\rangle$ es un estado con máxima proyección de espín a lo largo de una dirección dada por el vector unitario

$$\boldsymbol{\Omega} = \begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix}, \quad (1.49)$$

donde θ y ϕ son el ángulo polar y azimutal, respectivamente. Entonces, $|\boldsymbol{\Omega}\rangle$ es solución de la ecuación

$$(\mathbf{S} \cdot \boldsymbol{\Omega}) |\boldsymbol{\Omega}\rangle = S |\boldsymbol{\Omega}\rangle \quad (1.50)$$

Para construirlo, puede tomarse el vector de máxima proyección de espín a lo largo del eje z , $|S, S\rangle$, y rotarlo hacia la dirección $\boldsymbol{\Omega}$

$$|\boldsymbol{\Omega}\rangle \equiv e^{i\phi S^z} e^{i\theta S^y} |S, S\rangle \equiv R(\theta, \phi) |S, S\rangle. \quad (1.51)$$

Entonces, la ecuación (1.50) se satisface usando (1.45) y (1.51). Para obtener una expansión del estado coherente de espín $|\boldsymbol{\Omega}\rangle$ en términos de la base estándar $|S, m\rangle$, utilizamos los bosones de Schwinger (Ec. (1.42)) y sus propiedades de transformación frente a rotaciones (Ec. (1.46)) debido a que de esta forma el resultado es válido para todo valor de espín S . Obtenemos entonces

$$|\boldsymbol{\Omega}\rangle = R(\theta, \phi) \left[\frac{(a^\dagger)^{2S}}{\sqrt{(2S)!}} \right] |0\rangle = \frac{1}{\sqrt{(2S)!}} (ua^\dagger + vb^\dagger)^{2S} |0\rangle, \quad (1.52)$$

donde usamos que el vacío es invariante ante rotaciones. Utilizando el binomio de Newton $(A + B)^N = \sum_{n=0}^N \binom{N}{n} A^n B^{N-n}$ (ya que $[a, b] = 0$) y la ecuación (1.48) se sigue

$$|\boldsymbol{\Omega}\rangle = \sqrt{(2S)!} \sum_{m=-S}^S \frac{(\cos \frac{\theta}{2})^{S+m_S} (\sin \frac{\theta}{2})^{S-m_S}}{\sqrt{(S+m_S)!(S-m_S)!}} e^{im_S\phi} |S, m_S\rangle. \quad (1.53)$$

A partir de la formula anterior, calculamos la superposición (*overlap*) entre estados coherentes de espín

$$\begin{aligned} \langle \boldsymbol{\Omega}' | \boldsymbol{\Omega} \rangle &= (2S)! \sum_{m_S=-S}^S \frac{\left(\cos \frac{\theta'}{2} \cos \frac{\theta}{2} e^{i\frac{\phi-\phi'}{2}} \right)^{S+m_S} \left(\sin \frac{\theta'}{2} \sin \frac{\theta}{2} e^{-i\frac{\phi-\phi'}{2}} \right)^{S-m_S}}{(S+m_S)!(S-m_S)!} \\ &= \sum_{n=0}^{2S} \frac{\left(\cos \frac{\theta'}{2} \cos \frac{\theta}{2} e^{i\frac{\phi-\phi'}{2}} \right)^n \left(\sin \frac{\theta'}{2} \sin \frac{\theta}{2} e^{-i\frac{\phi-\phi'}{2}} \right)^{2S-n} (2S)!}{n!(2S-n)!} \\ &= \left(\cos \frac{\theta'}{2} \cos \frac{\theta}{2} e^{i\frac{\phi-\phi'}{2}} + \sin \frac{\theta'}{2} \sin \frac{\theta}{2} e^{-i\frac{\phi-\phi'}{2}} \right)^{2S} \\ &= \left(\frac{1 + \boldsymbol{\Omega}' \cdot \boldsymbol{\Omega}}{2} \right)^S \left(e^{-i\Psi} \right)^{2S}, \end{aligned} \quad (1.54)$$

donde en el último paso escribimos la descomposición polar del complejo y llamamos Ψ a su argumento

$$\Psi = \arctan \left(\cos \left(\frac{1}{2} (\theta' + \theta) \right) \sec \left(\frac{1}{2} (\theta' - \theta) \right) \tan \left(\frac{1}{2} (\phi' - \phi) \right) \right). \quad (1.55)$$

La ecuación (1.54) muestra que los estados coherentes no son ortogonales entre sí. A continuación y para concluir la introducción a los estados coherentes de espín calculamos la

resolución de la identidad en esta base, que resulta

$$\mathbb{1} = \int \frac{2S+1}{4\pi} d\Omega |\Omega\rangle \langle \Omega|, \quad (1.56)$$

donde la medida de integración,

$$\frac{2S+1}{4\pi} d\Omega = \frac{2S+1}{4\pi} \sin\theta d\theta d\phi, \quad (1.57)$$

se denomina medida de Haar del grupo $SU(2)$. Para probar (1.56), usamos la expresión (1.53) para escribir

$$\begin{aligned} \int \frac{2S+1}{4\pi} d\Omega |\Omega\rangle \langle \Omega| &= \\ \frac{2S+1}{2} (2S)! \sum_{m_S=-S}^S \frac{|S, m_S\rangle \langle S, m_S|}{(S+m_S)!(S-m_S)!} \int_0^\pi d\theta \sin\theta \cos(\theta/2)^{2(S+m_S)} \sin(\theta/2)^{2(S-m_S)}, \end{aligned} \quad (1.58)$$

donde para resolver la integral en $d\phi$ hay que notar que la resta (también la suma) de dos valores cualesquiera de m_S es un número entero, tanto para S entero como semi entero. Para resolver la integral que queda, la reescribimos de la siguiente manera

$$\int_0^\pi d\theta \sin\theta \cos(\theta/2)^{2(S+m_S)} \sin(\theta/2)^{2(S-m_S)} = \int_{-1}^1 d(\cos(\theta)) \left(\frac{1+\cos\theta}{2}\right)^{S+m_S} \left(\frac{1-\cos\theta}{2}\right)^{S-m_S}, \quad (1.59)$$

y finalmente el cambio de variables $t = (1+\cos\theta)/2$ permite usar una propiedad de las funciones Gamma conocida como función Beta,

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \quad (1.60)$$

donde $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$, $Re(z) > 0$. Nuevamente, la observación de que $S \pm m_S \in \mathbb{Z}$ permite reemplazar funciones Γ por factoriales, probando la ecuación (1.56). Por último, observamos que la resolución de la identidad permite calcular la traza de un operador \mathcal{O} en la base de estados coherentes de espín. Si $\{|n\rangle\}$ es una base ortonormal cualquiera, entonces

$$\begin{aligned} Tr \mathcal{O} &= \sum_n \langle n | \mathcal{O} | n \rangle \\ &= \left(\frac{2S+1}{4\pi}\right)^2 \int d\Omega d\Omega' \sum_n \langle n | \Omega \rangle \langle \Omega | \mathcal{O} | \Omega' \rangle \langle \Omega' | n \rangle \\ &= \frac{2S+1}{4\pi} d\Omega \langle \Omega | \mathcal{O} | \Omega \rangle. \end{aligned} \quad (1.61)$$

1.3 Integrales funcionales en sistemas de espines

Escribimos la función de partición de un sistema como

$$\mathcal{Z}[j] = Tr T_\tau \left(\exp \left[- \int_0^\beta d\tau \mathcal{H}(\tau) \right] \right) = \lim_{N_\epsilon \rightarrow \infty} Tr T_\tau \prod_{n=0}^{N_\epsilon-1} [1 - \epsilon \mathcal{H}(\tau_n)], \quad (1.62)$$

donde β es la inversa de la temperatura, T_τ es el operador de ordenamiento temporal, N_ϵ es el número de intervalos en los que se particiona el rango $[0, \beta]$ (se toma N_ϵ ‘grande’), $\epsilon = \beta/N_\epsilon$ es la longitud de cada intervalo y $\tau_n = n\epsilon$ es el tiempo imaginario discreto. El Hamiltoniano que caracteriza al sistema se escribe

$$\mathcal{H}(\tau) = \mathcal{H} - \sum_\alpha j^\alpha(\tau) S^\alpha, \quad (1.63)$$

incluyendo corrientes auxiliares $j(\tau)$. Luego, para calcular la traza elegimos la base de estados coherentes de espín, e insertamos $N_\epsilon - 1$ resoluciones de la identidad en la partición del intervalo de tiempo imaginario $[0, \beta]$, escribiendo

$$\begin{aligned} \mathcal{Z}[j] &= \left(\frac{2S+1}{4\pi} \right) \int d\Omega \langle \Omega | (1 - \mathcal{H}(\tau_{N_\epsilon-1})) \dots (1 - \mathcal{H}(\tau_0)) | \Omega \rangle \\ &= \left(\frac{2S+1}{4\pi} \right)^{N_\epsilon} \int \prod_{n=1}^{N_\epsilon} \left(d\Omega_{n-1} \langle \Omega_n | 1 - \epsilon \mathcal{H}(\tau_{n-1}) | \Omega_{n-1} \rangle \right), \end{aligned} \quad (1.64)$$

donde llamamos $|\Omega_n\rangle = |\Omega(\tau_n)\rangle$ y definimos

$$|\Omega(\beta)\rangle = |\Omega(0)\rangle, \quad (1.65)$$

notando que hay una sola integral para ambos tiempos. Luego, calculamos cada valor de expectación a primer orden en ϵ

$$\begin{aligned} \langle \Omega_n | e^{-\epsilon \mathcal{H}_{n-1}} | \Omega_{n-1} \rangle &= \langle \Omega_n | \Omega_{n-1} \rangle + (-\epsilon) \langle \Omega_n | \mathcal{H}_{n-1} | \Omega_{n-1} \rangle \\ &= \langle \Omega_n | \Omega_{n-1} \rangle + (-\epsilon) \langle \Omega_n | \mathcal{H}_{n-1} | \Omega_{n-1} \rangle \frac{\langle \Omega_n | \Omega_{n-1} \rangle}{\langle \Omega_n | \Omega_{n-1} \rangle} \\ &= \langle \Omega_n | \Omega_{n-1} \rangle \left(1 - \epsilon \frac{\langle \Omega_n | \mathcal{H}_{n-1} | \Omega_{n-1} \rangle}{\langle \Omega_n | \Omega_{n-1} \rangle} \right) \\ &= \langle \Omega_n | \Omega_{n-1} \rangle \left(1 - \epsilon \langle \Omega_n | \mathcal{H}_n | \Omega_n \rangle \right) \\ &= \langle \Omega_n | \Omega_{n-1} \rangle e^{-\epsilon \langle \Omega_n | \mathcal{H}_n | \Omega_n \rangle}, \end{aligned} \quad (1.66)$$

El Hamiltoniano clásico $H[\Omega(\tau)]$ se define como

$$H[\Omega(\tau)] = \langle \Omega_n | \mathcal{H}_n | \Omega_n \rangle. \quad (1.67)$$

Usando las ecuaciones (1.54) y (1.55) a primer orden en ϵ se obtiene

$$\langle \mathbf{\Omega}_n | \mathbf{\Omega}_{n-1} \rangle \approx e^{-iS\epsilon\dot{\phi}\cos\theta}. \quad (1.68)$$

Luego, tomando el límite $N_\epsilon \rightarrow \infty$, definimos la medida de la integral funcional como

$$\int \mathcal{D}\mathbf{\Omega} \equiv \lim_{N_\epsilon \rightarrow \infty} \left(\frac{2S+1}{4\pi} \right)^{N_\epsilon} d\mathbf{\Omega}_1 \dots d\mathbf{\Omega}_{N_\epsilon}, \quad (1.69)$$

y obtenemos para la función de partición la expresión formal

$$\begin{aligned} \mathcal{Z} &= \oint \mathcal{D}\mathbf{\Omega}(\tau) \exp(-\mathcal{S}[\mathbf{\Omega}]) \\ \mathcal{S}[\mathbf{\Omega}] &= -iS \int_0^\beta d\tau \dot{\phi} \cos\theta + \int_0^\beta d\tau H[\mathbf{\Omega}(t)], \end{aligned} \quad (1.70)$$

donde el símbolo \oint enfatiza la condición periódica de contorno (1.65). El primer término en la acción anterior se conoce como fase de Berry, una cantidad geométrica que surge aquí a partir de la no ortogonalidad de la base de estados coherentes de espín, mientras que el segundo término es la acción clásica del sistema, \mathcal{S}_{cl} . Es importante notar que el primer término de la acción anterior puede escribirse

$$\begin{aligned} \mathcal{S}_{BP} &= -S \int_0^\beta d\tau \dot{\phi} \cos\theta \\ &= -S \oint_{\phi_0}^{\phi_0} d\phi \cos\theta(\phi), \end{aligned} \quad (1.71)$$

mostrando que es una cantidad geométrica; depende de la trayectoria de $\mathbf{\Omega}(\tau)$ sobre la esfera unitaria. En particular, la ecuación (1.71) muestra que la fase de Berry representa S veces el área de la porción de esfera unitaria que contiene al polo norte encerrada por la trayectoria $\mathbf{\Omega}(\tau)$ del espín. Para verlo, podemos considerar un diferencial de área dA generado por el desplazamiento $d\phi$ del vector $\mathbf{\Omega}(\tau)$, que se escribe

$$dA(\tau) = d\phi \int_0^{\theta(\tau)} \sin\theta(\tau) d\theta(\tau) = d\phi(1 - \cos\theta(\tau)). \quad (1.72)$$

Luego, el área encerrada por la trayectoria resulta

$$A = \oint_{\phi_0}^{\phi_0} d\phi(1 - \cos\theta(\phi)). \quad (1.73)$$

Usando la condición (1.65) vemos que ambos resultados coinciden en ausencia de vorticidad.

1.4 Teoría de campos efectiva para la cadena de espín antiferromagnética

Consideramos el Hamiltoniano de Heisenberg en presencia de un término de anisotropía y campo magnético para la cadena de espín antiferromagnética

$$\mathcal{H} = J \sum_j \mathbf{S}_j \cdot \mathbf{S}_{j+1} + D \sum_j (S_j^z)^2 - h \sum_j S_j^z, \quad (1.74)$$

con $D > 0$, $J > 0$ y $h > 0$. Consideremos el modelo clásico que se obtiene al reemplazar cada operador de espín por un vector de \mathbb{R}^3 de módulo S , donde S es el número cuántico de espín. Parametrizamos entonces a cada vector con dos ángulos, polar y azimutal, según

$$\mathbf{S}_j = S \tilde{\mathbf{n}}_j = S(\cos \phi_j \sin \theta_j, \sin \phi_j \sin \theta_j, \cos \theta_j). \quad (1.75)$$

Luego, el Hamiltoniano toma la forma

$$\begin{aligned} \mathcal{H} = JS^2 \sum_j \left\{ \sin \theta_j \sin \theta_{j+1} \cos(\phi_{j+1} - \phi_j) + \cos \theta_j \cos \theta_{j+1} \right\} \\ + DS^2 \sum_j \cos^2 \theta_j - hS \sum_j \cos \theta_j. \end{aligned} \quad (1.76)$$

Al reemplazar los operadores de espín por vectores parametrizados por ángulos continuos nos referimos a que estamos tomando la versión clásica del Hamiltoniano del sistema, o bien, el Hamiltoniano clásico del sistema. Para minimizar a la energía como función de los ángulos, debemos resolver $\frac{\partial}{\partial \phi_j} H = \frac{\partial}{\partial \theta_j} H = 0$, $\forall j$, lo que lleva a una solución de la forma

$$\begin{aligned} \phi_{j+1} - \phi_j &= \pi \\ \cos \theta_0 &= \frac{h}{2S(2J + D)}, \end{aligned} \quad (1.77)$$

donde hemos usado la suposición $\theta_j = \theta_{j+1}$, $\forall j$, para condiciones de contorno periódicas. El estado fundamental clásico corresponde a una configuración con las componentes xy de los vectores alineadas de manera antiparalela. A una configuración como esta donde θ_0 no es necesariamente 0 o π se la denomina inclinada (*canted*, en inglés). Las fluctuaciones cuánticas se agregan sobre el estado fundamental clásico (1.77) introduciendo las variables $\phi(x_j)$ y $\delta\theta(x_j)$,

$$\begin{aligned} \phi_j &\rightarrow \frac{\pi}{a} x_j + \phi(x_j) \\ \theta_j &\rightarrow \theta_0 + \delta\theta(x_j), \end{aligned} \quad (1.78)$$

donde a es la distancia entre espines en la cadena y $x_j = ja$. En el límite continuo, las variables

ϕ_j y $\delta\theta(x_j)$ son campos $\phi(x)$ y $\delta\theta(x)$. Luego, los espines clásicos toman la forma

$$\begin{aligned} S^\pm &= S \sin(\theta_0 + \delta\theta(x_j)) (-1)^j e^{\pm i\phi(x_j)} \\ S^z &= S \cos(\theta_0 + \delta\theta(x_j)). \end{aligned} \quad (1.79)$$

Como $\phi(x_j)$ y $\delta\theta(x_j)$ son fluctuaciones, en adelante trabajaremos con una expansión a orden cuadrática en ambos campos. Por otra parte, como los operadores cuánticos tienen que cumplir el álgebra de $SU(2)$ $[S^z, S^\pm] = \pm S^\pm$, en el enfoque semiclásico necesitamos campos conjugados ϕ y $a\Pi$, tales que

$$i \left\{ S^z, S^\pm \right\}_{\phi, a\Pi} \equiv i \left(\frac{\partial S^z}{\partial \phi} \frac{\partial S^\pm}{\partial a\Pi} - \frac{\partial S^z}{\partial a\Pi} \frac{\partial S^\pm}{\partial \phi} \right) = \pm S^\pm, \quad (1.80)$$

donde estamos reemplazando los conmutadores cuánticos por corchetes de Poisson clásicos. Este no es el caso de $\delta\theta(x_j)$, para el cual $i \left\{ S^z, S^\pm \right\}_{\phi, \delta\theta} = \pm S(\sin(\theta_0 + \delta\theta) \cos \theta_0) S^\pm$. La variable conjugada de ϕ es

$$a\Pi(x_j) = -S \left(\delta\theta(x_j) \sin \theta_0 + \frac{1}{2} (\delta\theta(x_j))^2 \cos \theta_0 \right). \quad (1.81)$$

Los operadores de espín en términos de ϕ y $a\Pi$ son

$$\begin{aligned} S_j^z &= S \cos \theta_0 + a\Pi \\ S_j^\pm &= (-1)^j e^{\pm i\phi(x_j)} \left[S \sin \theta_0 - \frac{a\Pi(x_j)}{\tan \theta_0} - \frac{1}{2} \frac{S^2}{S^2 - m^2} \frac{(a\Pi)^2}{S \sin \theta_0} \right], \end{aligned} \quad (1.82)$$

donde $m = S \cos \theta_0$ es la magnetización clásica por sitio. Usando (1.82) puede verse que $i \left\{ S^+, S^- \right\}_{\phi, a\Pi} = 2S^z$ a primer orden en $a\Pi$. La acción efectiva Gaussiana o cuadrática del sistema es

$$\mathcal{S} = \int dx d\tau \left\{ \frac{aJ}{2} (S^2 - m^2) (\partial_x \phi)^2 + \frac{(2J + D)}{a} (a\Pi)^2 + i \left(\frac{S - m}{a} \right) \partial_\tau \phi - i (\partial_\tau \phi) \Pi \right\}. \quad (1.83)$$

El primer término en la acción efectiva es el término cinético, que penaliza las fluctuaciones espaciales de ϕ , y su coeficiente se denomina rigidez (*stiffness*) del campo. El segundo es un término de masa del campo $a\Pi$, terminología tomada de las teorías cuánticas de campos, que penaliza la magnitud del campo $a\Pi$, y el coeficiente que acompaña a $(a\Pi)^2$ se denomina la masa del campo. Los últimos dos términos corresponden a la parte topológica de la acción, que surge a causa de la no ortonormalidad de la base de estados coherentes de espín usada en este formalismo. Remarcamos aquí que la ausencia de un término de masa para el campo ϕ es consecuencia de la simetría $U(1)$ del Hamiltoniano, correspondiente a las rotaciones alrededor de la dirección del campo magnético aplicado. La acción efectiva está conectada con la función de partición del sistema según

$$\mathcal{Z} = \int \mathcal{D}\phi \mathcal{D}a\Pi \exp(-\mathcal{S}). \quad (1.84)$$

Los términos que dependen de Π en (1.83) pueden escribirse completando cuadrados de la

siguiente manera

$$\mathcal{S}_{\Pi} = \int dx d\tau a(2J + D) \left(\Pi + \frac{i\partial_{\tau}\phi}{2a(2J + D)} \right)^2 + \frac{1}{4a(2J + D)} (\partial_{\tau}\phi)^2. \quad (1.85)$$

Luego de la integración gaussiana del campo masivo Π la acción toma la forma

$$\mathcal{S} = \int dx d\tau \left\{ \frac{K_{\tau}}{2} (\partial_{\tau}\phi)^2 + \frac{K_x}{2} (\partial_x\phi)^2 + i \left(\frac{S - m}{a} \right) \partial_{\tau}\phi \right\}, \quad (1.86)$$

donde $K_{\tau} = \frac{1}{2a(2J+D)}$ y $K_x = Ja(S^2 - m^2)$.

Si $S - m$ es un entero no hay fase de Berry, porque la fase del término geométrico se vuelve trivial. Para verlo, hay que observar en (1.86) que el último término es una derivada total y usar la condición (1.65), que para ϕ corresponde a $\phi(\beta) = \phi(0) + 2q\pi$, con $q \in \mathbb{Z}$ la vorticidad asociada. Luego, sin fase de Berry la acción del sistema a temperatura cero es equivalente a un modelo XY en dos dimensiones, el cual puede estar en dos fases separadas por una transición de fase de Berezinskii-Kosterlitz-Thouless, dependiendo del valor de los coeficientes de rigidez K_x y K_{τ} . La condición

$$S - m \in \mathbb{Z}, \quad (1.87)$$

para la posible formación de un plateau de magnetización se conoce como criterio OYA[6], por las siglas de sus autores, Oshikawa, Yamanaka y Affleck. Para sistemas de espines con una celda unidad con N sitios de espín S , el criterio se modifica a $N(S - m) \in \mathbb{Z}$. En el marco de la teoría de campos en términos de estados coherentes de espín esta generalización puede obtenerse fácilmente[7].

En el sistema clásico los coeficientes de rigidez son dependientes de la temperatura, mientras que en el sistema cuántico dependen de los parámetros J , D , a , h , etc. Si los vórtices proliferan, entonces ϕ se desordena y adquiere una masa efectiva. Por lo tanto, su variable conjugada Π se localiza. Como Π es un campo con masa, típicamente se localiza a un valor nulo. Este proceso corresponde a la formación de un plateau de magnetización en el sistema.

Si en cambio hay una fase de Berry no trivial, es decir, $S - m$ no es un entero, entonces los vórtices generan contribuciones a la función de partición que interfieren destructivamente[8], lo que impide la condensación de defectos topológicos. Cabe destacar en este punto la posibilidad de que $S - m$ sea un número racional. En este caso los vórtices deben estar múltiplemente cuantizados para poder condensar. La bibliografía [8] muestra que estos casos particulares pueden dar origen a diferentes fases exóticas en antiferromagnetos cuánticos.

Antes de continuar trabajando la acción del sistema, hacemos un comentario sobre los campos conjugados y las ecuaciones de movimiento clásicas de la teoría. El momento canónico conjugado de ϕ está formalmente definido por

$$\frac{\partial \mathcal{L}}{\partial(\partial_{\tau}\phi)} = -i\Pi + i \frac{S - m}{a}, \quad (1.88)$$

usando la ecuación (1.83).

Si Π es integrado como en (1.86), el nuevo momento conjugado de ϕ es

$$\frac{\partial \mathcal{L}}{\partial(\partial_\tau \phi)} = K_\tau(\partial_\tau \phi) + i \frac{S - m}{a}. \quad (1.89)$$

Finalmente, la ecuación de Euler-Lagrange de la teoría (1.86) resulta

$$\partial_\tau^2 \phi_v + \frac{K_x}{K_\tau} \partial_x^2 \phi_v = 0. \quad (1.90)$$

Volviendo al tiempo real $t = i\tau$ tenemos que $e^{-\int d\tau \mathcal{L}} = e^{i\int dt \mathcal{L}} = e^{i\mathcal{S}}$. En este caso, vemos que la ecuación anterior corresponde a una ecuación de onda donde la velocidad de propagación de las excitaciones es $v = \sqrt{K_x/K_\tau}$. Por lo tanto, si el coeficiente de rigidez K_x se anula, las excitaciones no se propagan por la cadena. Más adelante veremos que este escenario se realiza en la teoría efectiva de antiferromagnetos que tienen como estado fundamental un cristal de magnones.

1.4.1 Transformación de dualidad

Aquí aplicamos una transformación de dualidad, descomponiendo a ϕ en dos campos con y sin vorticidad, respectivamente.

$$\phi = \phi_v + \phi_t, \quad \begin{cases} (\partial_\mu \partial_\nu - \partial_\nu \partial_\mu) \phi_v \neq 0 \\ (\partial_\mu \partial_\nu - \partial_\nu \partial_\mu) \phi_t = 0, \end{cases} \quad (1.91)$$

e introducimos un campo auxiliar de Hubbard-Stratonovich $\mathbf{J}(\tau, x) = (J_\tau(\tau, x), J_x(\tau, x))$ escribiendo en cada punto (τ, x) las llamadas transformaciones de Hubbard-Stratonovich

$$\begin{cases} \exp\left(-\frac{K_\tau}{2}(\partial_\tau \phi)^2\right) = \sqrt{2\pi K_\tau} \int_{-\infty}^{\infty} dJ_\tau \exp\left(-\frac{J_\tau^2}{2K_\tau} \pm iJ_\tau(\partial_\tau \phi)\right) \\ \exp\left(-\frac{K_x}{2}(\partial_x \phi)^2\right) = \sqrt{2\pi K_x} \int_{-\infty}^{\infty} dJ_x \exp\left(-\frac{J_x^2}{2K_x} \pm iJ_x(\partial_x \phi)\right). \end{cases} \quad (1.92)$$

De esta manera, la función de partición resulta

$$\mathcal{Z} = \oint \mathcal{D}\phi_t \mathcal{D}\phi_v \mathcal{D}J_\tau \mathcal{D}J_x \exp(-\mathcal{S}). \quad (1.93)$$

La nueva acción es

$$\mathcal{S} = \int dx d\tau \left\{ i \left(J_\tau + \left(\frac{S - m}{a} \right) \right) \partial_\tau \phi_v + i J_x \partial_x \phi_v + \frac{J_\tau^2}{2K_\tau} + \frac{J_x^2}{2K_x} - i \left[\partial_\tau \left(J_\tau + \left(\frac{S - m}{a} \right) \right) + \partial_x J_x \right] \phi_t \right\}, \quad (1.94)$$

donde hemos integrado por partes los términos con ϕ_t . Aquí, el campo sin vorticidad ϕ_t puede integrarse y usando la identidad $\int dx e^{ikx} = 2\pi \delta(k)$ en cada punto se obtiene la restricción

$$\partial_\mu \tilde{J}_\mu = 0, \quad (1.95)$$

donde $\tilde{J} = J_\mu + \delta_{\mu,0} \left(\frac{S - m}{a} \right)$. Luego, la acción se escribe

$$\mathcal{S} = \int dx d\tau \left\{ i\tilde{J}_\mu \partial_\mu \phi_v + \frac{1}{2K_\tau} \left(\tilde{J}_\tau - \left(\frac{S-m}{a} \right) \right)^2 + \frac{\tilde{J}_x^2}{2K_x} \right\}. \quad (1.96)$$

La restricción (1.95) se resuelve en una dimensión escribiendo

$$\tilde{J}_\mu = \epsilon_{\mu\nu} \partial_\nu \chi, \quad (1.97)$$

donde introducimos el campo χ que no tiene vorticidad. Integrando por partes el primer término en la acción y usando la sustitución $\chi \rightarrow \tilde{\chi} + \left(\frac{S-m}{a} \right) x$, obtenemos

$$\mathcal{S} = \int dx d\tau \left\{ \frac{1}{2K_\tau} (\partial_x \tilde{\chi})^2 + \frac{1}{2K_x} (\partial_\tau \tilde{\chi})^2 + iB \left(\tilde{\chi} + \left(\frac{S-m}{a} \right) x \right) \right\}, \quad (1.98)$$

donde $B = \epsilon_{\mu\nu} \partial_\mu \partial_\nu \phi_v = \sum_j 2\pi q_j \delta(\tau - \tilde{\tau}_j) \delta(x - \tilde{x}_j)$, siendo $\tilde{\tau}_j$ y \tilde{x}_j las coordenadas de tiempo y espacio del vórtice j -ésimo, y q_j su vorticidad. Usando B , la función de partición puede escribirse

$$\mathcal{Z} = \int \mathcal{D}\tilde{\chi} \sum_{\text{vortex conf}} e^{-\mathcal{S}_{XY}} \exp \left(2\pi i \int dx d\tau \left(\sum_j q_j \delta(\tau - \tilde{\tau}_j) \delta(x - \tilde{x}_j) \right) \left(\tilde{\chi} + \frac{S-m}{a} x \right) \right) \quad (1.99)$$

donde \mathcal{S}_{XY} está definida por los dos primeros términos en (1.98) y $\sum_{\text{vortex conf}} = \sum_{q_j} \int d\tilde{x}_j d\tilde{\tau}_j$, lo que significa que debemos sumar sobre todas las posibles configuraciones de vórtices. Luego, en una aproximación de baja energía consideramos el escenario donde solo puede haber como máximo dos vórtices con vorticidades opuestas $q = \pm 1$ o ningún vórtice, $q = 0$, obteniendo

$$\mathcal{Z} = \int \mathcal{D}\tilde{\chi} e^{-\mathcal{S}_{XY}} \left[1 + \int d\tilde{x}_j d\tilde{\tau}_j \left\{ e^{2\pi i \left(\tilde{\chi}_j + \frac{S-m}{a} \tilde{x}_j \right)} + e^{-2\pi i \left(\tilde{\chi}_j + \frac{S-m}{a} \tilde{x}_j \right)} \right\} \right]. \quad (1.100)$$

Luego, reexponenciando, obtenemos

$$\mathcal{Z} = \int \mathcal{D}\tilde{\chi} e^{-\mathcal{S}_{XY}} e^{2 \int dx d\tau \cos \left(\tilde{\chi} + \frac{S-m}{a} x \right)}. \quad (1.101)$$

Esta es una acción del tipo Sine-Gordon, y corresponde a la teoría dual de la teoría original (1.86). Es muy interesante notar que los coeficientes de rigidez de la teoría dual son los inversos a los coeficientes de rigidez de la teoría original. En ausencia de fase de Berry el coseno está conmensurado y tiende a fijar a $\tilde{\chi}$ en alguno de sus mínimos locales, lo que resulta en un plateau de magnetización si el coseno es relevante en el sentido del grupo de renormalización, es decir, si los coeficientes de rigidez son suficientemente pequeños. Para valores genéricos de $S - m$ las excitaciones magnéticas no tienen gap. La presencia de una fase de Berry conlleva nuevamente a una interferencia destructiva entre diferentes configuraciones de vórtices.

Interacción de Dzyaloshinskii–Moriya

En la sección 2.2 introduciremos las interacciones antisimétricas, o de Dzyaloshinskii–Moriya. Sin embargo, podemos adelantarnos por un momento y considerar un vector $\mathbf{D} // \hat{\mathbf{z}}$ uniforme en la cadena que da lugar a la perturbación

$$\begin{aligned}
H_{DM} &= \sum_j \mathbf{D} \cdot (\mathbf{S}_j \times \mathbf{S}_{j+1}) = D \int \frac{dx}{a} \frac{S_j^- S_{j+1}^+ - S_j^+ S_{j+1}^-}{2i} \\
&\equiv -D \int \frac{dx}{a} \text{Im} S_j^+ S_{j+1}^-.
\end{aligned} \tag{1.102}$$

En la ecuación anterior escribimos la parte imaginaria porque los espines son campos y ya no operadores. Usando (1.82) obtenemos a primer orden

$$H_{DM} = -D \int dx (\partial_x \phi) (S^2 - m^2). \tag{1.103}$$

La acción completa entonces es

$$S = \int dx d\tau \left\{ \frac{K_\tau}{2} (\partial_\tau \phi)^2 + \frac{K_x}{2} (\partial_x \phi)^2 + K'_x (\partial_x \phi) + i \left(\frac{S - m}{a} \right) \partial_\tau \phi \right\}, \tag{1.104}$$

donde $K'_x = -D(S^2 - m^2)$ y K_x y K_τ están definidos bajo la ecuación (1.86). Al estudiar las corrientes de espín en la cadena diente de sierra (sección 2.2) usaremos las últimas dos ecuaciones para interpretar informalmente la fase que llamamos ‘de pequeña rigidez’ (*low stiffness phase*).

2 Antiferromagnetos cuánticos frustrados en baja dimensión

2.1 Cadena Kagomé

Por un lado, esta geometría tiene interés teórico, pues puede pensarse como la reducción a una dimensión espacial de la famosa red de Kagomé en dos dimensiones. Sin embargo, existen dos compuestos sintetizados en laboratorio[9] que poseen átomos magnéticos dispuestos en tiras como la de la figura 2.1, suficientemente separadas entre sí como para despreciar la interacción entre ellas y realizar un modelo uno-dimensional.

Dependiendo de los acoplamientos magnéticos, el modelo de Heisenberg sobre esta geometría puede desarrollar dos plateaux de magnetización clásicos en $m = 1/5$ y $m = 3/5$, donde m es la magnetización total normalizada por espín. Nos referimos a ellos como clásicos en el sentido en que son reproducibles con una simulación Monte Carlo clásica del sistema usando el límite de alta anisotropía de Ising. Sobre ellos, los valores de expectación térmicos del modelo de Ising para las funciones de correlación de dos puntos arrojan valores muy similares a los valores de expectación cuánticos a temperatura cero, como mostramos en la referencia [10]. A continuación nos concentraremos en un plateau de magnetización cuántico del modelo, sin análogo clásico en el límite de Ising, que da lugar a lo que se conoce como un cristal de magnones.

2.1.1 Introducción: Cristal de magnones

Siguiendo a la bibliografía [11], consideramos el Hamiltoniano de Heisenberg XXZ sobre la cadena Kagomé

$$\mathcal{H} = \sum_{\langle i,j \rangle} J_{ij} \left\{ \Delta S_i^z S_j^z + \frac{1}{2} (S_i^+ S_j^- + S_i^- S_j^+) \right\} - h \sum_i S_i^z. \quad (2.1)$$

Para campo magnético h suficientemente grande, el estado fundamental del Hamiltoniano (2.1) es el vacío de magnones,

$$|0\rangle \equiv |\uparrow\uparrow\uparrow\uparrow\uparrow \dots\rangle. \quad (2.2)$$

Las excitaciones de menor energía sobre (2.2) serán estados de un magnón, $|1\rangle$, de la forma

$$|1\rangle = \sum_l a_l S_l^- |0\rangle. \quad (2.3)$$

Para ciertas combinaciones de J_{ij} en el Hamiltoniano (2.1), la mínima relación de dispersión de los magnones, $\epsilon_0(k)$, resulta independiente de k , lo que se conoce como una banda chata (*flat band*). En ese caso, los magnones pueden localizarse en una región finita de la cadena en

el espacio real. Estas excitaciones localizadas pueden tener N posiciones diferentes, donde N es el número de celdas unidad en la cadena¹. Luego, puede construirse un estado que contenga n excitaciones localizadas no interactuantes, siempre que estén suficientemente lejos entre sí, todas con la misma energía. El número máximo de excitaciones no interactuantes en la cadena Kagomé es $n_{m\acute{a}x} = \frac{N}{2}$. Es posible entonces que este conjunto de $n_{m\acute{a}x}$ excitaciones localizadas constituya la excitación de menor energía. En ese caso, se espera un salto en la magnetización $\delta m = 1/n_s = 0.2$ (con $n_s = 5$ el número de sitios en la celda unidad) justo antes de llegar a la saturación, para un número par de celdas en la cadena Kagomé. Denotando L a la región donde se localiza un magnón aislado, esquematizada en la figura 2.1, entonces en (2.3) $a_l \neq 0$ sólo para $l \in L$.

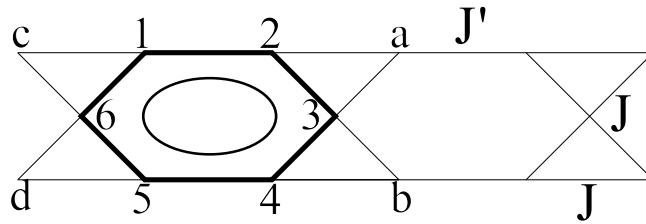


Figura 2.1: Esquema de la cadena Kagomé donde se señala con trazo grueso la zona L donde se localiza un magnón, simbolizado por la elipse. Los números simbolizan la indexación de los sitios de la zona L. Las letras simbolizan la indexación de los sitios que están en interacción con la zona L. En esta sección, por comodidad llamaremos J_l , $l = 1, \dots, 6$, al acoplamiento entre los espines \mathbf{S}_l y \mathbf{S}_{l+1} en L , con $\mathbf{S}_7 \equiv \mathbf{S}_1$. Además, llamaremos $J_{l\alpha}$ $\alpha = a, b, c, d$ al acoplamiento entre \mathbf{S}_l y \mathbf{S}_α .

Para hallar el estado $|1\rangle$, descomponemos el Hamiltoniano en tres partes,

$$\mathcal{H} = \mathcal{H}_L + \mathcal{H}_{L-R} + \mathcal{H}_R, \quad (2.4)$$

donde \mathcal{H}_L corresponde a la parte local, \mathcal{H}_{L-R} corresponde a interacción de la parte local con el resto de la cadena y \mathcal{H}_R corresponde al resto de la cadena. Luego, obtenemos

$$\begin{aligned} \mathcal{H}_L |1\rangle = & \left\{ \frac{\Delta}{4} \left(\sum_{l \in L} J_l \right) + h \left(1 - \frac{N_L}{2} \right) \right\} |1\rangle + \\ & \left(-\frac{\Delta}{2} \right) \sum_{l \in L} a_l (J_l + J_{l-1}) S_l^- |0\rangle + \frac{1}{2} \sum_{l \in L} (a_{l-1} J_{l-1} + a_{l+1} J_l) S_l^- |0\rangle, \end{aligned} \quad (2.5)$$

donde $N_L = 6$ y J_l es la constante de acoplamiento entre el espín l -ésimo y el espín $(l+1)$ -ésimo en L (ver figura 2.1). Y por otra parte

¹Aquí estamos considerando que la celda unidad de la cadena Kagomé está formada por 5 espines dispuestos en dos triángulos que comparten un vértice, como aquellos acoplados por J en la figura 2.1.

$$\begin{aligned} \mathcal{H}_{L-R} |1\rangle = & \left\{ -\frac{\hbar}{2} N_\alpha + \frac{\Delta}{4} \left(\sum_{\langle l, \alpha \rangle} J_{l\alpha} \right) \right\} |1\rangle + \\ & \left(\frac{-\Delta}{2} \right) \sum_{\langle l, \alpha \rangle} J_{l\alpha} a_l S_l^- |0\rangle + \frac{1}{2} \sum_{\langle l, \alpha \rangle} J_{l\alpha} a_l S_\alpha^- |0\rangle, \end{aligned} \quad (2.6)$$

con $\alpha = a, b, c, d$ (ver figura 2.1) y $N_\alpha = 4$. Observar que en el último término de (2.6), aparecen estados $S_\alpha^- |0\rangle$. La presencia de estos estados indica que el magnón abandona la zona L . Por lo tanto, para localizar los magnones en la zona L es necesaria la condición

$$\sum_{\langle l, \alpha \rangle} J_{l\alpha} a_l = 0, \quad \forall \alpha. \quad (2.7)$$

Para satisfacer la condición (2.7), puede elegirse

$$\begin{cases} a_l = \frac{(-1)^l}{\sqrt{6}} \\ J_{l\alpha} = J. \end{cases} \quad (2.8)$$

Fijando (2.8), $|1\rangle$ es autoestado de \mathcal{H} si y sólo si

$$J' = J \frac{2\Delta + 1}{\Delta + 1}. \quad (2.9)$$

Donde llamamos J' al acoplamiento que conecta celdas unidad contiguas (con la notación de la figura 2.1, $J' = J_1 = J_4$) y J a todo otro acoplamiento.

Con el objetivo de observar el salto en la curva de magnetización predicho analíticamente calculamos las curvas de magnetización utilizando DMRG (siglas de Density Matrix Renormalization Group) para tres anisotropías distintas, $\Delta = 1/2, 1, 3/2$ con $J = 1$ y $J' = J \frac{2\Delta+1}{\Delta+1}$, es decir, satisfaciendo la condición (2.9) y observamos el salto en la magnetización de magnitud $\delta m = 0.2$ justo antes de llegar a la saturación (figura 2.2).

Todos los cálculos numéricos con DMRG se llevaron adelante usando las librerías de ALPS² (Algorithms and Libraries for Physics Simulations). En el apéndice 5.1 presentamos un breve resumen de cómo funciona este algoritmo. En todos los cálculos con DMRG sobre la cadena Kagomé hemos conservado 500 estados para la descripción del estado fundamental del sistema, y hemos utilizado condiciones de contorno periódicas.

²<http://alps.comp-phys.org/mediawiki>

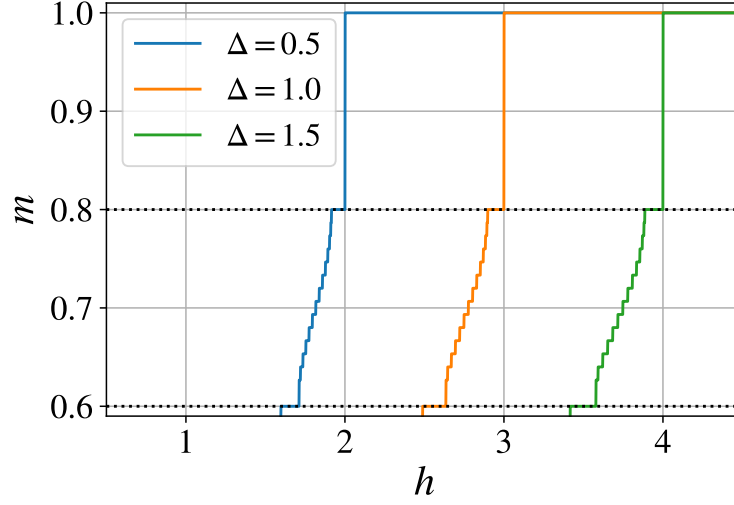


Figura 2.2: Curvas de magnetización para distintas anisotropías con $N = 30$ celdas unidad, $\Delta = 1/2, 1, 3/2$, $J = 1$ y $J' = J \frac{2\Delta+1}{\Delta+1}$ (satisfaciendo la condición (2.9)). En los tres los casos puede apreciarse el salto de amplitud $\delta m = 0.2$ justo antes de llegar a la saturación. Adicionalmente, se observa un plateau de magnetización en $m = 4/5$ para las tres anisotropías.

Además del salto de magnetización, se observa en la figura 2.2 la presencia de un plateau de magnetización en $m = 4/5$ para los tres valores de anisotropía. Este plateau fraccionario de magnetización está permitido por el criterio OYA si la celda unidad se duplica en tamaño y contiene 10 espines $1/2$. Para estudiarlo, utilizamos la técnica del hamiltoniano efectivo de baja energía. Para diferenciar con lo que sigue, llamamos ‘punto-I’ al conjunto de acoplamientos que satisfacen (2.9).

2.1.2 Hamiltoniano perturbativo efectivo

Por simplicidad, presentamos la construcción de un Hamiltoniano perturbativo de baja energía siguiendo a la bibliografía [12], donde la técnica fue introducida por primera vez. Entonces, antes de trabajar sobre la cadena Kagomé, usaremos ilustrativamente la escalera de espín $S = 1/2$ esquematizada en la figura 2.3, modelada según el siguiente Hamiltoniano

$$H = \sum_{n=1}^N \left\{ J \mathbf{S}_{1,n} \cdot \mathbf{S}_{2,n} + J_2 \sum_{\beta=1}^2 \mathbf{S}_{\beta,n} \cdot \mathbf{S}_{\beta,n+1} + 2J_1 \mathbf{S}_{1,n} \cdot \mathbf{S}_{2,(n+1)} - h \sum_{\beta=1}^2 S_{\beta,n}^z \right\}, \quad (2.10)$$

con condiciones de contorno periódicas, donde $n = 1, \dots, N$ indexa a los ‘escalones’ en la escalera. La celda unidad es un escalón o dímero (conjunto de dos espines) con un acoplamiento magnético $J > 0$. Las celdas se acoplan entre sí vía $J_2 > 0$ y $2J_1 > 0$, donde el último acoplamiento introduce frustración.

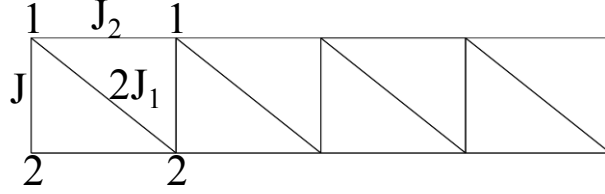


Figura 2.3: Esquema de la escalera de espín utilizada para construir un Hamiltoniano efectivo de baja energía. Los sitios en la cadena superior se indexan con 1 y los sitios de la cadena inferior con un 2.

Caso límite de dímeros no interactuantes

Considerando en primer lugar el caso límite $J_2 = J_1 = 0$, el Hamiltoniano anterior resulta

$$H^{(0)}(J, h) \equiv \sum_n H_n^{(0)}(J, h) = \sum_n \left\{ J \mathbf{S}_{1,n} \cdot \mathbf{S}_{2,n} - h \sum_{\beta=1}^2 S_{\beta,n}^z \right\}. \quad (2.11)$$

Es decir, se desacopla cada dímerno del resto de la cadena. Este puede estar en $2^2 = 4$ estados:

$$\begin{cases} |\tau^+\rangle \equiv |++\rangle \\ |\tau^0\rangle \equiv \frac{|+-\rangle + |-+\rangle}{\sqrt{2}} \\ |\tau^-\rangle \equiv |--\rangle \\ |s\rangle \equiv \frac{|+-\rangle - |-+\rangle}{\sqrt{2}}. \end{cases} \quad (2.12)$$

Para $h \equiv h_0 = J$ los estados $|\tau^+\rangle$ y $|s\rangle$ están degenerados con energía $\epsilon_0 = -\frac{3}{4}J$, mientras que $|\tau^-\rangle$ y $|\tau^0\rangle$ son estados excitados, con energías $\frac{5}{4}J$ y $\frac{1}{4}J$ respectivamente. En la figura 2.4 mostramos las energías de los 4 estados como función de h .

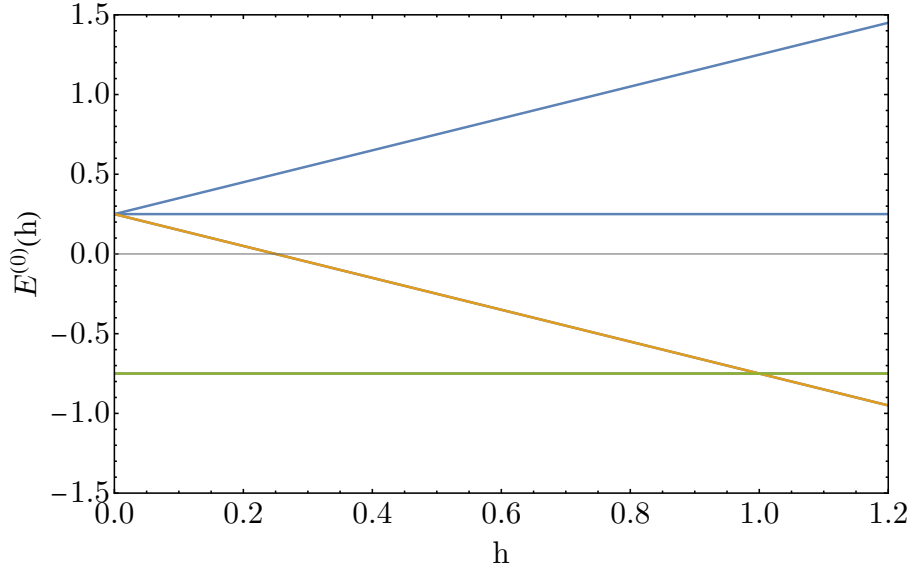


Figura 2.4: Energías de un dímero aislado como función del campo magnético aplicado, utilizando $J = 1$. En verde, el estado singlete, de espín cero. En naranja, el estado triplete, de espín 1. Ambos están degenerados en $h \equiv h_0 = J$.

Denotamos $|p_i\rangle$ a los 2^N estados degenerados de (2.11) con $h = h_0$ con mínima energía $E_0^{(0)} = N\epsilon_0$ y $|q_\alpha\rangle$ a los $4^N - 2^N$ estados excitados, de energías $E_\alpha^{(0)}$. Explícitamente, tenemos

$$\begin{aligned}
 |p_i\rangle &= \begin{cases} |s\rangle \otimes |s\rangle \otimes |s\rangle \otimes \dots \\ |\tau^+\rangle \otimes |\tau^+\rangle \otimes |\tau^+\rangle \otimes \dots \\ |\tau^+\rangle \otimes |s\rangle \otimes |s\rangle \otimes \dots \\ \dots \end{cases} \\
 |q_\alpha\rangle &= \begin{cases} |\tau^-\rangle \otimes |s\rangle \otimes |s\rangle \otimes \dots \\ |\tau^0\rangle \otimes |\tau^+\rangle \otimes |\tau^+\rangle \otimes \dots \\ |\tau^+\rangle \otimes |\tau^-\rangle \otimes |\tau^+\rangle \otimes \dots \\ \dots \end{cases}
 \end{aligned} \tag{2.13}$$

Por ejemplo, el vector $|q_1\rangle = |\tau^-\rangle \otimes |s\rangle \otimes |s\rangle \otimes \dots$ a campo $h \equiv h_0 = J$ tiene energía $E_1^{(0)}(h = h_0) = \frac{5}{4}J + (N - 1)\epsilon_0$

Aproximación perturbativa

Consideramos ahora $0 < J_2, J_1, (h - h_0) \ll J$. Es decir, consideramos el caso donde el acoplamiento dominante es J , y el campo magnético está cerca del campo magnético h_0 del cruce de niveles. A primer orden en teoría de perturbaciones, tenemos al Hamiltoniano que

podemos diagonalizar exactamente

$$H^{(0)}(J, h_0) = \sum_n \left\{ J \mathbf{S}_{1,n} \cdot \mathbf{S}_{2,n} - h_0 \sum_{\beta=1}^2 S_{\beta,n}^z \right\}, \quad (2.14)$$

y a la perturbación

$$H^{(1)} = \sum_{ij} |p_i\rangle \langle p_i| V |p_j\rangle \langle p_j|. \quad (2.15)$$

Y a segundo orden

$$H^{(2)} = \sum_{ij} \sum_{\alpha} |p_i\rangle \frac{\langle p_i| V |q_{\alpha}\rangle \langle q_{\alpha}| V |p_j\rangle}{E_0 - E_{\alpha}} \langle p_j|, \quad (2.16)$$

donde

$$V = \sum_n \left\{ 2J_1 \mathbf{S}_{1,n} \cdot \mathbf{S}_{2,(n+1)} + \sum_{\beta=1}^2 (J_2 \mathbf{S}_{\beta,n} \cdot \mathbf{S}_{\beta,n+1} - (h - h_0) S_{\beta,n}^z) \right\}. \quad (2.17)$$

Primer orden

A primer orden en teoría de perturbaciones los elementos de matriz de la perturbación en la base de estados degenerados se escriben

$$\langle p_i | H^{(1)} | p_j \rangle = J_2 \sum_{n,\beta} \langle p_i | \mathbf{S}_{\beta,n} \cdot \mathbf{S}_{\beta,n+1} | p_j \rangle + 2J_1 \sum_n \langle p_i | \mathbf{S}_{1,n} \cdot \mathbf{S}_{2,n+1} | p_j \rangle - (h - h_0) \sum_{\beta,n} \langle p_i | S_{\beta,n}^z | p_j \rangle. \quad (2.18)$$

Luego, escribiendo

$$\begin{aligned} \mathbf{S}_{\beta,n} &= \mathbf{1} \otimes \mathbf{1} \otimes \dots \otimes \overset{\text{sitio } n}{\mathbf{S}_{\beta}} \otimes \mathbf{1} \otimes \mathbf{1} \otimes \dots \\ |p_i\rangle &= |p_{i,1}\rangle \otimes |p_{i,2}\rangle \otimes |p_{i,3}\rangle \otimes \dots \end{aligned} \quad (2.19)$$

con $\mathbf{S} = \frac{1}{2}\boldsymbol{\sigma}$, $|p_{i,j}\rangle$ uno de los dos estados degenerados en el cruce y $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ las matrices de Pauli, se tiene

$$\begin{aligned} \langle p_i | H^{(1)} | p_j \rangle &= J_2 \sum_{n,\beta,\gamma} \dots \langle p_{i,n-1} | p_{j,n-1} \rangle \langle p_{i,n} | S_{\beta}^{\gamma} | p_{j,n} \rangle \langle p_{i,n+1} | S_{\beta}^{\gamma} | p_{j,n+1} \rangle \langle p_{i,n+2} | p_{j,n+2} \rangle \dots + \\ & 2J_1 \sum_{n,\gamma} \dots \langle p_{i,n-1} | p_{j,n-1} \rangle \langle p_{i,n} | S_1^{\gamma} | p_{j,n} \rangle \langle p_{i,n+1} | S_2^{\gamma} | p_{j,n+1} \rangle \langle p_{i,n+2} | p_{j,n+2} \rangle \dots + \\ & - (h - h_0) \sum_{n,\beta} \dots \langle p_{i,n-1} | p_{j,n-1} \rangle \langle p_{i,n} | S_{\beta}^z | p_{j,n} \rangle \langle p_{i,n+1} | p_{j,n+1} \rangle \dots \end{aligned} \quad (2.20)$$

donde cada operador de espín representado en la base de estados degenerados resulta en una matriz hermítica de 2×2 . La matriz hermítica de 2×2 más general posible se escribe

$$\begin{pmatrix} a & c + di \\ c - id & b \end{pmatrix} \quad (2.21)$$

con $a, b, c, d \in \mathbb{R}$. Por lo tanto, puede expandirse con coeficientes reales en la base $\{\sigma_0, \sigma\}$, con $\sigma_0 = \mathbf{1}_{2 \times 2}$, obteniéndose (a menos de una constante aditiva)

$$H_{eff} = \sum_n \tilde{J}_{xy} (S_n^x S_{n+1}^x + S_n^y S_{n+1}^y) + \tilde{J}_{zz} S_n^z S_{n+1}^z - \tilde{h} S_n^z, \quad (2.22)$$

con $\tilde{J}_{xy} = (-J_1 + J_2)$, $\tilde{J}_{zz} = \frac{J_1 + J_2}{2}$ y $\tilde{h} = \frac{J_1 + J_2}{4} - h + J$. Es decir que, a primer orden, se obtiene un Hamiltoniano efectivo correspondiente a una cadena de espines 1/2 de Heisenberg, anisotrópico, con interacciones a primeros vecinos. A órdenes superiores en teoría de perturbaciones aparecen interacciones a un número mayor de vecinos. Por otra parte, si en el cruce de niveles de la celda unidad se cruzaran más de dos estados, el Hamiltoniano efectivo correspondería a un sistema de espines de espín mayor que 1/2.

Campos críticos

A continuación mostramos como este modelo efectivo nos permite calcular los límites de algunos plateaux de magnetización en el sistema. Escribimos al Hamiltoniano efectivo a primer orden en teoría de perturbaciones en términos de operadores escalera

$$H_{eff} = \sum_n \tilde{J}_{xy} \frac{S_n^+ S_{n+1}^- + S_n^- S_{n+1}^+}{2} + \tilde{J}_{zz} S_n^z S_{n+1}^z - \tilde{h} \sum_n S_n^z. \quad (2.23)$$

A campo magnético $\tilde{h} \gg \tilde{J}_{xy}, \tilde{J}_{zz} > 0$ el estado fundamental de (2.23) es el vacío de magnones

$$|0\rangle = |\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \dots\rangle. \quad (2.24)$$

Consideramos la creación de un magnón en el sitio n

$$|n\rangle = S_n^- |0\rangle, \quad n = 1, 2, \dots, N. \quad (2.25)$$

Si nos restringimos a estados de un magnón, el Hamiltoniano (2.23) es diagonal en la base de momentos

$$|k\rangle = \frac{1}{\sqrt{N}} \sum_n e^{ikn} |n\rangle, \quad k \equiv k_i = \frac{2\pi}{N} i, \quad i = 1, \dots, N, \quad (2.26)$$

donde

$$\langle k | H_{eff} | k \rangle \equiv \epsilon(k) = \tilde{J}_{xy} \cos(k) + \tilde{J}_{zz} \frac{N-4}{4} - \tilde{h} \frac{N-2}{2}. \quad (2.27)$$

Por otra parte, el estado $|0\rangle$ tiene una energía

$$E_0 = \langle 0 | H_{eff} | 0 \rangle = N \left(\frac{\tilde{J}_{zz}}{4} - \frac{\tilde{h}}{2} \right), \quad (2.28)$$

y el gap de energía entre los estados $|0\rangle$ y $|k\rangle$ resulta

$$\epsilon(k) - E_0 = \tilde{J}_{xy} \cos(k) - \tilde{J}_{zz} + \tilde{h}. \quad (2.29)$$

Cuando las excitaciones de mínima energía se vuelven gapless, el estado $|0\rangle$ deja de ser el estado fundamental del sistema. Es decir, si

$$\text{Mín}(\epsilon(k) - E_0) = -\tilde{J}_{xy} - \tilde{J}_{zz} + \tilde{h} = 0, \quad (2.30)$$

entonces

$$\tilde{h} \equiv \tilde{h}_{c1} = \tilde{J}_{xy} + \tilde{J}_{zz}. \quad (2.31)$$

Llamamos entonces campo crítico \tilde{h}_{c1} al campo magnético $\tilde{h} > 0$ que cumple (2.30). Por simetría, el campo crítico $\tilde{h}_{c2} \ll -\tilde{J}_{xy}, -\tilde{J}_{zz} < 0$ por debajo del cual la cadena esta completamente polarizada en la dirección $-\tilde{z}$ es $\tilde{h}_{c2} = -\tilde{h}_{c1}$. Volviendo al problema original, este Hamiltoniano efectivo modela un cruce de dos estados, digamos en general, $|1\rangle$ y $|2\rangle$. Los campos críticos \tilde{h}_{c1} y \tilde{h}_{c2} se escriben en términos de los acoplamientos magnéticos del sistema J_{ij} y el campo magnético h . Conociendo esta dependencia, pueden hallarse campos críticos h_{c1} y h_{c2} tales que cuando $h > h_{c1}$ todas las plaquetas se encuentran en el estado $|2\rangle$ y cuando $h < h_{c2}$ todas las plaquetas están en el estado $|1\rangle$. Por lo tanto, estos campos críticos marcan *bordes* de plateaux de magnetización en el sistema original.

Cadena de Heisenberg de espín 1/2

El modelo efectivo obtenido a partir de trabajar a primer orden en teoría de perturbaciones resulta la cadena de Heisenberg anisotrópica de espines 1/2, con interacciones a primeros vecinos. Este modelo tiene solución exacta, conocida como Bethe Ansatz [13][14]. El panel izquierdo de la figura 2.5 muestra un esquema del diagrama de fases $h/|J_{xy}|$ versus Δ , a $T = 0$, tomado de la referencia [15]. Hay tres fases: la ferromagnética (F), la fase del líquido de Luttinger (LL) y la fase tipo Néel (N).

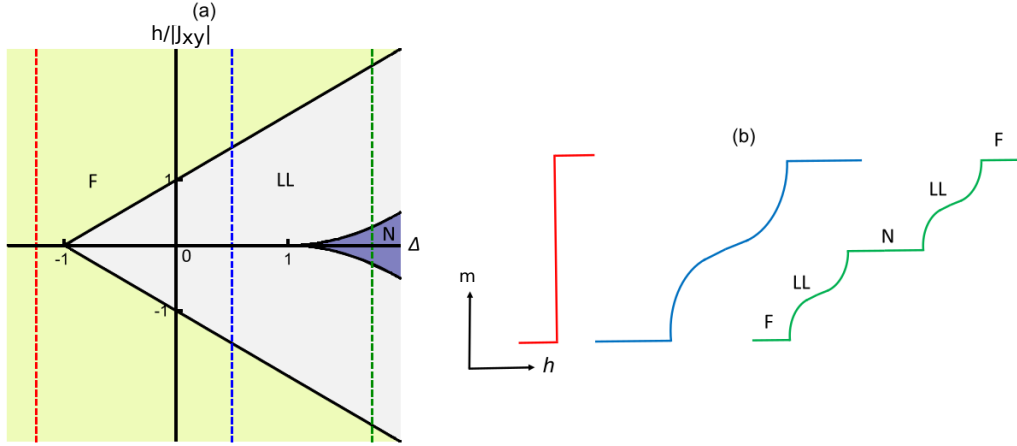


Figura 2.5: Figura (a): Diagrama de fases a temperatura cero según campo magnético h/J_{xy} y anisotropía $\Delta = J_{zz}/|J_{xy}|$ del Bethe Ansatz. En amarillo, la fase ferromagnética (F); en blanco, la fase de ‘Líquido de Luttinger’ (LL); en azul, la fase tipo Néel (N). Con líneas verticales punteadas se marcan 3 valores de Δ que dan lugar a diferentes tipos de curvas de magnetización, esquematizadas en la figura (b) y descritas en el texto general.

El panel derecho de la figura 2.5 muestra tres curvas de magnetización que corresponden a mover el campo magnético en cada una de las tres rectas punteadas verticales (roja, azul y verde) del diagrama de fases, para distintos valores de Δ .

- La curva roja, que corresponde a la fase F, muestra un salto abrupto desde magnetización (normalizada) $m = -1$ a $m = +1$, en $h = 0$.
- La curva azul corresponde a comenzar en la fase F para h grande y negativo, donde $m = -1$, luego cruzar por la fase LL, mostrando la forma característica de la transición, hasta llegar a la fase F, con $m = +1$.
- Finalmente, la curva verde representa el caso en que el sistema pasa por la fase N donde aparece un plateau de magnetización en $m = 0$.

El borde entre la fase F y la fase LL (o línea crítica) está dado por la relación

$$h_{L-F} = \pm |J_{xy}|(1 + \Delta), \quad (2.32)$$

que coincide con (2.31), mientras que el borde entre las fases LL y N está dado por la relación

$$h_{L-N} = \pm |J_{xy}| \sinh(g) \sum_{n=-\infty}^{\infty} \frac{(-1)^n}{\cosh(ng)}, \quad (2.33)$$

con $g = \text{arcosh}(\Delta)$ [15]. Resaltamos que en la fase N hay una ruptura espontánea de la simetría de traslación del sistema, ya que la celda unidad en ese caso contiene dos espines. Más abajo interpretaremos el diagrama de fases y las curvas de magnetización de la figura 2.5 aplicado a la cadena Kagomé.

Espinones

Si la cadena se encuentra en el estado fundamental de magnetización nula, las excitaciones sobre él son *espinones*; cuasipartículas de masa nula, carga nula y espín 1/2 (fermiones). Definimos $|n\rangle$ como un estado tipo Néel con una pared de dominio en la posición n .

$$\begin{aligned} |n^-\rangle &= |\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\dots\rangle \\ |n^+\rangle &= |\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\dots\rangle \end{aligned} \quad (2.34)$$

Destacamos que si la cadena tiene N sitios, un estado con una pared de dominio sólo puede obtenerse con N impar. Por lo tanto, $|n^\pm\rangle$ tiene magnetización total $M = \pm 1/2$. No escala con el número N de espines, porque hay un único espín desapareado. Además, $|n^\pm\rangle$ no es un autoestado del Hamiltoniano. Definimos

$$|k^\pm\rangle = \sum_n e^{ikn} |n^\pm\rangle, \quad k = \frac{2\pi}{L}n, \quad n = 0, 1, \dots, N-1 \quad (2.35)$$

Luego, evaluamos la acción del Hamiltoniano sobre, por ejemplo, $|k^-\rangle$. Por un lado, tomando la parte xy del Hamiltoniano, tenemos

$$H_{xy} |k^-\rangle = \sum_{m,n} \frac{J_{xy}}{2} e^{ikn} (S_m^+ S_{m+1}^- + S_m^- S_{m+1}^+) |n^-\rangle \quad (2.36)$$

Observamos que

- si $m = n \rightarrow (S_m^+ S_{m+1}^- + S_m^- S_{m+1}^+) |n^-\rangle = 0$
- si $m + 1 = n \rightarrow (S_m^+ S_{m+1}^- + S_m^- S_{m+1}^+) |n^-\rangle = |n - 2\rangle$
- si $m - 1 = n \rightarrow (S_m^+ S_{m+1}^- + S_m^- S_{m+1}^+) |n^-\rangle = |n + 2\rangle$

Notamos que los términos $m + 1 = n$ y $m - 1 = n$ mueven la pared de dominio dos posiciones en la cadena. Por otra parte,

- si $m + 2 = n \rightarrow (S_m^+ S_{m+1}^- + S_m^- S_{m+1}^+) |n^-\rangle = |\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\dots\rangle$
- si $m + 3 = n \rightarrow (S_m^+ S_{m+1}^- + S_m^- S_{m+1}^+) |n^-\rangle = |\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\dots\rangle$,

entonces, tenemos que si $m \neq n \pm 1$, la aplicación $H_{xy} |n^-\rangle$ no da como resultado un estado de una única pared de dominio. Luego, obtenemos

$$H_{xy} |k^\pm\rangle = J_{xy} \cos(2k) |k^\pm\rangle + \sum, \quad (2.37)$$

donde \sum es una suma sobre estados que no tienen la forma de una única pared de dominio moviéndose por la cadena. Por otra parte, evaluamos

$$H_z |k^\pm\rangle = \sum_{m,n} e^{ikn} (J_{zz} S_m^z S_{m+1}^z - h S_m^z) |n^\pm\rangle, \quad (2.38)$$

y obtenemos

$$H_z |k^\pm\rangle = \frac{J_{zz}}{2} |k^\pm\rangle - \frac{NJ_{zz}}{4} |k^\pm\rangle \mp \frac{h}{2} |k^\pm\rangle. \quad (2.39)$$

Finalmente, si consideramos la aproximación en que las excitaciones sobre el estado fundamental de magnetización nula son paredes de dominio moviéndose por un estado de Néel, la energía que llevan se escribe

$$\epsilon^\pm(k) \approx J_{xy} \cos(2k) + \frac{J_{zz} \mp h}{2} \quad (2.40)$$

donde no consideramos el término de orden N^1 en (2.39) porque está asociado a la energía del estado fundamental (de manera análoga a lo que ocurría en el caso de excitaciones magnónicas). Al considerar los campos h_c para los cuales la excitación es *gapless*, se obtiene

$$h_c = \pm(2J_{xy} - J_{zz}). \quad (2.41)$$

Estos campos críticos nos permiten encontrar los bordes de plateaux fraccionarios, que corresponden a la fase de Néel en el modelo efectivo. Además, si el modelo efectivo no es una cadena de espines 1/2 a primeros vecinos, la solución de Bethe no nos sirve, por lo que es importante contar con este enfoque aproximado alternativo.

2.1.3 Primeras predicciones

Para el caso de la cadena Kagomé, escribimos el Hamiltoniano efectivo de baja energía a primer orden en teoría de perturbaciones para estudiar el caso homogéneo $J_i = 1$, $J_4 \ll J_i$, $i = 1, 2, 3$. En la figura 2.6 mostramos las $2^5 = 32$ energías de una celda unidad aislada. El estado fundamental para $h < h_{c1}$ está triplemente degenerado, por lo que en el cruce no puede obtenerse un Hamiltoniano efectivo de espines 1/2 con la técnica presentada en la sección anterior, sino de espines 3/2, ya que en el cruce cada plaqueta puede estar en 4 estados. Estudiamos el segundo cruce de dos niveles en $h = h_{c2} = 2.5$.

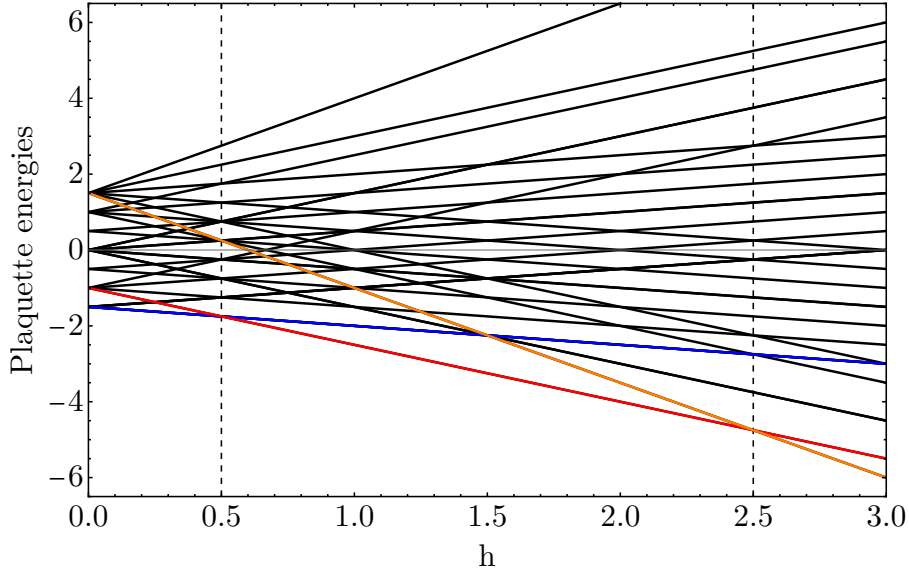


Figura 2.6: Energías de la plaqueta aislada como función del campo magnético externo h aplicado, para $J_i = 1$, $i = 1, 2, 3$. Para $h > 0$ hay tres estados fundamentales, de magnetizaciones $m = 1/5, 3/5, 1$, remarcados en azul, rojo y naranja respectivamente. Con líneas a trazos verticales marcamos los campos h_{c1} , h_{c2} a los cuales el estado fundamental esta degenerado por cruce de niveles.

Trabajando de manera análoga a la sección 2.1.2, llegamos a un Hamiltoniano efectivo correspondiente a una cadena de espines $1/2$ de Heisenberg anisotrópica como (2.22), con

$$\begin{cases} \tilde{J}_{xy} = 0.1 J_4 \\ \tilde{J}_{zz} = 0.005 J_4 \\ \tilde{h} = -2.5 + h - 0.0950 J_4 \end{cases} \quad (2.42)$$

En la figura 2.7 presentamos cuatro curvas de magnetización calculadas con DMRG, donde $J_4 = 0.05; 0.1; 0.15$ y 0.2 . Para mejor visualización, solo mostramos la parte con $m \geq 0.6$. Con líneas punteadas mostramos la predicción teórica de los campos críticos usando la técnica del Hamiltoniano efectivo con las ecuaciones (2.42) y (2.31).

Observando los acoplamientos efectivos (2.42), vemos que debido a la simpleza del acoplamiento entre las celdas en la cadena Kagomé, mediante J_4 , el parámetro de anisotropía $\tilde{\Delta} = \tilde{J}_{zz}/\tilde{J}_{xy} = 0.05$ es independiente de J_4 . Esto significa que con este modelo no podemos movernos horizontalmente en el diagrama de fases 2.5 para buscar plateaux fraccionarios.

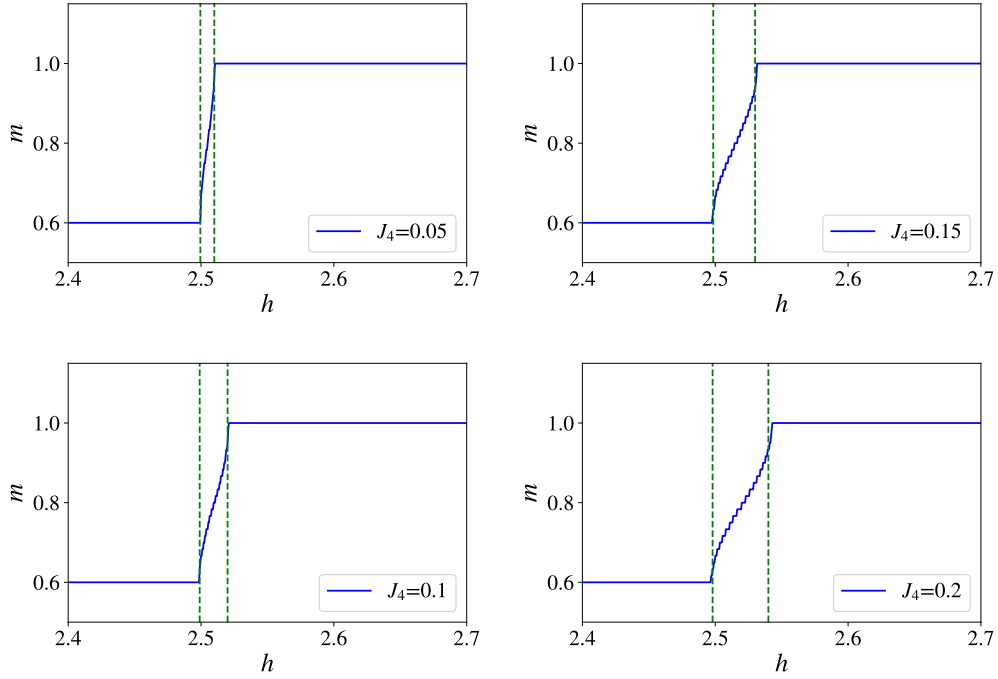


Figura 2.7: Curvas parciales de magnetización ($m > 3/5$) para $J_4 = 0.05, 0.1, 0.15, 0.2$, obtenidas con DMRG en sistemas con 200 sitios. En línea a trazos verde, la predicción teórica de los bordes de plateaux utilizando (2.31) y (2.42)

2.1.4 Plateaux fraccionarios

Celda unidad tipo lápiz

Como para el punto-I el acoplamiento que conecta celdas contiguas es el más grande, propusimos una celda unidad diferente que lo contiene, también de cinco espines, esquematizada en la figura 2.8. La llamamos ‘celda unidad tipo lápiz’. Resaltamos que esta celda unidad no está frustrada, mientras que en el acoplamiento entre celdas contiguas aparece una competencia entre los acoplamientos denominados J_2 y K_2 . Para distinguir, llamaremos a la celda unidad anterior ‘celda unidad de dos triángulos’.

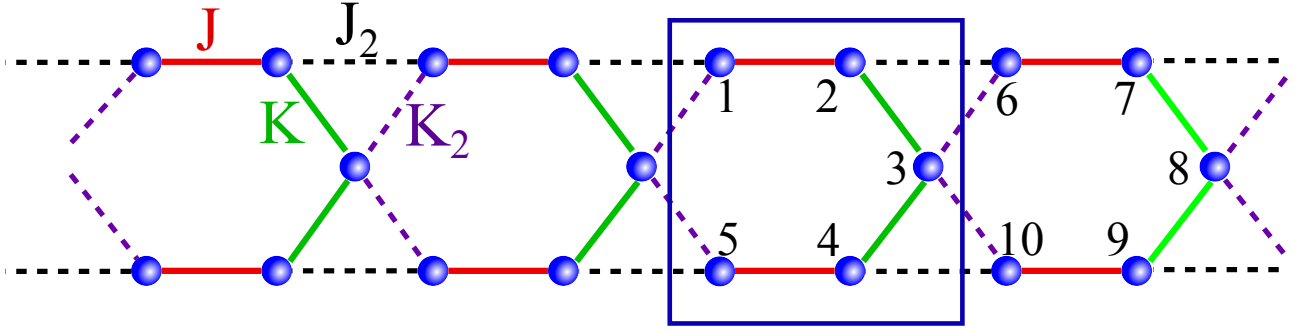


Figura 2.8: Esquema de la cadena Kagomé usando la celda unidad ‘tipo lápiz’, encuadrada con líneas azules. Los números denotan la indexación de los sitios en la cadena. La celda unidad posee solo dos acoplamientos diferentes J y K , mientras que los acoplamientos J_2 y K_2 acoplan celdas tipo lápiz contiguas.

Haciendo la separación $H = H_0 + H_{int}$ escribimos, en el caso isotrópico $\Delta = 1$,

$$H_0 = \sum_n \left[J(\vec{S}_{n,1} \cdot \vec{S}_{n,2} + \vec{S}_{n,5} \cdot \vec{S}_{n,4}) + K(\vec{S}_{n,2} \cdot \vec{S}_{n,3} + \vec{S}_{n,3} \cdot \vec{S}_{n,4}) - h_0 \sum_{m=1}^5 S_{n,m}^z \right] \quad (2.43)$$

y

$$H_{int} = \sum_n \left[J_2(\vec{S}_{n,2} \cdot \vec{S}_{n+1,1} + \vec{S}_{n,4} \cdot \vec{S}_{n+1,5}) + K_2(\vec{S}_{n,3} \cdot \vec{S}_{n+1,1} + \vec{S}_{n,3} \cdot \vec{S}_{n+1,5}) - (h - h_0) \sum_{m=1}^5 S_{n,m}^z \right]. \quad (2.44)$$

Luego, los acoplamientos del Hamiltoniano efectivo de baja energía a primer orden en teoría de perturbaciones con esta celda unidad, resultan

$$\begin{cases} \tilde{J}_{xy} = 0.289898K_2 - 0.322474J_2 \\ \tilde{J}_{zz} = 0.0519949J_2 + 0.0420204K_2 \\ \tilde{h} = -h + 0.266889J_2 + 0.0627653J_2 + 0.21569K_2 + 0.0677526K_2 + 2.11237. \end{cases} \quad (2.45)$$

Por definición, el parámetro de anisotropía es

$$\tilde{\Delta} = \frac{\tilde{J}_{zz}}{|\tilde{J}_{xy}|} \quad (2.46)$$

donde se toma el módulo en \tilde{J}_{xy} porque una rotación en π a lo largo del eje z para la sub red de espines de índice par (o alternativamente, impar) en la cadena es equivalente a cambiar

$\tilde{J}_{xy} \rightarrow -\tilde{J}_{xy}$, ya que la cantidad $S_n^x S_{n+1}^x + S_n^y S_{n+1}^y$ cambia de signo y $S_n^z S_{n+1}^z$ no lo hace. En la figura 2.9 mostramos el gráfico de anisotropía vs J_2 , fijando $J = 1.5$, $K = 1$ y $K_2 = 1$. Si bien al tomar $K_2 = 1$ no estamos en el régimen perturbativo, el modelo efectivo predice correctamente la existencia de plateaux fraccionarios, como mostramos más adelante. Con línea a trazos horizontal marcamos el valor $\tilde{\Delta} = 1$ por encima del cual según el diagrama de fases del Bethe ansatz (Fig 2.5) hay plateau fraccionario. Con línea a trazos vertical, marcamos el valor de J_2 para el cual diverge la anisotropía según (2.45) y (2.46).

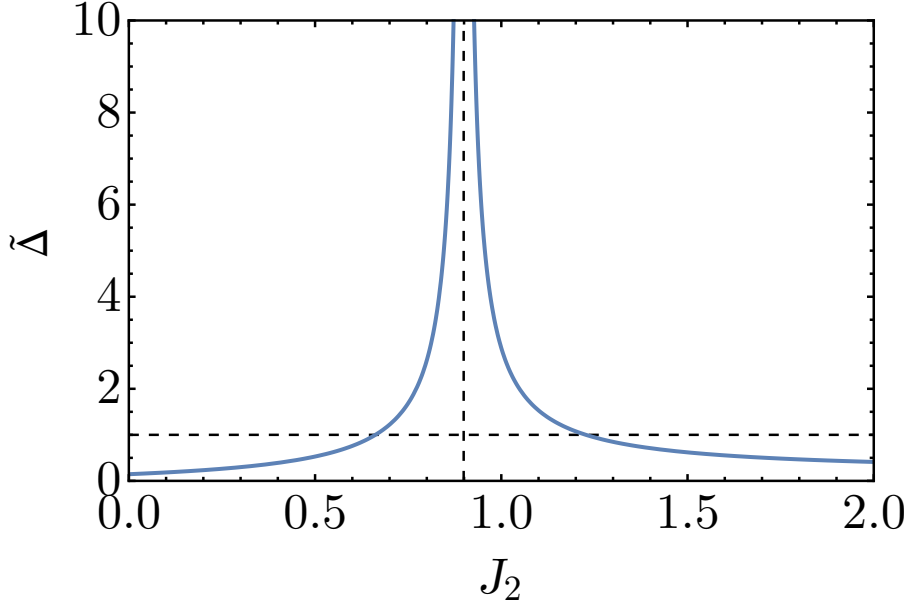


Figura 2.9: Anisotropía efectiva como función de J_2 , habiendo fijado $K_2 = 1$, $J = 1.5$ y $K = 1$. Con línea a trazos horizontal marcamos el valor $\tilde{\Delta} = 1$ por encima del cual según el diagrama de fases de Bethe ansatz (Fig 2.5) hay plateau fraccionario. Con línea a trazos vertical, marcamos el valor de J_2 para el cual diverge la anisotropía según (2.45) y (2.46)

Según (2.45) y (2.46), tomando $J_2 = K_2$ se tiene $\tilde{\Delta} \approx 2.88 > 1$. Es decir que el modelo efectivo predice que si los couplings que acoplan celdas tipo lápiz contiguas son iguales entre sí, hay plateau fraccionario. Para saber en qué región del plano (J_2, K_2) el modelo efectivo predice la presencia de un plateau fraccionario hicimos dos cálculos. Por un lado, pueden considerarse los campos críticos h_{c1} y h_{c2} que vuelven gapless a la dispersión de espinoes³ (2.40) y forzar la igualdad $h_{c1} = h_{c2}$. Las dos soluciones son:

$$\begin{cases} J_2 = 0.77162K_2 \\ J_2 = 1.04868K_2 \end{cases} \quad (2.47)$$

marcadas con rojo en la figura 2.10. Dentro de la región delimitada por las rectas rojas el modelo predice que hay un plateau fraccionario de magnetización.

Alternativamente, puede considerarse el caso $\tilde{\Delta} = 1$ en el diagrama de fases del bethe Ansatz (Fig 2.5). Si el modelo efectivo no pasa por la fase N, entonces la cadena Kagomé no presenta

³recordar que este es un resultado aproximado, pues sólo consideramos la relación de dispersión de una pared de dominio moviéndose por el estado de Néel.

plateau fraccionario. Las soluciones correspondientes son:

$$\begin{cases} J_2 = 0.661943K_2 \\ J_2 = 1.22715K_2 \end{cases} \quad (2.48)$$

marcadas con azul en la figura 2.10. En este caso aparece una región más grande, que contiene a la anterior, en la cual el modelo predice que hay plateau fraccionario. En la figura 2.10 marcamos en rosa el punto-I, $(J_2, K_2) = (1, 1)$, que corresponde a la solución exacta de magnones localizados, la cual cae dentro de las dos regiones; aunque cabe resaltar que el cálculo perturbativo tiene validez sólo en el régimen $J_2, K_2 \ll J = 1.5, K = 1$.

A continuación calculamos correlaciones y observables usando DMRG sobre la recta $J_2 = K_2$, (contenida en la región con plateau fraccionario según el modelo efectivo), para $J_2 = K_2 = 0.1, 0.2, \dots, 0.9, 1$ (puntos representados con círculos en la figura 2.10). Es decir, partiendo del régimen perturbativo y llegando hasta el punto-I, representado por un círculo rosa en $J_2 = K_2 = 1$; en la figura 2.10.

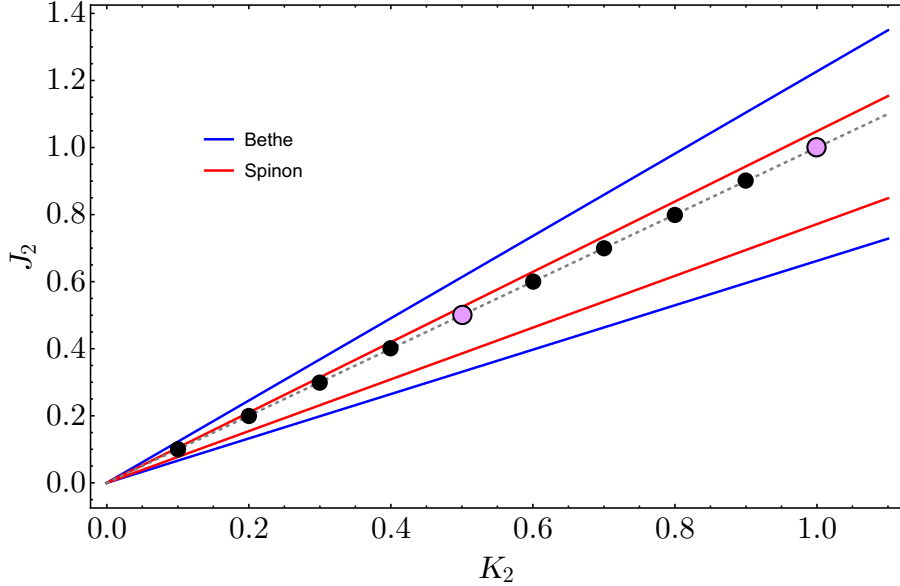


Figura 2.10: Con rectas azules, las soluciones para $\tilde{\Delta} = 1$, (Ec. (2.48)) que indican la desaparición del plateau fraccionario según Bethe Ansatz. Con rectas rojas, las soluciones a una relación de dispersión gapless de espinones sobre el estado de Néel del modelo efectivo donde los dos campos críticos coinciden (Ec. (2.47)). Entre rectas del mismo color se forman regiones donde el modelo efectivo predice la existencia de un plateau fraccionario de magnetización en $m = 4/5$. Con círculos marcamos los puntos que elegimos para calcular curvas de magnetización y correlaciones $\langle S^+ S^- \rangle$ utilizando DMRG. En particular, con círculos rosas representamos los puntos- I y II, en donde hay soluciones exactas de magnones localizados.

Para cada círculo en la figura 2.10 calculamos las correlaciones $\langle S_i^+ S_j^- \rangle$ sobre el plateau de magnetización en $m = 4/5$, presentadas en la figura 2.11, donde podemos ver que los magnones están localizados en Hexágonos para $J_2 = K_2 = 1 \equiv \tilde{J}$, y al bajar progresivamente \tilde{J} los magnones se localizan progresivamente en regiones tipo lápiz.

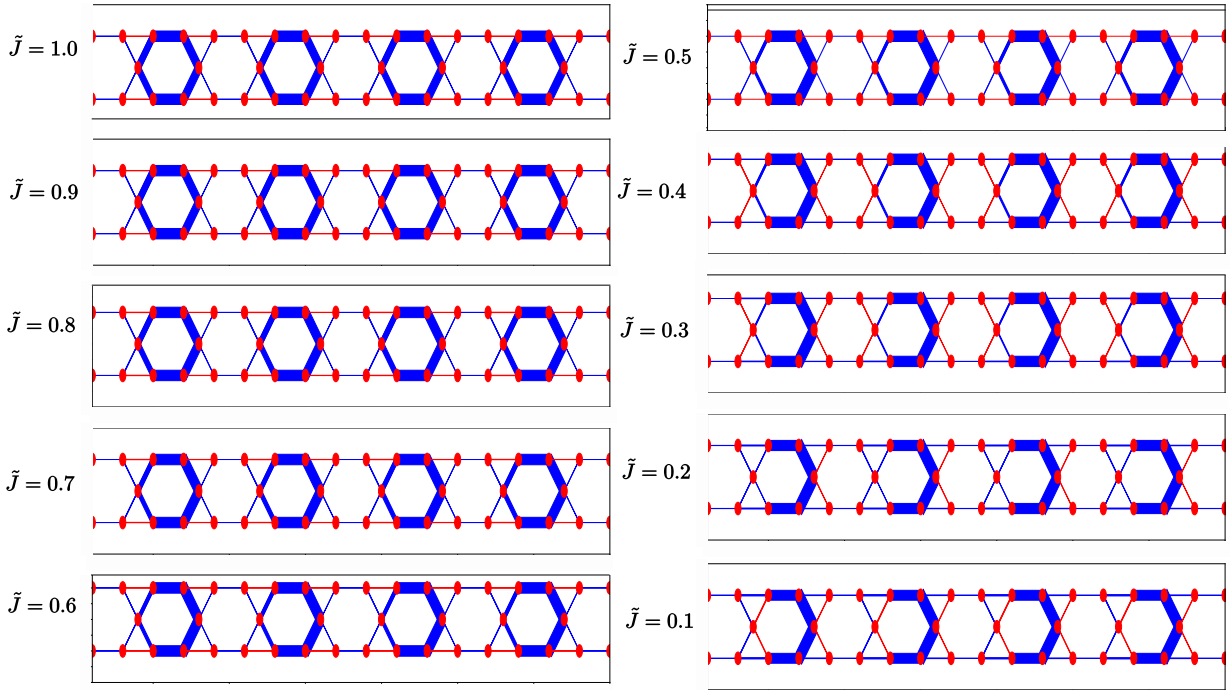


Figura 2.11: Correlaciones $\langle S_i^+ S_j^- \rangle$ representadas en el ancho de línea, para $\Delta = 1$, $J = 3/2$, $K = 1$ y $K_2 = J_2 \equiv \tilde{J} = 0.1, 0.2, \dots, 1$ sobre el plateau de magnetización en $m = 4/5$ calculadas con DMRG, en sistemas de 24 celdas unidad tipo lápiz. Líneas azules y rojas corresponden a correlaciones negativas y positivas respectivamente.

Solución exacta número 2 (punto-II)

Para $J = 3/2$, $K = 1$ y $J_2 = K_2 = 1/2$, con $\Delta = 1$ (círculo rosa inferior en la figura 2.10), observamos un salto de magnetización de $\delta m = 1/5$ justo antes de llegar a saturación (figura 2.12), idéntico a aquel observado para el punto-I (figura 2.2).

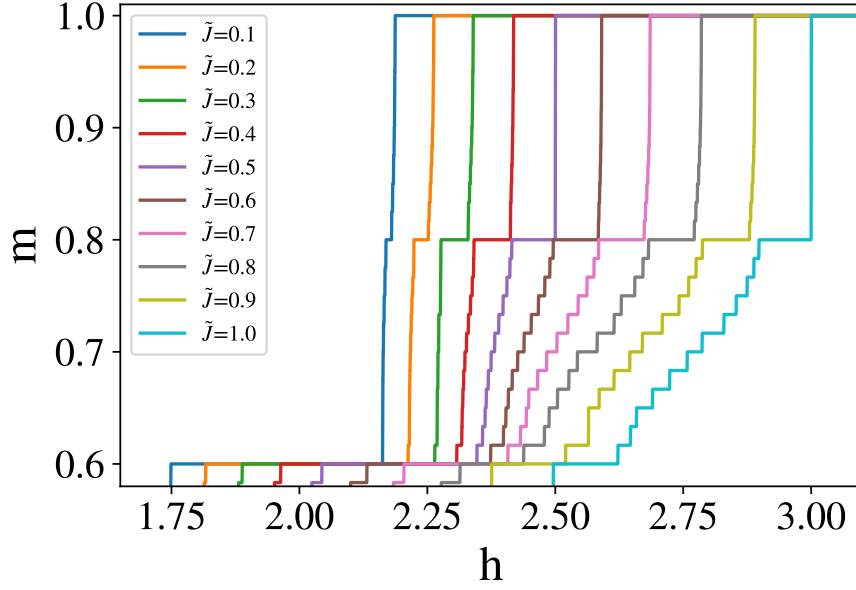


Figura 2.12: Curvas parciales de magnetización calculadas con DMRG para $J = 3/2$, $K = 1$ y $J_2 = K_2 \equiv \tilde{J} = 0.1, 0.2, \dots, 1$, usando 24 celdas unidad tipo lápiz. Puede verse que para $\tilde{J} = 0.5$ hay un salto abrupto desde $m = 4/5$ a saturación idéntico al caso del punto-I ($\tilde{J} = 1$; curva celeste).

Retomando la notación de la figura 2.1 el salto se observa para los acoplamientos:

J_1	J_2	J_3	J_4	J_5	J_6	$J_{l\alpha}$
$3/2$	1	1	$3/2$	$1/2$	$1/2$	$1/2$

Tabla 2.1: Conjunto de acoplamientos para los cuales se observa un salto de magnetización de $\delta m = 1/5$, usando la notación de la figura 2.1. J_l acopla el espín l -ésimo con el espín $(l+1)$ -ésimo de la denominada zona L, representada con trazo grueso en la figura 2.1.

La presencia de este salto nos llevó a buscar una solución exacta del Hamiltoniano (2.1) (con $\Delta = 1$) de la forma

$$|1\rangle = \sum_l a_l S_l^- |0\rangle, \quad a_l \in L \quad (2.49)$$

con coeficientes a_l a determinar. Utilizando los acoplamientos de la tabla 2.1 y la ecuación (2.6), la condición necesaria para la localización de los magnones (2.7) se escribe:

$$\begin{cases} a_2 + a_3 = 0 & (\alpha = a) \\ a_3 + a_4 = 0 & (\alpha = b) \\ \frac{1}{2}a_1 + a_6 = 0 & (\alpha = c) \\ \frac{1}{2}a_5 + a_6 = 0 & (\alpha = d) \end{cases} \quad (2.50)$$

De los seis coeficientes en (2.3), quedan solo dos libres; por ejemplo a_1 y a_2 . Forzando que $|1\rangle$ sea autoestado de (2.1), obtenemos la relación

$$\frac{a_2}{a_1} = -3/2 \quad (2.51)$$

Normalizando, obtenemos los coeficientes presentados en la tabla 2.2

a_1	a_2	a_3	a_4	a_5	a_6
-1/3	+1/2	-1/2	+1/2	-1/3	+1/6

Tabla 2.2: Coeficientes del estado (2.49) correspondientes a la solución exacta del punto-II a $\Delta = 1$.

Por lo tanto, hallamos una segunda solución exacta de magnones localizados. Nombramos ‘punto-II’ al conjunto de couplings $\{ J = 3/2, K = 1, J_2 = K_2 = 1/2 \}$ (o, equivalentemente, los acoplamientos de la tabla 2.1, con la notación de la figura 2.1.) Una vez hallado el punto-II, buscamos generalizar esta solución para el caso anisotrópico ($\Delta \neq 0$). Para ello, propusimos que los acoplamientos magnéticos tienen la forma

$$J = \frac{2\Delta + 1}{\Delta + 1} K \quad (2.52a)$$

$$K = 1 \quad (2.52b)$$

$$J_2(\Delta) = K_2(\Delta) \quad (2.52c)$$

En esta propuesta conservamos la relación entre J y K que vale en el punto-I (eq 2.52a), fijamos como referencia el valor de K a la unidad y asumimos que los acoplamientos que conectan celdas tipo lápiz son iguales (como ocurre en el punto-II) y tengan dependencia en Δ (2.52c). Encontramos dos soluciones, representadas en la figura 2.13:

- Solución-I (azul en la figura 2.13):

$$J_2 = K_2 = K \equiv 1$$

No hay dependencia en Δ . Esta solución es, por supuesto, exactamente la que corresponde al punto-I.

- Solucion-II (amarillo en la figura 2.13):

$$J_2 = K_2 = \frac{\sqrt{12\Delta^3(\Delta + 1) + 1} - (2\Delta + 1)}{2\Delta(\Delta + 1)} \quad (2.53)$$

Para $\Delta < 1/\sqrt{3}$ la solución corresponde a tener acoplamientos $J_2 = K_2$ negativos (ferromagnéticos).

Para $\Delta = 1/\sqrt{3}$ marcado con línea punteada, $J_2 = K_2(\Delta) = 0$, que corresponde al caso de celdas tipo lápiz aisladas.

Para $\Delta = 1$, la solución corresponde al llamado punto-II. (coeficientes en tabla 2.2 y acoplamientos en tabla 2.1).

Para $\Delta = 2$, el estado solución es nuevamente $a_l = \frac{(-1)^l}{\sqrt{6}}$, $l = 1, \dots, 6$, con $\{J_2 = \frac{5}{2}K, K = 1, J_2 = K_2 = 1\}$, como mostramos en la figura 2.14

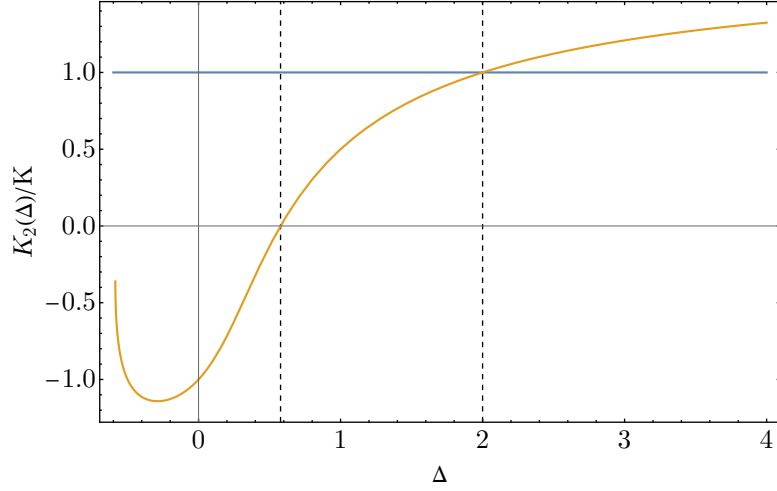


Figura 2.13: Soluciones de (2.52c). En azul, la solución correspondiente al punto-I. En amarillo, la solución correspondiente al punto-II. Con línea a trazos vertical marcamos los valores $\Delta = 1/\sqrt{3}$ y $\Delta = 2$, donde la solución-II cambia de signo y los coeficientes de ambas soluciones coinciden, respectivamente.

En la figura 2.14 mostramos los coeficientes $a_l(\Delta)$, $l = 1, \dots, 6$ que corresponden a la solución-II. Debido a que $a_1(\Delta) = a_5(\Delta)$ y $a_2(\Delta) = a_4(\Delta)$ solo se observan 4 curvas. Observamos que para $\Delta = 2$, los coeficientes coinciden con aquellos correspondientes al punto-I ($a_l = \frac{(-1)^l}{\sqrt{6}}$), marcados con línea a trazos.

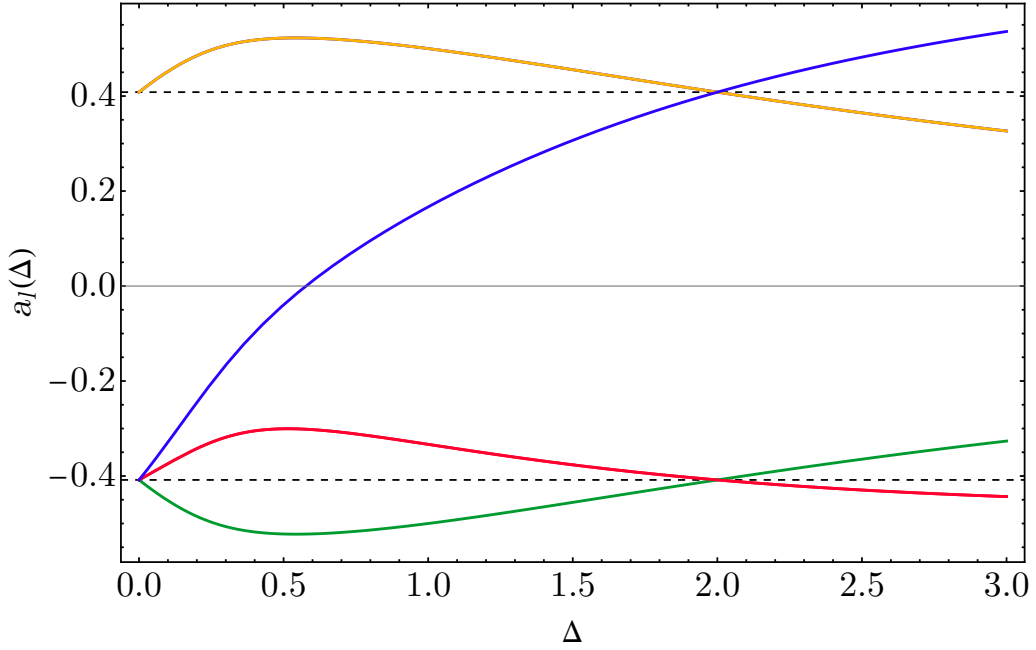


Figura 2.14: Coeficientes $a_l(\Delta)$, $l = 1, \dots, 6$ que corresponden a la solución-II. Debido a que $a_1(\Delta) = a_5(\Delta)$ y $a_2(\Delta) = a_4(\Delta)$ solo se observan 4 curvas. Observamos que para $\Delta = 2$, las coeficientes coinciden con aquellas correspondientes al punto-I ($a_l = \frac{(-1)^l}{\sqrt{6}}$), marcados con línea a trazos. a_1, a_2, a_3 y a_6 están graficados en rosa, amarillo, verde y violeta, respectivamente.

2.1.5 Conclusiones

Estudiamos el comportamiento magnético del Hamiltoniano de Heisenberg antiferromagnético a primeros vecinos sobre la geometría 1D conocida como cadena Kagomé. Calculamos los límites de diferentes plateaux de magnetización del sistema usando un Hamiltoniano efectivo de baja energía basado en teoría de perturbaciones en el límite de celdas unidad tipo lápiz débilmente interactuantes. Esta técnica resultó útil para predecir la presencia de un plateau fraccionario de magnetización en $m = 4/5$, un plateau que no puede explicarse en el límite de Ising. Este plateau tiene origen en la presencia de magnones localizados, como mostramos calculando la función de correlación $\langle S_i^+, S_j^- \rangle$ con DMRG. Además, encontramos un estado fundamental exacto con $m = 4/5$ que es una generalización del estado encontrado por Schulenburg et al.[11], compartiendo las características de producir un salto en la curva de magnetización hacia saturación, y ser un estado producto, o factorizado, generado por la fuerte frustración geométrica.

Desde un punto de vista más general, es conocido que los cristales de magnones están presentes en varios sistemas magnéticos frustrados en 1, 2 y 3 dimensiones espaciales[11]. Estos sistemas comparten la presencia de una banda chata que produce un estado fundamental factorizado, como consecuencia de la naturaleza cuántica de las excitaciones. El diagrama de fases de la cadena Kagomé presenta además plateaux de magnetización clásicos[10], que pueden observarse en una simulación Monte Carlo del sistema en el límite de alta anisotropía de Ising, es decir, sin fluctuaciones cuánticas.

Para concluir, mencionamos que la riqueza del diagrama de fases del sistema, donde pueden

encontrarse plateaux clásicos, cuánticos, metamagnetismo y fases sin gap, lo hace ideal para estudiar la influencia de perturbaciones como un campo magnético transversal o interacciones antisimétricas (Dzyaloshinskii-Moriya). En la siguiente sección abordaremos el estudio de la presencia de interacciones antisimétricas en otra escalera de espín que presenta banda chata, la cadena diente de sierra.

2.2 Cadena diente de sierra

2.2.1 Introducción

Interacción de Dzyaloshinskii–Moriya

En sistemas magnéticos, las interacciones de Dzyaloshinskii–Moriya corresponden a la versión antisimétrica de la interacción de intercambio de Heisenberg. El origen de esta interacción fue propuesto fenomenológicamente por Dzyaloshinskii[16] y derivado formalmente por Moriya[17], quien mostró que las interacciones antisimétricas son consecuencia del acoplamiento espín-órbita en materiales. Si bien la presencia de estas interacciones tiene que estar permitida por ciertas reglas de simetría introducidas en el trabajo de Moriya[17], su presencia es usual en sistemas reales y puede causar efectos diversos[18, 19], como inducir una magnetización uniforme o alternada, causar un ferromagnetismo débil en sistemas antiferromagnéticos, abrir un gap magnético, destruir o estabilizar órdenes no colineales, etc.

Para introducir formalmente a las interacciones antisimétricas, podemos considerar a una cadena de espines con el siguiente Hamiltoniano

$$H = \sum_{\langle i,j \rangle} \left[J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j + \mathbf{D}_{ij} \cdot (\mathbf{S}_i \times \mathbf{S}_j) \right], \quad (2.54)$$

donde \mathbf{D} es el vector de Dzyaloshinskii–Moriya, que rompe simetría $SU(2)$. Si $\mathbf{D}_{ij} \parallel \hat{\mathbf{z}}$, entonces usando el álgebra de los operadores de espín el Hamiltoniano anterior se escribe

$$H = \sum_{\langle i,j \rangle} J_{ij} \left(\frac{(1 + i\theta_{ij})S_i^+ S_j^- + (1 - i\theta_{ij})S_i^- S_j^+}{2} + S_i^z S_j^z \right), \quad (2.55)$$

donde $\theta_{ij} = D_{ij}^z / J_{ij}$. En el límite $\theta_{ij} \ll 1$ podemos escribir

$$H = \sum_{\langle i,j \rangle} J_{ij} \left(\frac{e^{i\theta_{ij}} S_i^+ S_j^- + e^{-i\theta_{ij}} S_i^- S_j^+}{2} + S_i^z S_j^z \right). \quad (2.56)$$

Introducir interacciones DM en muchos casos suele tomarse directamente en el sentido de la ecuación (2.56), donde el acoplamiento entre espines usual J_{ij} gana una fase $e^{i\theta_{ij}}$. θ_{ij} se suele llamar el ‘campo de gauge’, debido a su interpretación al mapear el sistema de espines a un sistema de partículas.

Lenguaje de partículas

Utilizamos un mapeo exacto entre sistemas de espines $1/2$ y sistemas de bosones de núcleo duro (*hard-core bosons*), dado por

$$\begin{aligned} b_j^\dagger &\equiv S_j^- \\ n_j = b_j^\dagger b_j &\equiv 1/2 - S_j^z, \end{aligned} \quad (2.57)$$

donde para esta convención $|0\rangle \equiv |+\rangle$ y $|1\rangle \equiv |-\rangle$. La nomenclatura ‘núcleo duro’ significa que en el sistema bosónico no es posible crear dos bosones en el mismo estado, entendiéndose que el costo energético para hacerlo es muy alto, debido a una interacción repulsiva entre ambos. Esta característica es esencial en dos aspectos. En primer lugar, permite hacer coincidir las dimensiones de los espacios de Hilbert entre el sistema de partículas y el sistema de espines. En segundo lugar, permite conservar el álgebra de $SU(2)$ luego del mapeo, ya que

$$[S^z, S^+] \equiv [1/2 - b^\dagger b, b] = -[b^\dagger, b] = b \equiv S^+, \quad (2.58)$$

pero, en cambio

$$[S^+, S^-] \equiv [b, b^\dagger] = 1 \neq 2S^z. \quad (2.59)$$

El álgebra de $SU(2)$ solo se preserva en el subespacio físico, usando explícitamente la restricción de núcleo duro. Para verlo, solo hay que aplicar $[S^+, S^-]$ a los estados $|+\rangle$ y $|-\rangle$ y $[b, b^\dagger]$ a los estados $|0\rangle$ y $|1\rangle$.

El Hamiltoniano (2.56) puede leerse entonces en términos de operadores bosónicos según

$$H = \sum_{\langle i,j \rangle} J_{i,j} \left(\frac{e^{i\theta_{i,j}} b_i b_j^\dagger + e^{-i\theta_{i,j}} b_i^\dagger b_j}{2} + (1/2 - n_i)(1/2 - n_j) \right). \quad (2.60)$$

El término $S^z S^z$ del Hamiltoniano de espines corresponde a un término de interacción entre bosones en sitios vecinos, mientras que los términos de fluctuaciones corresponden a términos de *hopping* en el sistema de partículas. El campo θ_{ij} aquí se ve explícitamente como un potencial vector para el sistema bosónico, usando lo que se conoce como sustitución de Peierls⁴. Ya que podemos entender al vector \mathbf{D} como un potencial vector, es de esperar entonces encontrar libertades de Gauge. En efecto, el término de DM puede ser removido⁵ si su integral de línea (discretizada) vale cero para una dada curva cerrada abc, \dots, yza , es decir,

$$\mathbf{D}_{ab} + \mathbf{D}_{bc} + \dots + \mathbf{D}_{yz} + \mathbf{D}_{za} = 0.$$

Por otra parte, la presencia de un campo magnético en el sistema de espines que da lugar a un término de la forma $-\hbar \sum_j S_j^z$ se escribe en término de bosones según $\hbar \sum_j n_j + \text{cte}$, por lo que el campo magnético en el sistema de espines se mapea a un potencial químico en el sistema bosónico. En nuestro trabajo no utilizaremos el mapeo a partículas para calcular observables, sino solo de manera conceptual para interpretar la presencia de interacciones antisimétricas

⁴Para una motivación, ver por ejemplo las notas de Feynman <https://www.feynmanlectures.caltech.edu>, volumen III, capítulo 21.

⁵en inglés, *gauged away*.

como un potencial vector, y sus consecuencias.

Modelos de Hubbard en presencia de un potencial vector sobre la cadena sawtooth[20] y otras geometrías cuasi-1-dimensionales han recibido mucha atención[21, 22, 23, 24] debido a la presencia de diversas fases quirales y topológicas y su relación con física de superconductores.

Corriente de espín

Definiendo la densidad de espín como $\rho_l \equiv S_l^z$, la ecuación de continuidad en el esquema de Heisenberg se escribe

$$\partial_t \rho_l \equiv -i [S_l^z, H] = -\nabla \cdot \mathbf{j}_l, \quad (2.61)$$

definiendo así a la corriente de espín, \mathbf{j} , como aquella que satisface (2.61). Para una cadena de espines de Heisenberg con el Hamiltoniano (2.56), introducir un vector \mathbf{D} uniforme a lo largo del eje z da lugar a la siguiente ecuación de continuidad

$$-i [S_l^z, H] = -\frac{J}{2} ((ie^{i\theta} S_l^+ S_{l+1}^- + h.c.) - (ie^{i\theta} S_{l-1}^+ S_l^- + h.c.)), \quad (2.62)$$

donde, reconociendo la derivada discretizada como la diferencia entre sitios vecinos, la corriente de espín resulta (módulo una constante multiplicativa)

$$j_l = ie^{i\theta} \langle S_l^+ S_{l+1}^- \rangle + h.c. \quad (2.63)$$

Observamos que si bien la densidad de espín local $\rho_l = S_l^z$ no es una cantidad conservada, la densidad total o global sí lo es, ya que es inmediato probar a partir de las ecuaciones anteriores que $\partial_t \sum_{l=1}^L \rho_l = 0$.

2.2.2 Cristal de magnones

Consideramos el Hamiltoniano de Heisenberg sobre la cadena diente de sierra dado por

$$H = J \sum_l \left\{ \mathbf{S}_{l,1} \mathbf{S}_{l,2} + \mathbf{S}_{l,2} \mathbf{S}_{l,3} + \mathbf{S}_{l,3} \mathbf{S}_{l,4} + \mathbf{S}_{l,4} \mathbf{S}_{l+1,1} + \alpha \left(\mathbf{S}_{l,1} \mathbf{S}_{l,3} + \mathbf{S}_{l,3} \mathbf{S}_{l+1,1} \right) \right\} - h \sum_l \sum_{l'=1}^4 S_{l,l'}^z, \quad (2.64)$$

con $J = 1$, $l = 1, \dots, L$ indexando cada celda unidad, $\alpha \in \mathbb{R}$ un parámetro que regula la frustración y $l' = 1, \dots, 4$ indexando los espines dentro de cada celda unidad como muestra la figura 2.15-a). En el Hamiltoniano anterior estamos tomando una interacción de intercambio con anisotropía Δ , es decir

$$\mathbf{S}_{l,k} \mathbf{S}_{l',k'} = \frac{S_{l,k}^+ S_{l',k'}^- + S_{l,k}^- S_{l',k'}^+}{2} + \Delta S_{l,k}^z S_{l',k'}^z.$$

Para

$$\alpha = \alpha_c \equiv 1/\sqrt{2(1+\Delta)}, \quad (2.65)$$

el estado fundamental del sistema en un dado rango de campo magnético aplicado consiste en un arreglo periódico de magnones localizados en regiones finitas de la cadena, esquematizado en la figura 2.15-b), es decir, un cristal de magnones como el de la cadena Kagomé de la sección anterior. El estado fundamental puede encontrarse exactamente[25] y corresponde a un estado producto donde los magnones no interactúan entre sí.

La anisotropía en el modelo no afecta las propiedades del estado fundamental siempre que se verifique la condición (2.65). Sin embargo, para construir la teoría de campos efectiva es conveniente estudiar al sistema usando $\Delta < 1$ para evitar configuraciones clásicas colineales, que no permiten ser tratadas con nuestra formulación en términos de integrales de camino y estados coherentes de espín. Por esta razón en lo que sigue fijaremos $\Delta = 1/2$.

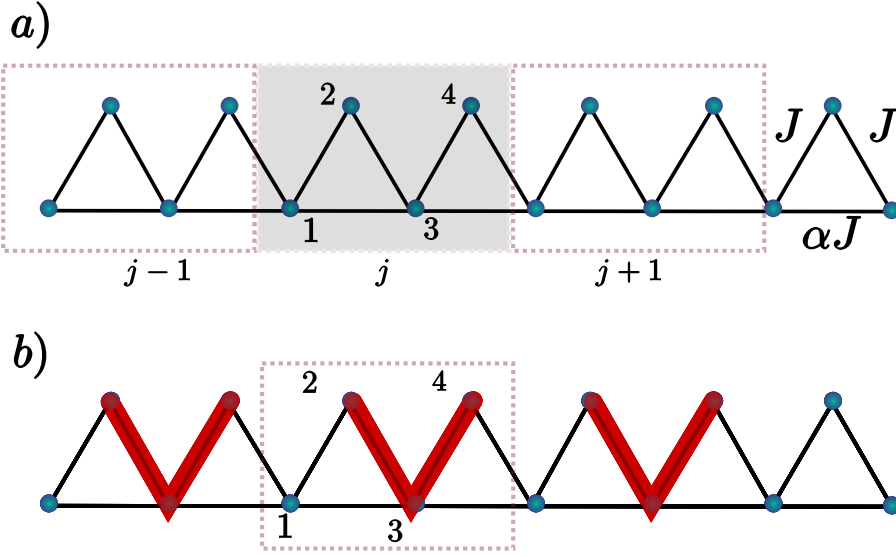


Figura 2.15: Esquema de la cadena diente de sierra y los acoplamientos elegidos. a) Celda unidad de cuatro sitios utilizada a lo largo del trabajo. (b) Representación del cristal de magnones. Las zonas rojas representan las zonas donde hay una probabilidad finita de encontrar un magnón, mientras que los sitios indexados con un 1 en cada celda unidad corresponden a espines completamente polarizados.

2.2.3 Teoría de campos efectiva

Para obtener la teoría de campos efectiva primero identificamos la configuración clásica de mínima energía. Para hacer esto comenzamos con un triángulo aislado formado por los espines \mathbf{S}_1 , \mathbf{S}_2 y \mathbf{S}_3 de la celda unidad (ver figura 2.15). Por simetría, tomamos los ángulos polares de \mathbf{S}_1 y \mathbf{S}_3 iguales, es decir, $\theta_1^0 = \theta_3^0 \equiv \theta_B$, y $\phi_2^0 - \phi_1^0 = \phi_2^0 - \phi_3^0 \equiv \phi^0$ para los ángulos azimutales. Para la cadena completa, en cada celda tomamos $\theta_2^0 = \theta_4^0 \equiv \theta_A$, y un único ángulo azimutal relativo entre los espines de índice par y los de índice impar.

Tomando $\Delta = 1/2$, el estado fundamental clásico en presencia de un campo magnético aplicado consiste en una configuración ‘inclinada’⁶, con $\phi^0 = \pi$ y ángulos polares dependientes del campo magnético aplicado, $\theta_{A,B}(h)$, como es usual[8, 7, 26]. Escribimos a los operadores de espín en términos de los ángulos polar y azimutal como vectores en \mathbb{R}^3 de longitud S según

⁶en inglés, *canted*, donde los espines no están en la dirección del eje z.

$$\mathbf{S}_{l,l'} = S(\cos \phi_{l,l'} \sin \theta_{l,l'}, \sin \phi_{l,l'} \sin \theta_{l,l'}, \cos \theta_{l,l'}). \quad (2.66)$$

donde S corresponde al número cuántico de espín, y parametrizamos las fluctuaciones cuánticas alrededor de la configuración clásica según

$$\begin{aligned} \phi_{l,l'} &= \phi_{l'}^0 + \phi_{l'}(x_l) \\ \theta_{l,l'} &= \theta_{l'}^0 + \theta_{l'}(x_l). \end{aligned} \quad (2.67)$$

Las variables conjugadas para contruir la teoría son

$$\begin{cases} \phi_{l'}(x_l) \\ a\Pi_{l'}(x_l) = -S\left(\theta_{l'}(x_l) \sin \theta_{l'}^0 + \frac{1}{2}\theta_{l'}(x_l)^2 \cos \theta_{l'}^0\right), \end{cases} \quad (2.68)$$

donde a es la distancia entre celdas. Los operadores de espín se escriben entonces

$$\begin{cases} S_{l,l'}^z = S \cos \theta_{l'}^0 + a\Pi_{l'}(x_l) \\ S_{l,l'}^\pm = S(-1)^{l'} e^{\pm i\phi_{l'}(x_l)} \left(\sin \theta_{l'}^0 - \frac{a\Pi_{l'}(x_l)}{S \tan \theta_{l'}^0} - \frac{1}{2} \frac{1}{1 - \cos^2 \theta_{l'}^0} \frac{a^2 \Pi_{l'}^2(x_l)}{S^2 \sin \theta_{l'}^0} \right). \end{cases} \quad (2.69)$$

Escribimos la teoría a orden cuadrático en fluctuaciones. Los términos de primer orden se anulan debido a que agregamos fluctuaciones sobre el estado fundamental clásico, y descartamos los términos constantes. La acción completa del sistema es $\mathcal{S} = \mathcal{S}_{cl} + \mathcal{S}_{BP}$. El primer término $\mathcal{S}_{cl} = \int d\tau H(\tau)$ corresponde a la acción clásica del sistema, y el segundo término \mathcal{S}_{BP} es el término geométrico o topológico de fase de Berry. La acción efectiva del sistema toma entonces la forma

$$\begin{aligned} \mathcal{S} &= \int dx d\tau \frac{J}{a} \left\{ \frac{S^2}{2} (\gamma - \alpha\beta) (a\partial_x \phi_1)^2 + \frac{S^2}{2} \phi_l (M_\phi)_{l,l'} \phi_{l'} + \frac{a^2}{2} \Pi_l (M_\Pi)_{l,l'} \Pi_{l'} \right\} + \\ & i \int dx d\tau \left\{ (\partial_\tau \phi_l) \frac{S - m_l}{a} - (\partial_\tau \phi_l) \Pi_l \right\}, \end{aligned} \quad (2.70)$$

donde estamos haciendo implícita la suma sobre índices repetidos $l, l' = 1, \dots, 4$, $m_l = \cos \theta_l^0$, $\beta = \sin^2 \theta_B$, $\gamma = \sin \theta_A \sin \theta_B$. M_ϕ y M_Π corresponden las matrices de masas de los campos ϕ_l y Π_l . Ambas matrices de masas pueden simetrizarse, y podemos entonces diagonalizar M_ϕ usando la transformación unitaria

$$U = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}, \quad (2.71)$$

tal que $\phi'_l = U_{l,l'} \phi_{l'}$ y $M'_\phi = U M_\phi U^t = \text{diag}(m_1, m_2, m_3, m_4)$, donde

$$m_1 = 0, \quad m_2 = 2\gamma, \quad m_3 = 4\gamma, \quad m_4 = 2(\gamma - 2\alpha\beta). \quad (2.72)$$

Las ecuaciones (2.71) y (2.72) nos dicen que el campo $\phi'_1 = (\phi_1 + \phi_2 + \phi_3 + \phi_4)/2$ no tiene masa. Esto corresponde a la simetría $U(1)$ del Hamiltoniano, siendo ϕ'_1 el modo de Goldstone correspondiente. Luego, los campos $\phi'_2 = (\phi_4 - \phi_2)/\sqrt{2}$ y $\phi'_3 = (-\phi_1 + \phi_2 - \phi_3 + \phi_4)/2$ tienen masa y en el límite de baja energía están ‘congelados’ y pueden aproximarse por cero. Finalmente, la combinación $\phi'_4 = \phi_3 - \phi_1$ puede tener una masa nula si

$$\alpha = \frac{\gamma}{2\beta}. \quad (2.73)$$

Si m_4 se anula, el campo ϕ'_4 es libre de fluctuar, es decir, se deslocaliza, y por lo tanto su campo conjugado $\Pi'_4 = \Pi_3 - \Pi_1$ se localiza. Si resulta energéticamente favorable para el sistema que el campo Π'_4 se localice en algún valor finito, entonces el proceso se denomina desbalance de espín (*spin imbalance*), ya que los campos Π_l parametrizan las fluctuaciones cuánticas en la magnetización. En efecto, $\Pi'_4 \neq 0$ corresponde a una diferencia en las magnetizaciones locales de los espines 1 y 3, rompiendo espontáneamente la simetría de la red. Este escenario es corroborado por nuestros cálculos numéricos en siguiente la subsección.

Por otra parte, si nos enfocamos en el término cinético, vemos que hay un único término, cuyo coeficiente $\gamma - \alpha\beta = \frac{m_4}{2} + \alpha\beta$ es mínimo cuando $m_4 = 0$. La banda chata puede interpretarse aquí como la anulación del valor renormalizado de este coeficiente, como mostramos al final de la sección 1.4. Más adelante veremos que esto tiene consecuencias importantes al agregar las interacciones antisimétricas.

2.2.4 Resultados numéricos (DMRG)

Comenzamos mostrando el estudio numérico del desbalance de espín presentado en la subsección anterior, usando DMRG. Todos nuestros cálculos numéricos están hechos con condiciones de contorno periódicas, salvo que se especifique lo contrario, y se llevaron adelante usando las librerías de ALPS. En el apéndice 5.1 presentamos un breve resumen de cómo funciona este algoritmo.

La figura 2.16 muestra la magnetización local $\langle S_l^z \rangle$ para cada espín en la celda unidad como función de la magnetización total normalizada del sistema, para $\alpha = \alpha_c$, en ausencia de interacciones antisimétricas. El valor más grande mostrado de m es $1/2$, donde el sistema está gapeado, previo al salto hacia saturación. Allí, el espín \mathbf{S}_1 está completamente polarizado. La ruptura de la simetría de traslación de la red puede ser claramente observada solo sobre el plateau. Aunque el fenómeno de la factorización exacta y el salto en la curva de magnetización es exclusivo al valor crítico $\alpha = \alpha_c$, es importante remarcar que el desbalance de espín y el plateau de magnetización no lo son, como muestra el panel interno de la figura 2.16. En efecto, el panel interno muestra nuevamente las magnetizaciones locales $\langle S_l^z \rangle$ para cada espín en la celda unidad como función de la magnetización total normalizada pero para $\alpha = \alpha_c + \frac{\alpha_c}{10} \equiv \tilde{\alpha}$. El sistema ya no tiene un salto en su curva de magnetización, pero el desbalance de espín sigue presente. Esto es natural desde el punto de vista de la teoría de campos, en donde se espera que exista una región finita del espacio de parámetros para la cual el valor renormalizado de m_4 se anule, y corresponde a la presencia de una fase de cristal de magnones, como aquella presente en la cadena Kagomé en la sección 2.1.4.

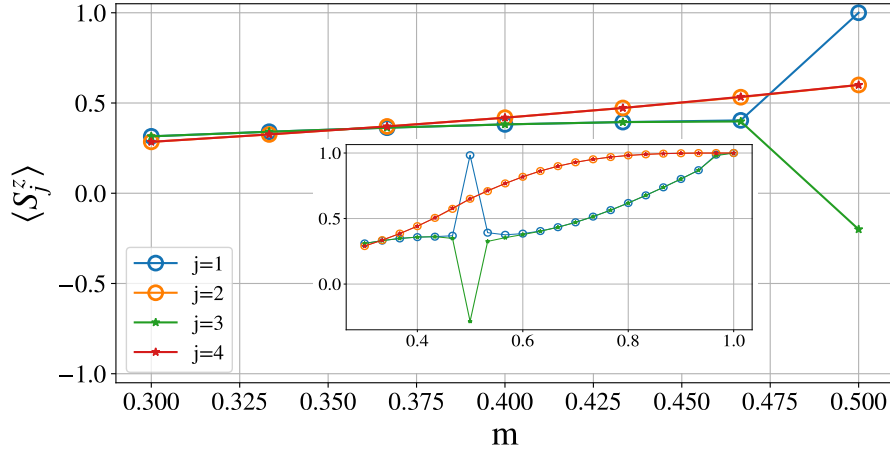


Figura 2.16: Panel principal: Magnetizaciones locales normalizadas en la celda unidad como funciones de la magnetización total normalizada para $D = 0$, $\alpha = \alpha_c$ y $N = 60$ sitios. El desbalance de espín entre los espines \mathbf{S}_1 y \mathbf{S}_3 ocurre en el plateau de magnetización, donde $m = 1/2$. Las magnetizaciones inferiores no están mostradas por claridad. Panel interno: mismo gráfico pero fijando $\alpha = \alpha_c + \frac{\alpha_c}{10} \equiv \tilde{\alpha}$. En ambas figuras $\langle S_2^z \rangle = \langle S_4^z \rangle$ y las curvas respectivas están superpuestas.

En lo que sigue, presentamos un estudio numérico de la corriente de espín del estado fundamental del sistema en presencia de interacciones antisimétricas. Agregamos entonces al Hamiltoniano (2.64) el siguiente término

$$H_{DM} = D\hat{z} \cdot \sum_l \left\{ \mathbf{S}_{l,1} \times \mathbf{S}_{l,2} + \mathbf{S}_{l,2} \times \mathbf{S}_{l,3} + \mathbf{S}_{l,3} \times \mathbf{S}_{l,4} + \mathbf{S}_{l,4} \times \mathbf{S}_{l+1,1} + \alpha \left(\mathbf{S}_{l,1} \times \mathbf{S}_{l,3} + \mathbf{S}_{l,3} \times \mathbf{S}_{l+1,1} \right) \right\} \quad (2.74)$$

que corresponde a la versión antisimétrica del Hamiltoniano original.

Calculamos numéricamente la energía del estado fundamental que corresponde a cada sector de magnetización para diferentes valores de D , para construir las curvas de magnetización del sistema. La figura 2.17 presenta los resultados, donde puede verse que en ausencia de interacciones antisimétricas la magnetización del estado fundamental como función del campo magnético aplicado presenta un salto desde $m = 1/2$ a saturación. En $m = 1/2$ el sistema presenta gap y el estado fundamental corresponde al cristal de magnones citado en 2.2.2. El plateau es consistente con el criterio OYA tomando una celda unidad de 4 sitios (ver figura 2.15). Aunque el plateau está presente para todos los valores de D estudiados, el ancho del plateau depende de D . Notablemente, a $m = 0.7$ no hay dependencia en D en la curva de magnetización.

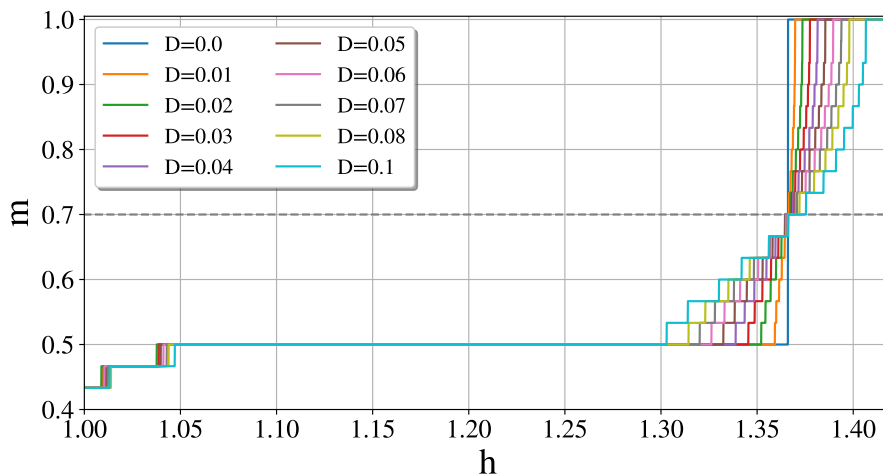


Figura 2.17: Magnetización del estado fundamental como función del campo magnético aplicado para diferentes valores de D , con $\alpha = \alpha_c$ y $N = 60$ sitios. Sectores de magnetización inferiores no se muestran por claridad.

En lo que sigue, estudiaremos las corrientes de espín del sistema definidas por

$$j_{l,l'} = ie^{i\theta_{l,l'}} \langle S_l^+ S_{l'}^- \rangle + h.c. \quad (2.75)$$

En las figuras 2.18, 2.19 y 2.20 fijamos la magnetización del sistema a diferentes valores. Es natural en este caso mapear el sistema de espines a un sistema de partículas que modelan las excitaciones magnéticas, y que pueden ser fermiónicas o bosónicas, en donde la magnetización fija se traduce en una cantidad fija de partículas, o un dado ‘llenado’.

La figura 2.18 muestra la corriente de espín obtenida numéricamente con DMRG como función de D para $\alpha = \alpha_c$ y magnetización fija. La nomenclatura ‘*Bottom spin current*’ corresponde a la corriente de espín que circula ‘por debajo’ en la cadena, es decir, entre los espines \mathcal{S}_1 y \mathcal{S}_3 en la celda unidad. Esta corriente tiene el signo opuesto a la corriente que conecta espines pares e impares en la celda unidad. Sobre el plateau de magnetización, $m = 1/2$, y para pequeños valores de D , el sistema presenta gap y la corriente de espín es lineal con D , o bien con el flujo magnético en lenguaje de partículas. Esta es la fase de Meissner, por su analogía con física de superconductores donde la corriente es proporcional al potencial vector⁷.

El panel incrustado de la figura 2.18 muestra cómo sobre el plateau la corriente deja de ser lineal con D para D suficientemente grande. Bajo el plateau de magnetización, para $m = 0.3$ y $m = 0.4$ el sistema está en su fase sin gap de líquido de Luttinger. Enfatizamos aquí que con condiciones de contorno periódicas el sistema forma un anillo. Luego, la presencia de D con una circulación finita a lo largo del anillo corresponde a la presencia de un flujo magnético para el sistema pensado como un sistema de partículas. Tomamos entonces de la teoría de líquido de Luttinger[28] el resultado de la corriente estacionaria de partículas inducida por un flujo magnético atravesando un anillo,

⁷Es interesante resaltar que en el efecto Meissner la corriente es expulsada del interior del superconductor para un campo externo suficientemente pequeño. Un sistema bosónico inmerso en un pequeño campo magnético sobre la geometría cuasi-1D conocida como escalera de dos patas presenta corrientes solo en las ‘patas’ de la escalera, y no sobre los escalones[27]. Al aumentar el campo magnético aplicado se encuentran arreglos de vórtices de corrientes, como en un superconductor de tipo II.

$$j = \frac{uK}{L} \left(\nu - 2 \frac{\Phi}{\Phi_0} \right),$$

donde u es la velocidad de las ondas de espín, K es el parámetro de Luttinger, L es el perímetro del anillo, $\Phi_0 = hc/e$ es un quantum de flujo, y $\nu = N_+ - N_-$ es la diferencia entre el número de partículas moviéndose a derecha e izquierda, N_+ y N_- respectivamente. Lo importante aquí es por un lado que en el equilibrio ν es elegido por el sistema para minimizar la energía, pudiendo variar solo por valores enteros; y por otro lado, puede probarse que la corriente estacionaria j es discontinua y tiene periodicidad Φ_0 como función del flujo magnético Φ .

Numéricamente podemos observar cómo las discontinuidades en la corriente de espín en la fase de líquido de Luttinger son una consecuencia de las condiciones de contorno periódicas. Como ya mencionamos más arriba, en lenguaje de partículas, para tener efectos del campo de gauge, el sistema de partículas equivalente a nuestro sistema de espines tiene que tener una circulación finita de potencial vector \mathbf{A} a lo largo de una curva cerrada.

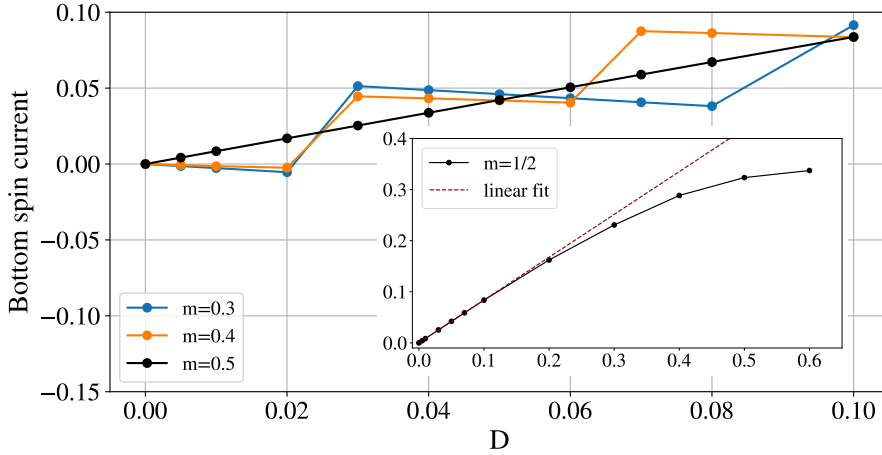


Figura 2.18: Panel principal: Corriente de espín como función de D , con $\alpha = \alpha_c$ y $N = 60$ sitios, a magnetización fija menor o igual a $m = 1/2$. Panel incrustado: Corriente de espín como función de D para $m = 1/2$ donde puede observarse el apartamiento del régimen lineal para un D suficientemente grande.

Para la cadena diente de sierra tenemos dos curvas cerradas con una circulación finita de D . Primero, cada triángulo (como el formado por los espines \mathbf{S}_1 , \mathbf{S}_2 y \mathbf{S}_3 en la celda unidad) y segundo, el sistema completo pensado como un anillo con condiciones de contorno periódicas. El primer camino hace que las corrientes de la fase de líquido de Luttinger de la figura 2.18 no sean periódicas en D . Es importante observar que el segundo camino es mucho más largo que el primero, por lo que en la figura 2.18 para D pequeño no alcanzamos a observar el primer salto en la corriente a causa de la estructura de celda. La figura 2.19 muestra la corriente de espín como función de D para $m = 0.3$ y $m = 0.4$ para condiciones de contorno abiertas y cerradas, mostrando que en este rango de D las discontinuidades en las corrientes de espín desaparecen con condiciones de contorno abiertas.

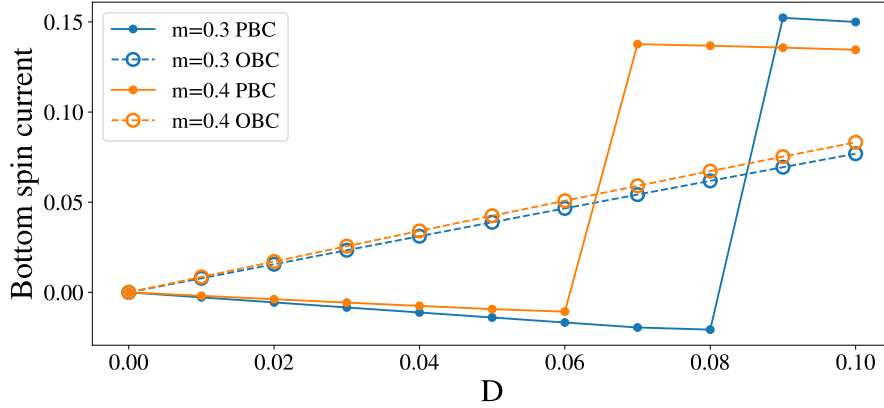


Figura 2.19: Corriente de espín como función de D a magnetización fija para un sistema de 20 sitios. En línea a trazos y continua, los resultados obtenidos usando condiciones de contorno abiertas y cerradas, respectivamente.

Ahora discutimos la situación a magnetización constante por encima del plateau en $m = 1/2$. La figura 2.20 muestra la corriente de espín como función de D para $\alpha = \alpha_c$ para $1/2 \leq m \leq 1$. Para $m = 1/2$ el sistema está en la fase de Meissner, y la curva se muestra como referencia. Para $1/2 < m < 1$ a $D = 0$ no hay corriente de espín, pero para D finito observamos un salto, consecuencia de la banda chata en el sistema. El salto es seguido de un régimen aproximadamente lineal, que corresponde a lo que se denomina la contribución diamagnética a la corriente, debido a que la pendiente de cada recta es proporcional a la densidad de bosones (observar que la pendiente disminuye conforme m aumenta). Con $\alpha = \alpha_c$ y $D = 0$ el sistema de partículas tiene una banda chata y el sistema de espines no accede a ningún estado con $1/2 < m < 1$ porque es energéticamente favorable estar o bien en el plateau de magnetización de $m = 1/2$ o bien saturado, $m = 1$, dependiendo del campo magnético h aplicado. En la figura 2.20 cómo a magnetización constante $1/2 < m < 1$ la corriente de espín resulta discontinua como función de D en el origen.

En el panel incrustado de la figura 2.20 mostramos un gráfico igual al del panel principal donde solo modificamos el valor de α , corriéndolo desde α_c a un valor arbitrario $\tilde{\alpha} = \alpha_c + \frac{\alpha_c}{10}$. Para $\alpha = \tilde{\alpha}$ el sistema no tiene una banda chata a $D = 0$ y las corrientes de espín resultan regulares en el origen como función de D .

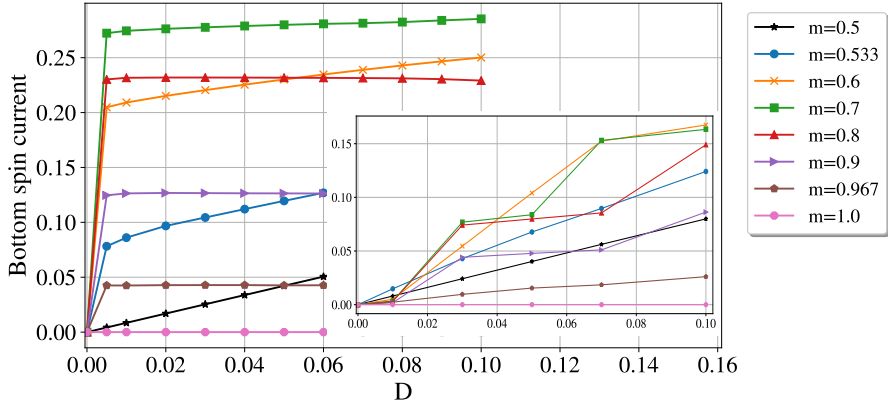


Figura 2.20: Panel principal: Corriente de espín como función de D con $\alpha = \alpha_c$, $N = 60$ sitios a magnetización fija. Panel incrustado: misma figura eligiendo $\alpha = \tilde{\alpha} = \alpha_c + \frac{\alpha_c}{10}$

Si ahora dejamos la magnetización libre y estudiamos la corriente de espín del estado fundamental como función del campo magnético externo h aplicado, el comportamiento puede compararse con lo presentado anteriormente. La figura 2.21 muestra la corriente de espín para el estado fundamental con $\alpha = \alpha_c$ donde pueden observarse claramente tres regímenes. Para $0 < h < h_1(D)$ el sistema está en su fase sin gap de líquido de Luttinger, con $m < 1/2$. La corriente de espín es discontinua como función de h a causa de las condiciones de contorno periódicas del sistema. Para $h_1 < h < h_2$ el sistema está en la fase de Meissner, con $m = 1/2$. Puede verse que el espaciado entre curvas es constante debido a que la diferencia entre valores de D elegidos en el grafico es constante y la corriente es lineal con D sobre el plateau. Finalmente, para $h_2 < h < h_{sat}$ el sistema está en lo que denominamos una fase de pequeña rigidez (en inglés, *low-stiffness phase*), con $1/2 < m < 1$. La teoría de campos efectiva de la sección anterior permite tener una intuición de por qué la corriente de espín es mayor a campo magnético grande. La corriente de espín en la teoría de campos se escribe $j(x) \propto \partial_D \mathcal{H} \propto \partial_x \phi_1$. Para $\alpha = \alpha_c$ y D finito, la rigidez (el coeficiente del término cinético en la teoría) se vuelve finita pero pequeña, y el costo energético de tener corrientes de espín es pequeño.

El pico alrededor de $h = 1.36$ corresponde al máximo de la corriente de espín para $m = 0.7$ en la figura 2.20. En el mismo valor de magnetización todas las curvas de magnetización se cruzan en la figura 2.17

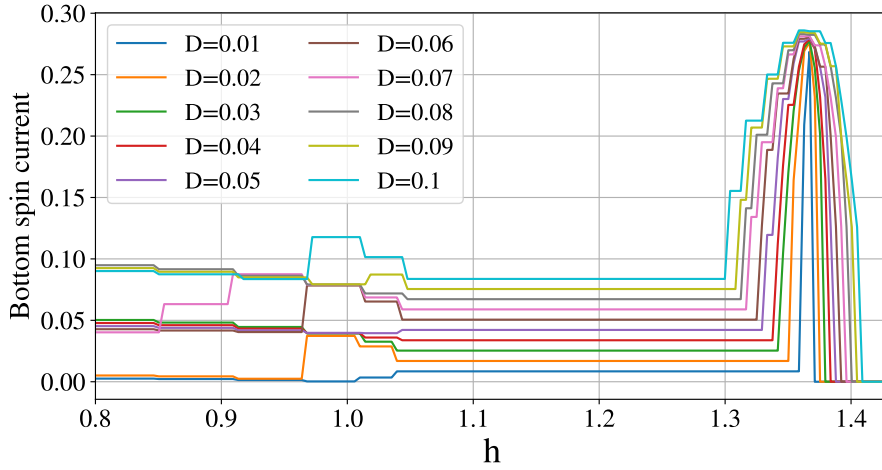


Figura 2.21: Corriente de espín del estado fundamental como función del campo magnético externo aplicado h , para diferentes valores de D , con $\alpha = \alpha_c$ y $N = 60$ sitios.

Bandas chatas en cadenas tipo Kagomé

La figura 2.22 muestra tres cadenas de espín que presentan un estado fundamental factorizado exacto correspondiente a un cristal de magnones. La figura *a*) es la cadena diente de sierra estudiada previamente, colocada para comparar, mientras que las cadenas *b*) y *c*) son dos cadenas tipo Kagomé presentadas en la referencia [25]. Allí muestran que para acoplamientos magnéticos específicos, elegidos en esta sección, la magnetización presenta saltos hacia saturación de magnitudes $\delta m = 1/3$ y $\delta m = 1/2$ para las cadenas *b*) y *c*) de la figura 2.22, respectivamente.

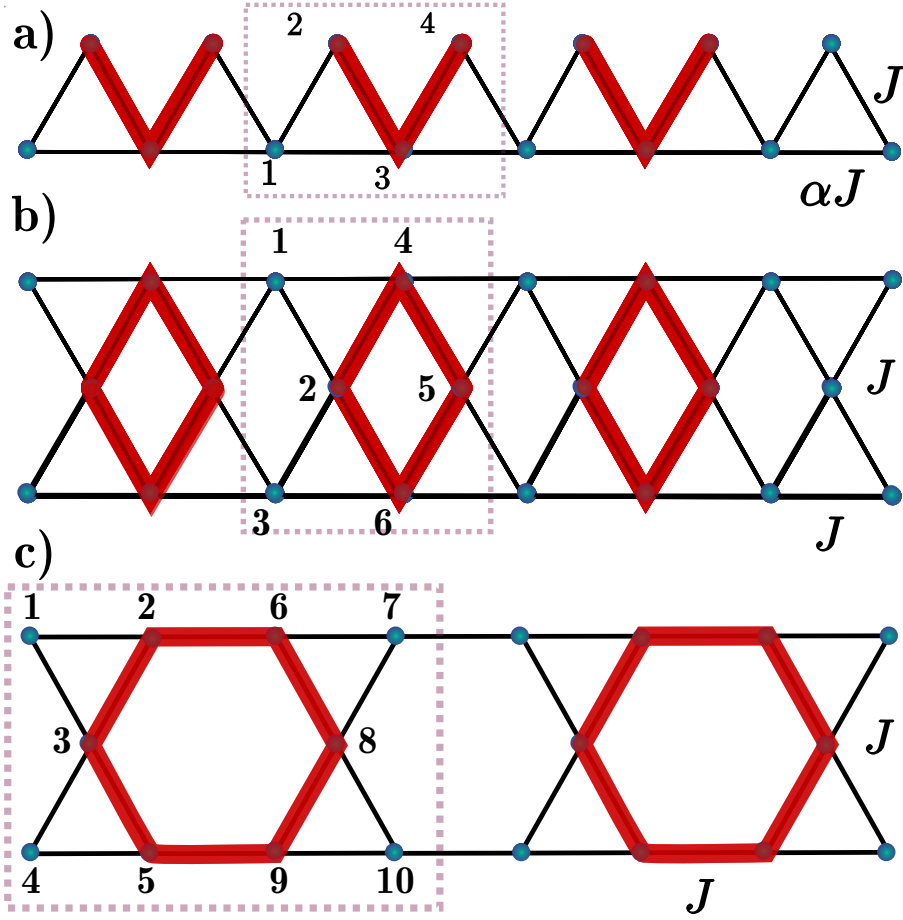


Figura 2.22: Cadenas de espín que para $\alpha = \alpha_c$ y $J' = J(2\Delta + 1)/(\Delta + 1)$ exhiben banda chata. En a) la cadena diente de sierra, por comparación. En b) y c) dos cadenas de tipo Kagomé. con línea punteada marcamos la celda unidad que rompe la simetría de traslación de la red subyacente. En trazo rojo grueso, marcamos los sitios donde hay una probabilidad finita de encontrar un magnón en cada cadena en presencia de la banda chata. Con números, mostramos la indexación utilizada para los espines dentro de cada celda unidad.

Siguiendo los pasos del caso de la cadena Sawtooth, estudiamos la presencia de modos deslocalizados escribiendo las matrices de masa de las cadenas de tipo Kagomé de la figura 2.22. Consideramos primero la cadena b). Por simetría, tomamos el anzats para los ángulos polares clásicos dado por $\theta_1^{(b)} = \theta_3^{(b)} = \theta_4^{(b)} = \theta_6^{(b)} \equiv \theta_A^{(b)}$ y $\theta_2^{(b)} = \theta_5^{(b)} \equiv \theta_B^{(b)}$. Para los ángulos azimutales clásicos, tomamos $\phi_2^{(b)} = \phi_5^{(b)}$ y $\phi_l^{(b)} - \phi_2^{(b)} \equiv \phi^{(b)}$, para $l = 1, 3, 4, 6$. La cadena tipo Kagomé de la figura 2.22-b) presenta un plateau de magnetización con magnones localizados en $m = 2/3$. El estado fundamental clásico para este sector de magnetización tiene nuevamente $\phi^{(b)} = \pi$. Agregando fluctuaciones clásicas y tomando el límite continuo, construimos la matriz de masas $M_\phi^{(b)}$ para los campos $\phi_l^{(b)}$. Como antes, esta matriz simétrica se diagonaliza por la transformación unitaria

$$P^{(b)} = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{2\sqrt{3}} \end{pmatrix}, \quad (2.76)$$

dando

$$M_\phi^{(b)'} = P M_\phi^{(b)} P^t = \text{Diag}(0, m_2^{(b)}, m_3^{(b)}, m_4^{(b)}, m_5^{(b)}, m_6^{(b)}),$$

donde

$$\begin{cases} m_2^{(b)} = m_3^{(b)} = 2(\gamma^{(b)} - 2\beta^{(b)}) \\ m_4^{(b)} = 2\gamma^{(b)} \\ m_5^{(b)} = 4\gamma^{(b)} \\ m_6^{(b)} = 6\gamma^{(b)} \end{cases} \quad (2.77)$$

con $\beta^{(b)} = \sin^2 \theta_B^{(b)}$ y $\gamma^{(b)} = \sin \theta_A^{(b)} \sin \theta_B^{(b)}$. Los campos transformados son $\phi_l'^{(b)} = \sum_{l'} P_{l,l'}^{(b)} \phi_{l'}^{(b)}$. Las ecuaciones (2.76) y (2.77) muestran que las combinaciones $\phi_j'^{(b)}$ son masivas para $j = 4, 5, 6$, por lo que en el límite de baja energía las podemos fijar a cero.

El campo sin masa $\phi_1'^{(b)}$ representa como antes al bosón de Goldstone asociado a la simetría $U(1)$ del Hamiltoniano. La masa $m_2^{(b)}$ puede anularse para valores de $\theta_A^{(b)}$ y $\theta_B^{(b)}$ tales que $\gamma^{(b)} = 2\beta^{(b)}$. Luego, los campos $\phi_2'^{(b)} = (\phi_6^{(b)} - \phi_3^{(b)})/\sqrt{2}$ y $\phi_3'^{(b)} = (\phi_4^{(b)} - \phi_1^{(b)})/\sqrt{2}$ se deslocalizan simultáneamente. Esto lleva nuevamente a un mecanismo de desbalance de espín que describe a las excitaciones localizadas en el sistema (ver figura 2.22).

Para la estructura de tipo Kagomé de la figura 2.22-c) tomamos un estado fundamental clásico con una celda unidad de 10 sitios. Los ángulos polares del estado fundamental son $\theta_l^{(c)} \equiv \theta_A^{(c)}$ para $l \in \{1, 2, 4, 5, 6, 7, 9, 10\}$, y $\theta_3^{(c)} = \theta_8^{(c)} \equiv \theta_B^{(c)}$. Los ángulos azimutales son $\phi_l^{(c)} \equiv \phi^{(c)}$ para $l \in \{1, 2, 4, 5, 8\}$ y $\phi_{l'}^{(c)} - \phi^{(c)} = \pi$ para $l' \in \{3, 6, 7, 9, 10\}$.

Siguiendo los pasos de los casos previos, construimos la matriz de masas para los campos $\phi_l^{(c)}$. Puede verse que las combinaciones $\phi_1^{(c)} - \phi_2^{(c)}$, $\phi_4^{(c)} - \phi_5^{(c)}$, $\phi_7^{(c)} - \phi_6^{(c)}$, $\phi_{10}^{(c)} - \phi_9^{(c)}$ pueden simultáneamente volverse no masivos eligiendo apropiadamente $\theta_A^{(c)}$ y $\theta_B^{(c)}$, en el límite de baja energía. La deslocalización de estos modos típicamente lleva a una región en el espacio de parámetros con desbalance de espín, dentro de la cual el estado producto de magnones localizados es un punto particular.

Los modos deslocalizados en ambas cadenas de tipo Kagomé son descriptos como en la cadena diente de sierra, a causa de la banda chata presente en los tres sistemas. Más aún, la banda chata significa una rigidez nula. Entonces, la corriente de espín en estas cadenas de tipo Kagomé debe también ser discontinua en $D = 0$ a magnetización fija para sectores previos a saturación, y una fase de baja rigidez es también esperable.

2.2.5 Conclusiones

Hemos estudiado a la cadena diente de sierra en el caso donde la frustración induce una banda chata, y cómo las interacciones antisimétricas afectan este sistema. A través de una teoría

de campos semiclásica, hemos descripto las excitaciones localizadas del sistema a través de un mecanismo de desbalance de espín. Un aspecto central de este mecanismo es la presencia de modos angulares deslocalizados, cuya presencia se manifiesta al diagonalizar la matriz de masas de la teoría, que exhibe autovalores nulos. Esta matriz puede ser evaluada para diferentes geometrías, incluso en dimensiones⁸ mayores a 1, donde también pueden encontrarse masas nulas. En particular, lo hemos hecho para dos cadenas de tipo Kagomé en las cuales es conocido que pueden presentar banda chata, mostrando la relación entre los modos deslocalizados y las correspondientes excitaciones magnéticas localizadas.

Por otra parte, estudiamos numéricamente la corriente de espín sobre la cadena diente de sierra introduciendo interacciones antisimétricas en el Hamiltoniano, y encontramos tres regímenes diferentes. A valores pequeños del campo magnético aplicado, el sistema está en la fase Líquido de Luttinger. Aquí la corriente de espín muestra saltos como función del parámetro D . Este comportamiento se explica dentro de la teoría de Líquido de Luttinger. A campos magnéticos intermedios, la magnetización (normalizada) presenta un plateau en $m = 1/2$. En el plateau, el estado fundamental es un cristal de magnones con gap en ausencia de interacciones antisimétricas. Para D finito, la corriente de espín es constante como función de h sobre el plateau y proporcional a D como consecuencia del término diamagnético de la corriente de espín. Esta es la fase de Meissner. Finalmente, a campo magnético alto, la dependencia en D es notoria. A $D = 0$ hay un salto en la curva de magnetización. Los sectores magnéticos entre $m = 1/2$ y $m = 1$ son salteados por el sistema. Entonces para $D = 0$ no hay fase de alto campo magnético. Sin embargo, al agregar un pequeño D la curva de magnetización es suave como función de h . La corriente de espín del estado fundamental es en este caso también suave, y tiene un pico alrededor de $h/J = 1.36$. En términos de la teoría de campos efectiva, el sistema no tiene modos delocalizados y tiene una rigidez finita aunque pequeña. Si en cambio la magnetización se mantiene fija en $1/2 < m < 1$, entonces la corriente de espín presenta un salto en $D = 0$, como consecuencia de la banda chata.

2.3 Bicapa hexagonal

En esta sección abordamos el estudio semiclásico del Hamiltoniano de Heisenberg en una geometría bidimensional que denominamos bicapa hexagonal⁹, y que introduciremos a la brevedad. Antes de comenzar, destacamos que en dos dimensiones varias de las técnicas más exitosas en el estudio de sistemas cuánticos en una dimensión fallan, como es el caso de DMRG y de bosonización. Por otra parte, Monte Carlo Cuántico no es aplicable a sistemas frustrados, y diagonalización exacta solo permite alcanzar tamaños que típicamente son demasiado pequeños como para extraer información precisa. Por estas razones, nuestro enfoque semiclásico en términos de teorías de campos efectivas de baja energía es especialmente valioso en dimensiones mayores a 1.

⁸En la sección siguiente abordaremos esta situación para la denominada ‘bicapa hexagonal’.

⁹En esta tesis llamaremos hexagonal a la estructura geométrica bidimensional también conocida como panal de abejas.

2.3.1 Introducción

$\text{Bi}_3\text{Mn}_4\text{O}_{12}(\text{NO}_3)$ consiste en un arreglo de bicapas de redes hexagonales no distorcionadas formadas por iones Mn^{+4} de espín $3/2$, como mostramos en la figura 2.23-a) y 2.23-b). En ausencia de campo magnético el material no muestra orden de largo alcance, al menos a temperaturas mayores que 0.4 K[29], lo que sugiere la presencia de fuerte frustración. Al aplicar un campo magnético, el sistema adquiere orden antiferromagnético de largo alcance alrededor de los 6 T[29].

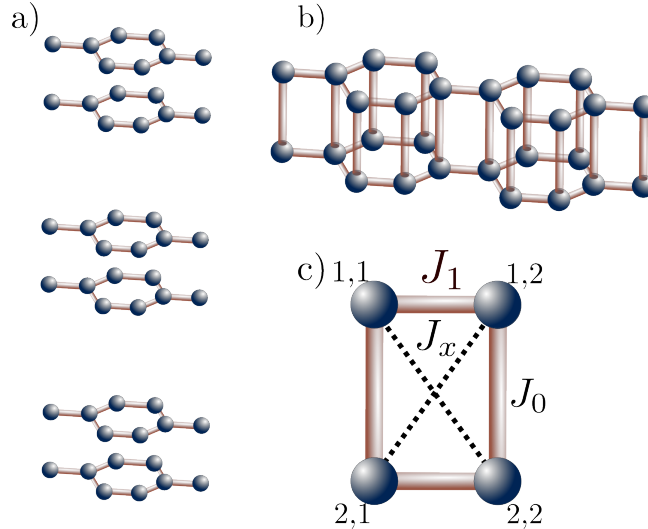


Figura 2.23: La geometría de bicapa considerada. a) esquematiza el arreglo de bicapas presente en el material. b) muestra una bicapa aislada. c) muestra la celda unidad de cuatro espines, señalando la indexación utilizada para cada sitio en cada capa y cada acoplamiento.

A través de mediciones de dispersión inelástica de neutrones sobre $\text{Bi}_3\text{Mn}_4\text{O}_{12}(\text{NO}_3)$ en presencia de campo magnético, se han estimado los acoplamientos magnéticos para un modelo de Heisenberg de espines $3/2$ con interacciones a primeros vecinos[30]. Estos son consistentes con aquellos determinados por calculos ab initio usando DFT[31], y sugieren que no hay frustración significativa en el plano del arreglo hexagonal sino que la interacción entre capas probablemente juega un papel importante en desestabilizar el orden magnético. La dispersión de neutrones muestra la presencia de orden antiferromagnético (Néel) de corto alcance a bajas temperaturas y la presencia de una transición de fase magnética en donde el corto alcance se expande en orden de largo alcance[29]. Los resultados experimentales sugieren que el estado colineal de Néel se vuelve más estable al aumentar levemente la temperatura, es decir, las fluctuaciones térmicas estabilizan el orden antiferromagnético[29] en lo que se conoce como orden por desorden.

Modelo

Motivados por los resultados experimentales, consideraremos el Hamiltoniano

$$\begin{aligned}
H = & \sum_{\mathbf{r}} \sum_{i=0}^2 \left\{ \frac{J_0}{3} (\mathbf{S}_{1,1}(\mathbf{r}) \cdot \mathbf{S}_{2,1}(\mathbf{r}) + \mathbf{S}_{1,2}(\mathbf{r}) \cdot \mathbf{S}_{2,2}(\mathbf{r})) \right. \\
& + J_1 (\mathbf{S}_{1,1}(\mathbf{r}_i) \cdot \mathbf{S}_{1,2}(\mathbf{r}) + \mathbf{S}_{2,1}(\mathbf{r}_i) \cdot \mathbf{S}_{2,2}(\mathbf{r})) \\
& \left. + J_x (\mathbf{S}_{1,1}(\mathbf{r}_i) \cdot \mathbf{S}_{2,2}(\mathbf{r}) + \mathbf{S}_{2,1}(\mathbf{r}_i) \cdot \mathbf{S}_{1,2}(\mathbf{r})) \right\} \quad (2.78)
\end{aligned}$$

$$- h \sum_{\mathbf{r}} \left\{ S_{1,1}^z(\mathbf{r}) + S_{1,2}^z(\mathbf{r}) + S_{2,1}^z(\mathbf{r}) + S_{2,2}^z(\mathbf{r}) \right\}, \quad (2.79)$$

donde $J_1 > 0$ conecta espines en cada capa hexagonal, $J_0 > 0$ conecta capas entre sí, $J_x > 0$ regula la frustración entre capas y h es el campo magnético externo aplicado. El eje z corresponde a la dirección del campo magnético, $i = 0, 1, 2$ corresponde a $\mathbf{r}_{(0,1,2)} = \mathbf{r} + (\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2)$, siendo \mathbf{e}_1 y \mathbf{e}_2 los vectores primitivos de la red triangular. Con nuestra notación, el término proporcional a J_0 está contado 3 veces, por lo que introducimos el factor $1/3$. Para $\text{Bi}_3\text{Mn}_4\text{O}_{12}(\text{NO}_3)$, los resultados numéricos y experimentales sugieren que J_1 es el acoplamiento dominante[30], mientras que J_0 es aproximadamente tres veces más grande que J_x .

2.3.2 Teoría de campos efectiva

Escribimos una teoría de campos efectiva usando como antes el formalismo de integrales de camino en términos de estados coherentes de espín. Nuevamente los espines están representados por vectores de \mathbb{R}^3 de módulo S , $\mathbf{S} = S(\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta)$, e introducimos las fluctuaciones cuánticas sobre la configuración clásica de mínima energía. Nuestra parametrización para la configuración clásica de mínima energía en presencia de un campo magnético externo consiste en una configuración de Néel inclinada, donde para cada celda unidad fijamos

$$\begin{aligned}
\phi_{l,\eta}^0(\mathbf{r}) &= \pi(l + \eta) \\
\theta_{l,\eta}^0(\mathbf{r}) &= \theta^0(h, J_0, J_1),
\end{aligned} \quad (2.80)$$

con $l = 1, 2$ y $\eta = 1, 2$. Esta configuración corresponde al límite de baja frustración $J_x \ll J_0, J_1$. La energía clásica del sistema a $T = 0$ se minimiza entonces por $\cos \theta_0 = h/2(J_0 + 3J_1)$. En el límite continuo las fluctuaciones cuánticas se agregan sobre el estado fundamental clásico cambiando $\phi_{l,\eta}^0(x, y) \rightarrow \pi(l + \eta) + \phi_{l,\eta}(x, y)$ y $\theta_{l,\eta}^0(x, y) \rightarrow \theta^0 + \delta\theta_{l,\eta}(x, y)$. Los campos canónicamente conjugados de la teoría son $\phi_{l,\eta}$ y $a\Pi_{l,\eta} = -S(\delta\theta \sin \theta^0 + \frac{1}{2}(\delta\theta)^2 \cos \theta_0)$, donde a es la distancia entre espines conectados por J_1 en la celda unidad. Los operadores de espín escritos a orden cuadrático en $\phi_{l,\eta}$ y $\Pi_{l,\eta}$ son

$$\begin{aligned}
S_{l,\eta}^z &= S \cos \theta_0 + a\Pi_{l,\eta} \\
S_{l,\eta}^\pm &= (-1)^{l+\eta} e^{\pm i\phi_{l,\eta}} \left(S \sin \theta^0 - \frac{m}{S \sin \theta^0} a\Pi_{l,\eta} - \frac{1}{2} \frac{S^2}{S^2 - m^2} \frac{1}{S \sin \theta^0} (a\Pi_{l,\eta})^2 \right),
\end{aligned} \quad (2.81)$$

donde $m = S \cos \theta_0$. Reemplazando (2.81) en el Hamiltoniano y conservando términos a orden cuadrático obtenemos la acción efectiva

$$\mathcal{S} = \mathcal{S}_{cl} + \mathcal{S}_{BP}, \quad (2.82)$$

donde \mathcal{S}_{cl} es la acción clásica del sistema que contiene los términos cinéticos y de masas, es decir, $\mathcal{S}_{cl} = \mathcal{S}_K + \mathcal{S}_M$. Por comodidad, renombramos a los campos con un único índice según

$$\phi_{i,j} \rightarrow \phi_{i(i-1)+j}, \quad (2.83)$$

y obtenemos para la acción efectiva las siguientes expresiones:

$$\begin{aligned} \mathcal{S}_K &= \int d\tau \frac{d^2x}{\nu} \frac{K}{2a^2} \sum_{j=1}^2 \left((\mathbf{e}_j \cdot \nabla \phi_1)^2 + (\mathbf{e}_j \cdot \nabla \phi_3)^2 \right) \\ &= \int d\tau \frac{d^2x}{\nu} \frac{K}{2} \frac{3}{2} \left(3(\partial_x \phi_1)^2 + 3(\partial_x \phi_3)^2 + (\partial_y \phi_1)^2 + (\partial_y \phi_3)^2 \right), \end{aligned} \quad (2.84)$$

con $\tau \in [0, \beta]$ el tiempo imaginario, β la inversa de la temperatura (donde la constante de Boltzmann es absorbida), $\nu = (9\sqrt{3}a^2)/2$, $K = (Sa)^2(J_1 - J_x) \sin^2(\theta^0)$, $\mathbf{e}_j = a\frac{\sqrt{3}}{2}(\sqrt{3}, (-1)^{j+1})$, $j = 1, 2$, y

$$\mathcal{S}_M = \int d\tau \frac{d^2x}{\nu} \left(\frac{1}{2} \phi_i (M_\phi)_{ij} \phi_j + \frac{a^2}{2} \Pi_i (M_\Pi)_{ij} \Pi_j \right), \quad (2.85)$$

donde estamos usando la notación de Einstein para índices repetidos, y M_ϕ y M_Π son matrices de masa simétricas. El segundo término en (2.82) es la fase de Berry, dada por

$$\mathcal{S}_{BP} = -i(S - m) \int d\tau \frac{d^2x}{\nu} \sum_j \partial_\tau \phi_j + i \int d\tau \frac{d^2x}{\nu} \sum_j (\partial_\tau \phi_j) a \Pi_j. \quad (2.86)$$

La matriz de masas M_ϕ es diagonalizada por la transformación

$$W = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{pmatrix}, \quad (2.87)$$

donde $\phi'_j = \sum_{k=1}^4 W_{jk} \phi_k$, $j = 1, \dots, 4$ y $M'_\phi = 4WM_\phi W^{-1} = \text{diag}(m_1, m_2, m_3, m_4)$, con

$$\begin{cases} m_1 = 0 \\ m_2 = 8S^2 \sin^2(\theta^0)(J_0 + 3J_1) \\ m_3 = 8S^2 \sin^2(\theta^0)(J_0 - 3J_x) \\ m_4 = 24S^2 \sin^2(\theta^0)(J_1 - J_x). \end{cases} \quad (2.88)$$

El anulamiento de m_1 vale a todo orden en fluctuaciones a causa de la simetría $U(1)$

del Hamiltoniano. La misma transformación diagonaliza también a M_{Π} , es decir, $M'_{\Pi} = 4WM_{\Pi}W^{-1} = \text{diag}(\mu_1, \mu_2, \mu_3, \mu_4)$ con

$$\begin{cases} \mu_1 = 8(J_0 + 3J_1) \\ \mu_2 = 8(J_0 + 3J_1) \cot^2(\theta^0) \\ \mu_3 = 8(-J_0 + 3J_1 + (J_0 - 3J_x) \csc^2(\theta^0)) \\ \mu_4 = 8(J_0 - 3J_1 + 3(J_1 - J_x) \csc^2(\theta^0)). \end{cases} \quad (2.89)$$

La acción efectiva completa luego de la transformación W toma la forma

$$\begin{aligned} \mathcal{S} = \int d\tau \frac{d^2x}{\nu} \left\{ \frac{K}{2} 3 \left(3(\partial_x \phi'_2 - \partial_x \phi'_3)^2 + 3(\partial_x \phi'_1 - \partial_x \phi'_4)^2 + (\partial_y \phi'_2 - \partial_y \phi'_3)^2 + (\partial_y \phi'_1 - \partial_y \phi'_4)^2 \right) + \right. \\ \left. \frac{1}{2} \sum_{j=2}^4 \left(m_j \phi_j'^2 \right) + \frac{1}{2} \sum_{j=1}^4 \left(\mu_j (a\Pi_j')^2 \right) - 4i(S - m)(\partial_\tau \phi'_1) + 4i \sum_{j=1}^4 (\partial_\tau \phi'_j) a\Pi_j' \right\}. \end{aligned} \quad (2.90)$$

En general, todos los términos de masa de los campos Π_j' son no nulos, y por lo tanto sus fluctuaciones son penalizadas energéticamente. Estos campos pueden integrarse exactamente (la integral es Gaussiana y hay que completar cuadrados), dando lugar a la acción

$$\begin{aligned} \mathcal{S} = \int d\tau \frac{d^2x}{\nu} \left\{ \frac{K}{2} 3 \left(3(\partial_x \phi'_2 - \partial_x \phi'_3)^2 + 3(\partial_x \phi'_1 - \partial_x \phi'_4)^2 + (\partial_y \phi'_2 - \partial_y \phi'_3)^2 + (\partial_y \phi'_1 - \partial_y \phi'_4)^2 \right) + \right. \\ \left. \frac{1}{2} \sum_{j=2}^4 \left(m_j \phi_j'^2 \right) + \frac{1}{2} \sum_{j=1}^4 \left(\frac{16}{\mu_j} (\partial_\tau \phi_j')^2 \right) - 4i(S - m)(\partial_\tau \phi'_1) \right\}. \end{aligned} \quad (2.91)$$

Para $J_x \ll J_0, J_1$ el sistema está débilmente frustrado y tenemos masas no nulas $m_i > 0$, $\mu_j > 0$, $i = 2, \dots, 4$, $j = 1, \dots, 4$. Como mencionamos antes, el campo ϕ'_1 no tiene masa, y permanece así para $4(S - m) \notin \mathbb{Z}$ y $T = 0$.

El factor $1/4$ en la transformación (2.87) es elegido para darle al campo ϕ'_1 la correcta periodicidad [7]. Si los campos ϕ'_2 , ϕ'_3 y ϕ'_4 son masivos, pueden fijarse a cero en el límite de baja energía, lo que implica $\phi_1 = \phi_2 = \phi_3 = \phi_4$. En este caso la acción depende solo del campo sin masa $\phi'_1 = \phi_1$. Así, la acción (2.91) corresponde a un modelo XY clásico con un término geométrico adicional. Si $T = 0$, la teoría está en tres dimensiones, mientras que para temperatura finita $0 \leq \tau \leq \beta < \infty$, y la teoría es 2-dimensional, donde el teorema de Mermin-Wagner prohíbe el orden de largo alcance.

La condición $4(S - m) \notin \mathbb{Z}$ corresponde a la presencia de una fase de Berry no trivial. Para $T = 0$, los vórtices son el único mecanismo disponible para destruir el orden antiferromagnético de largo alcance para las componentes de espín en el plano representadas por el campo ϕ'_1 . Como los vórtices están prohibidos para valores genéricos de la magnetización, el orden de largo alcance se preserva a $T = 0$. A temperatura finita, el largo alcance se transforma en cuasi largo alcance. Si $4(S - m)$ es entero la fase de Berry se vuelve trivial y puede desecharse. Los vórtices de ϕ'_1 pueden proliferar si eso es energéticamente favorable (es decir, si la rigidez es

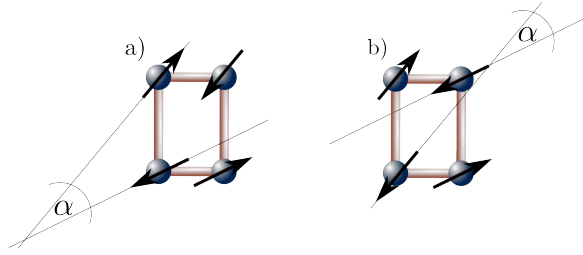


Figura 2.24: Representación esquemática de los modos cero correspondientes a a) $J_x = J_0/3$ y b) $J_1 = J_x$.

suficientemente pequeña), desordenando al campo ϕ'_1 y abriendo un gap en el sistema. La fase resultante tiene orden antiferromagnético de corto alcance.

La situación descrita aquí cambia cuando los acoplamientos generan competencia entre términos del Hamiltoniano. A continuación discutimos qué ocurre cuando el acoplamiento en el plano de los arreglos hexagonales es el acoplamiento dominante.

2.3.3 Orden por desorden

Si J_1 es el acoplamiento dominante, como en $\text{Bi}_3\text{Mn}_4\text{O}_{12}(\text{NO}_3)$ [30], al incrementar J_x encontramos que $m_3 = 0$ para $J_x = J_0/3$ (ver ecuación (2.88)), mientras que $\mu_j > 0$, $j = 1, \dots, 4$.

Es fácil ver que la condición $J_x = J_0/3$ implica no solo una masa nula en la teoría efectiva sino que también corresponde a un modo cero del Hamiltoniano usando la parametrización (2.80). En efecto, si tomamos el Hamiltoniano (2.79) evaluado en $J_x = J_0/3$, podemos hacer una variación en la parametrización clásica (2.80) introduciendo los parámetros reales $\alpha_{l,\eta}$ según

$$\phi_{l,\eta}^0 = (l + \eta)\pi + \alpha_{l,\eta}, \quad (2.92)$$

donde $l = 1, 2$, $\eta = 1, 2$. Si tomamos $\alpha_{l,1} = \alpha_{l,2}$, $l = 1, 2$, entonces el Hamiltoniano es exclusivamente función de θ^0 . Como consecuencia de la simetría $U(1)$ del sistema podemos fijar $\alpha_{2,1} = 0$, dejando un único parámetro libre $\alpha_{1,1} \equiv \alpha$. El parámetro α corresponde al ángulo relativo entre espines en las dos capas hexagonales como indicamos en la figura 2.24-a), y entra en la teoría efectiva solo a través del término cinético, ya que es un modo cero de la teoría para fluctuaciones uniformes en espacio. Aquí hacemos la aproximación de baja energía $\phi'_2 = \phi'_4 = 0$ para los campos masivos, que implica $\phi_2 = \phi_1$ y $\phi_4 = \phi_3$, y estudiamos la teoría sin masa

$$\mathcal{S} = \sqrt{3} \int d\tau d^2x \left\{ \frac{\tilde{K}}{2} \left((\nabla\phi_a)^2 + (\nabla\phi_s)^2 \right) + \frac{1}{2\tilde{\mu}_s} (\partial_\tau\phi_s)^2 + \frac{1}{2\tilde{\mu}_a} (\partial_\tau\phi_a)^2 + (-4i) \frac{(S-m)}{\nu} (\partial_\tau\phi_s) \right\}, \quad (2.93)$$

donde $\tilde{K} = \frac{3}{\nu}(Sa)^2(J_1 - J_x \cos \alpha) \sin^2 \theta^0$, $\tilde{\mu}_{s,a} = \frac{\nu}{16}\mu_{s,a}$, $\phi_s = (\phi_1 + \phi_3)/2$, $\phi_a = (-\phi_1 + \phi_3)/2$. Resaltamos que toda la dependencia en α está en el parámetro \tilde{K} , y que hemos hechos el rescaleo $x \rightarrow x/3$ para tener coeficientes de rigidez iguales en ambas direcciones espaciales.

A continuación descartamos el factor $\sqrt{3}$ fuera de la integral en la última ecuación. La notación ϕ_s y ϕ_a es elegida para enfatizar la presencia de una combinación simétrica y una

antisimétrica de los campos. A este orden en fluctuaciones la función de partición se factoriza según

$$Z = \left(\int \mathcal{D}\phi_s e^{-\mathcal{S}[\phi_s]} \right) \left(\int \mathcal{D}\phi_a e^{-\mathcal{S}[\phi_a]} \right) = Z_s Z_a, \quad (2.94)$$

y puede ser computada analíticamente. Aquí seguimos el enfoque de la referencia [26], y expresamos ϕ_b , $b \in \{s, a\}$ en términos del momento cristalino k y las frecuencias de Matsubara $\omega_n = \frac{2\pi n}{\beta}$, $n \in \mathbb{Z}$, según

$$\phi_b(\mathbf{r}, \tau) = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi\beta} \int d^2k e^{i\mathbf{k}\mathbf{r}} e^{-i\omega_n\tau} \phi_b(\mathbf{k}, \omega_n). \quad (2.95)$$

Esta representación es válida solo para campos que satisfacen $\phi_b(\mathbf{r}, 0) = \phi_b(\mathbf{r}, \beta)$, es decir, sin vorticidad, e implica $S_{BP} = 0$. La acción completa resulta

$$S = \frac{1}{\beta} \sum_b \sum_{n=-\infty}^{\infty} \frac{1}{2} \int d^2k \phi_b^*(\mathbf{k}, \omega_n) \phi_b(\mathbf{k}, \omega_n) \left(\tilde{K}k^2 + \frac{\omega_n^2}{\tilde{\mu}_b} \right). \quad (2.96)$$

Calculamos la integral funcional gaussiana en términos del campo adimensional reescalado $\phi'(k, \omega_n) = \phi(k, \omega_n)\Gamma/\beta$, donde $\Gamma \propto 1/a^2$, obteniendo

$$\log(Z_b) = \frac{-1}{2} \int \frac{d^2k}{\Gamma} \sum_{n=-\infty}^{\infty} \log \left[\frac{\beta}{\Gamma} \left(\tilde{K}k^2 + \frac{\omega_n^2}{\tilde{\mu}_b} \right) \right]. \quad (2.97)$$

Para sumar la serie, hacemos uso de la identidad

$$\int_1^{\beta\tilde{K}k^2/\Gamma} \frac{dt^2}{\frac{\omega_n^2}{\tilde{\mu}_b\Gamma/\beta} + t^2} = \log \left(\frac{\beta}{\Gamma} \tilde{K}k^2 + \frac{\omega_n^2}{\tilde{\mu}_b\Gamma/\beta} \right) - \log \left(1 + \frac{\omega_n^2}{\tilde{\mu}_b\Gamma/\beta} \right). \quad (2.98)$$

Así, la energía libre de Helmholtz es

$$F = -\frac{1}{\beta} \log(Z) = \frac{1}{\beta} \int \frac{d^2k}{\Gamma} \sum_b \log \left(\sinh \left(\frac{1}{2} k\beta \sqrt{\tilde{K}\tilde{\mu}_b} \right) \right), \quad (2.99)$$

donde hemos descartado las contribuciones de vacío, es decir, los términos independientes del momento k y del ángulo α . La última ecuación puede ser reescrita como

$$F = \int \frac{d^2k}{\Gamma} \sum_b \left\{ \frac{1}{2} k\beta \sqrt{\tilde{K}\tilde{\mu}_b} + \frac{1}{\beta} \log \left(1 - e^{-k\beta \sqrt{\tilde{K}\tilde{\mu}_b}} \right) \right\}, \quad (2.100)$$

donde el primer término es la contribución cuántica a la energía libre, F_Q , y el segundo es la contribución térmica, F_β . La primera puede ser integrada directamente, dando

$$F_Q = \frac{\pi}{3\Gamma} \sqrt{\tilde{K}} \Lambda^3 \sum_b \sqrt{\tilde{\mu}_b}, \quad (2.101)$$

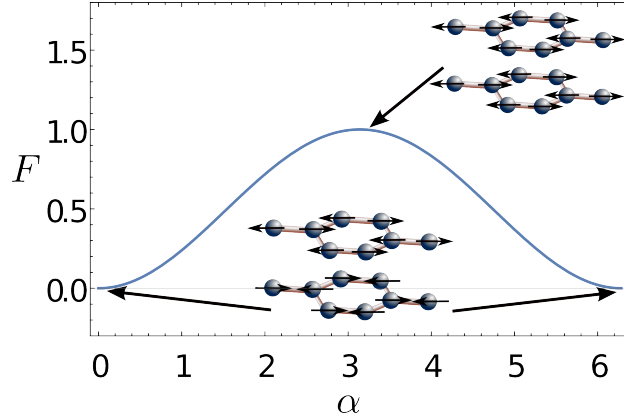


Figura 2.25: Energía libre correspondiente a la ecuación (2.104), con $J_1 = 1$, $J_x = 1/10$, $J_0 = 3J_x$, $\theta^0 = \pi/4$ y $\beta = 5$. La energía libre presenta mínimos en $\alpha = 0, 2\pi$, correspondientes a configuraciones antiparalelas en los acoplamientos verticales. El máximo en $\alpha = \pi$ corresponde a una configuración paralela en los acoplamientos verticales. Para la presentación, hemos normalizado la energía libre tal que $F(\alpha = 0) = 0$ y su valor máximo es 1.

donde $\Lambda = 2\pi/a$ es el valor de corte de momentos. Para F_β tenemos

$$F_\beta = \frac{2\pi}{\beta^3 \Gamma \tilde{K}} \sum_b \frac{1}{\tilde{\mu}_b} \int_0^{\Lambda \beta \sqrt{\tilde{K} \tilde{\mu}_b}} dx x \log(1 - e^{-x}), \quad (2.102)$$

donde hemos definido la variable adimensional $x = k\beta\sqrt{\tilde{K}\tilde{\mu}_b}$. Aquí tomamos el límite de baja temperatura $\beta\Lambda\sqrt{\tilde{K}\tilde{\mu}_b} \gg 1$ y obtenemos

$$F_\beta = (-\zeta(3)) \frac{2\pi}{\beta^3 \Gamma \tilde{K}} \sum_b \frac{1}{\tilde{\mu}_b}, \quad (2.103)$$

donde $-\zeta(3) \equiv \int_0^\infty dx x \log(1 - e^{-x}) \approx -1.2$, siendo $\zeta(s)$ la función zeta de Riemann, para $Re(s) > 1$. La energía libre Gaussiana completa en términos de los parámetros microscópicos desnudos de la teoría y del ángulo α es

$$F = 6\sqrt{2}\pi^4 \left(\sqrt{J_1 - J_x} + \sqrt{J_1 + J_x} \right) \sin(\theta^0) \sqrt{J_1 - J_x \cos(\alpha)} - \frac{32\pi J_1 \zeta(3) \csc^2(\theta^0)}{81\beta^3 (J_1^2 - J_x^2) (J_1 - J_x \cos(\alpha))}. \quad (2.104)$$

En la energía libre la contribución térmica y la cuántica tienen un único mínimo en $\alpha = 0$, por lo que la configuración de Néel es seleccionada tanto por las fluctuaciones térmicas como por las cuánticas. Adicionalmente, vemos en (2.104) que en esta aproximación, para $J_1 = J_x$ la contribución cuántica se anula y la térmica diverge, sugiriendo que cerca de $J_1 = J_x = J_0/3$ las contribuciones térmicas dominan frente a las cuánticas. Un análisis del sistema clásico a

temperatura finita sobre el punto¹⁰ $J_1 = J_x = J_0/3$ fue realizado en la referencia [32]. Para esta elección particular de acoplamientos magnéticos nuestra teoría de campos efectiva es inestable, dado que $m_3 = m_4 = \mu_3 = \mu_4 = 0$. Sin embargo, la presencia de un término de anisotropía $H_D = \sum_{\mathbf{r}, l, \eta} D(S_{l, \eta}^z)^2$, con $D > 0$, proveería una contribución finita a las cuatro masas μ_j , $j = 1, \dots, 4$, estabilizando las fluctuaciones en la magnetización. Este escenario está más allá del objetivo del presente estudio debido a que el punto $J_1 = J_x = J_0/3$ está lejos de las estimaciones experimentales y numéricas de los acoplamientos magnéticos para $\text{Bi}_3\text{Mn}_4\text{O}_{12}(\text{NO}_3)$, donde, como mencionamos, J_1 es el acoplamiento dominante. Además, nuestro estado fundamental clásico corresponde al límite de baja frustración mientras que para $J_1 = J_x = J_0/3$ la frustración es dominante y lleva a estados fundamentales clásicos más generales[32].

Hasta el momento los resultados de esta sección se aplican para campos ϕ_s, ϕ_a libres de vorticidad. Si en cambio proliferan los vórtices de ϕ_s el sistema entra en una fase gapeada de corto alcance. Para que esto esté permitido la teoría no debe tener fase de Berry, es decir, debe cumplirse la condición $4(S - m) \in \mathbb{Z}$. En este escenario el campo ϕ_s se desordena y como consecuencia también lo hace el campo ϕ_a , ya que ambas combinaciones no son completamente independientes.

2.3.4 Fase de dímeros

En esta sección consideramos el caso donde J_0 es el acoplamiento dominante. Si bien nos apartamos del material $\text{Bi}_3\text{Mn}_4\text{O}_{12}(\text{NO}_3)$, damos una descripción a la llamada fase de dímeros del sistema desde un enfoque de teorías de campos. Para $J_1 = J_x \ll J_0$, el estado fundamental a $T = 0$ puede determinarse exactamente y corresponde a un producto de singletes entre espines acoplados por J_0 [33]. Para verlo, introducimos los operadores

$$\mathbf{L}_\eta = \mathbf{S}_{1, \eta} + \mathbf{S}_{2, \eta} \quad \mathbf{K}_\eta = \mathbf{S}_{1, \eta} - \mathbf{S}_{2, \eta}, \quad (2.105)$$

con $\eta = 1, 2$, donde $[L_\eta^\alpha, L_{\eta'}^\beta] = i\epsilon^{\alpha\beta\gamma} L_\eta^\gamma \delta_{\eta, \eta'}$, $[L_\eta^\alpha, K_{\eta'}^\beta] = i\epsilon^{\alpha\beta\gamma} K_{\eta'}^\gamma \delta_{\eta, \eta'}$ y $[K_\eta^\alpha, K_{\eta'}^\beta] = i\epsilon^{\alpha\beta\gamma} L_\eta^\gamma \delta_{\eta, \eta'}$ en la misma celda unidad \mathbf{r} , siendo $\epsilon^{\alpha\beta\gamma}$ el tensor completamente antisimétrico de Levi-Civita y $\epsilon^{xyz} = 1$. Luego, reescribimos el Hamiltoniano (2.79),

$$\begin{aligned} H = & -2NJ_0S(S+1) + \frac{1}{2} \sum_{\mathbf{r}, i} \left\{ \frac{J_0}{3} \left(\mathbf{L}_1^2(\mathbf{r}_i) + \mathbf{L}_2^2(\mathbf{r}) \right) + \right. \\ & \left. (J_1 + J_x) \mathbf{L}_1(\mathbf{r}_i) \cdot \mathbf{L}_2(\mathbf{r}) + (J_1 - J_x) \mathbf{K}_1(\mathbf{r}_i) \cdot \mathbf{K}_2(\mathbf{r}) \right\} \\ & - h \sum_{\mathbf{r}} \left\{ \mathbf{L}_1^z(\mathbf{r}) + \mathbf{L}_2^z(\mathbf{r}) \right\}, \end{aligned} \quad (2.106)$$

con N el número de celdas unidad en el sistema y S el número cuántico de espín. Del Hamiltoniano anterior podemos ver que para $J_1 = J_x$ y $h = 0$, solo sobrevive la dependencia en $\mathbf{L}_\eta(\mathbf{r})$. Luego, puede verse[33] que el estado fundamental corresponde al producto de singletes $|\psi\rangle = \bigotimes_{\mathbf{r}} |s_1\rangle_{\mathbf{r}} |s_2\rangle_{\mathbf{r}}$ con $\mathbf{L}_\eta(\mathbf{r})|s_\eta\rangle_{\mathbf{r}} = 0$, y $|s_\eta\rangle_{\mathbf{r}} = \sum_{m=-S}^S (-1)^{S-m} |m, -m\rangle / \sqrt{2S+1}$. Aquí

¹⁰Hacia el final del trabajo, también estudian el régimen $J_1 = J_x < J_0/3$, que abordaremos en la siguiente sección.

$|m, -m\rangle$ etiqueta un producto de autoestados de $\mathbf{S}_{1,\eta}^z(\mathbf{r})$ y $\mathbf{S}_{2,\eta}^z(\mathbf{r})$ sobre el dímero η de la celda unidad ubicada en \mathbf{r} . Este resultado es válido para todo espín S , y para $J_0 \gg J_1 = J_x$ el *valence bond solid* descrito es el estado fundamental del sistema tanto en ausencia de campo como en presencia de un campo pequeño comparado contra el gap del sistema. Si aumentamos el valor de $J_1 = J_x$ habrá una transición de fase en cierto J_1^* donde la naturaleza del estado fundamental y el valor de J_1^* a priori dependen de S .

A continuación mostramos que a un nivel semiclásico la elección $J_1 = J_x$ da lugar a un modo cero de la teoría efectiva. Si en (2.91) J_0 es el acoplamiento dominante e incrementamos el valor de J_x , la teoría se vuelve singular en $J_1 = J_x$. En este caso tanto la masa m_4 como la rigidez K se anulan simultáneamente. En contraste, μ_j permanece positivo para $j = 1, \dots, 4$ (ver (2.89)). Como antes, la condición $J_1 = J_x$ corresponde a un modo cero del Hamiltoniano parametrizado por (2.80). En este caso podemos nuevamente hacer una variación en la parametrización agregando los parámetros reales $\alpha_{l,\eta}$, $l = 1, 2$, $\eta = 1, 2$ de la ecuación (2.92). Si fijamos $\alpha_{1,\eta} = \alpha_{2,\eta}$, $\eta = 1, 2$, la energía clásica a $T = 0$ es función solo de θ^0 . Nuevamente, a causa de la simetría $U(1)$ del Hamiltoniano podemos tomar $\alpha_{2,1} \equiv \alpha$ y $\alpha_{1,2} = 0$. Este modo cero está presentado gráficamente en la figura 2.24-b). El parámetro α no entra en la teoría efectiva esta vez, debido a que el coeficiente de rigidez se anula.

Para $J_1 = J_x$, en el límite de baja energía los dos campos masivos ϕ'_2 y ϕ'_3 en (2.91) pueden fijarse a cero, llevando a $\phi_1 = \phi_3$ y $\phi_2 = \phi_4$. Si consideramos el caso $m = 0$ entonces no hay fase de Berry y los vórtices en ϕ'_1 pueden proliferar si eso es energéticamente favorable, desordenando el sistema y abriendo un gap. En este escenario, la formación de singletes en el sistema se manifiesta en la teoría efectiva en el campo sin masa $\phi'_4 = (\phi_2 - \phi_1)/2$, que se deslocaliza, lo que lleva a que su campo canónicamente conjugado y masivo $\Pi'_4 = (\Pi_2 - \Pi_1)/2$ se localice a un valor nulo. Además, el anulamiento de la rigidez produce una banda chata en la relación de dispersión de las excitaciones magnéticas, lo que implica que están localizadas en el espacio de coordenadas, es decir, no se propagan a través del sistema, como ocurría en la cadena Sawtooth y en la cadena Kagomé.

Como comentario final podemos mencionar una estimación para el valor de $J_1 = J_x$ que marca el fin de la fase de dímeros. Es fácil ver de (2.89) y (2.88) que para $J_1 = J_x > J_0/3$ las masas m_3, μ_3 y μ_4 se vuelven simultáneamente negativas. Esto significa que nuestra teoría no interactuante es inestable con respecto a las fluctuaciones ϕ_3, Π_3 y Π_4 , y que el sistema debería estar en un estado fundamental descrito por otra teoría.

2.3.5 Conclusiones

Hemos estudiado al Hamiltoniano de Heisenberg en la geometría hexagonal bicapa con acoplamientos dentro y fuera del plano hexagonal (J_1 y J_0) en presencia de frustración (J_x), escribiendo una teoría de campos efectiva para describir el comportamiento magnético de $\text{Bi}_3\text{Mn}_4\text{O}_{12}(\text{NO}_3)$. Aunque exploramos diferentes regímenes en el espacio de parámetros de la teoría, prestamos especial atención al caso donde los acoplamientos magnéticos están cerca de aquellos estimados experimental y numéricamente[30] en $\text{Bi}_3\text{Mn}_4\text{O}_{12}(\text{NO}_3)$, donde el acoplamiento dominante es J_1 . En este caso, encontramos que el modelo clásico presenta un modo cero parametrizado por el ángulo relativo entre espines de cada capa. La teoría efectiva de baja energía de las fluctuaciones cuánticas puede escribirse en términos de un campo simétrico, relacionado con la magnetización global, y un campo antisimétrico, relacionado con el desbalance de espín

entre capas. Mostramos que la presencia de fluctuaciones cuánticas y térmicas selecciona la configuración colineal de la variedad degenerada de estados fundamentales clásicos. Este resultado vale para configuraciones sin vorticidad, y coincide con el orden magnético observado experimentalmente[29] cuando se aplica un alto campo magnético. Si la magnetización del sistema es tal que $4(S - m) \in \mathbb{Z}$ los vórtices pueden proliferar si la rigidez renormalizada es suficientemente pequeña, llevando al sistema a una fase gapeada de corto alcance. Este mecanismo podría explicar las correlaciones antiferromagnéticas de corto alcance observadas en $\text{Bi}_3\text{Mn}_4\text{O}_{12}(\text{NO}_3)$ a bajas temperaturas y bajo campo.

Por otra parte, mostramos que si J_0 es el acoplamiento dominante, el conocido[33] estado fundamental factorizado compuesto por un arreglo de singletes sobre los enlaces J_0 es interpretado en la teoría como una rigidez nula, como ya habíamos observado en una dimensión espacial para la cadena diente de sierra. Una diferencia con respecto a los cristales de magnones de la cadena Sawtooth y Kagomé, es que en la bicapa hexagonal el arreglo de singletes no rompe simetría de traslación de la red y no duplica el tamaño de la celda unidad. La teoría efectiva en este caso 2-dimensional no tiene un desbalance de espín. Como fue demostrado para el estado fundamental exacto factorizado, nuestro enfoque también es para S genérico, y no se restringe a $S = 3/2$ como es el caso del material.

3 Aprendizaje automático de fases magnéticas

Comenzamos a trabajar en la implementación de inteligencia artificial al estudio de sistemas magnéticos motivados por los resultados de Carrasquilla y Melko [34], donde los autores estudian un ferromagneto de Ising en la red cuadrada utilizando redes neuronales. A partir de una simulación Monte Carlo¹ del sistema puede construirse un conjunto de datos con el que pueden utilizarse las herramientas que originalmente fueron desarrolladas para el procesamiento de imágenes, como el reconocimiento facial, segmentación semántica, generación y transformación de imágenes, etc². Las herramientas de inteligencia artificial exceden al procesamiento de imágenes, pudiendo abarcar áreas como el procesamiento, comprensión y generación de textos, predicciones en finanzas y en clima, estudios de mercado, manejo automático de autos, etc[37].

En materia condensada, el objetivo usual de la implementación de inteligencia artificial es estudiar y caracterizar fases y transiciones de fase desde un enfoque de ciencia de datos (en inglés, *data driven approach*). Sin embargo, las aplicaciones pueden ser muy variadas, como la predicción de estructuras cristalinas[38, 39], la tomografía de estados cuánticos[40], la aceleración de simulaciones Monte Carlo[41], y más[42].

3.1 Introducción a las redes neuronales artificiales

3.1.1 Redes neuronales completamente conectadas

La red neuronal más simple consiste en una red neuronal denominada completamente conectada. La tarea habitual de este objeto consiste en aprender a hacer una clasificación, o regresión logística. El proceso de aprendizaje consiste en tener un conjunto X de N datos de entrada, x , etiquetados cada uno con su correspondiente clasificación, y , y lograr que la red neuronal aprenda a predecir correctamente la etiqueta que se le indica, a través del ajuste de sus parámetros. De esta manera, una vez entrenada, la red puede predecir a qué clase pertenece una dada muestra nueva x' .

Por el momento trabajamos con datos de entrada $x \in \mathbb{R}^{d_x}$, con d_x un número natural que corresponde a la dimensión de nuestro conjunto de datos. x_j denota entonces la componente j -ésima del vector x , con $j = 1, \dots, d_x$. La etiqueta y que señala a qué clase pertenece el dato x es un vector de dimensión d_y , donde d_y es la cantidad de clases posibles en la clasificación.

¹En el apéndice 5.2 desarrollamos los conceptos teóricos relacionados con la simulación computacional de sistemas clásicos via métodos Monte Carlo.

²También hay algoritmos basados en herramientas provenientes del área de inteligencia artificial que no necesitan conjuntos de datos, por ejemplo [35] y [36]

Si x pertenece a la clase j_0 , entonces su etiqueta satisface³ $y_j = \delta_{j,j_0}$.

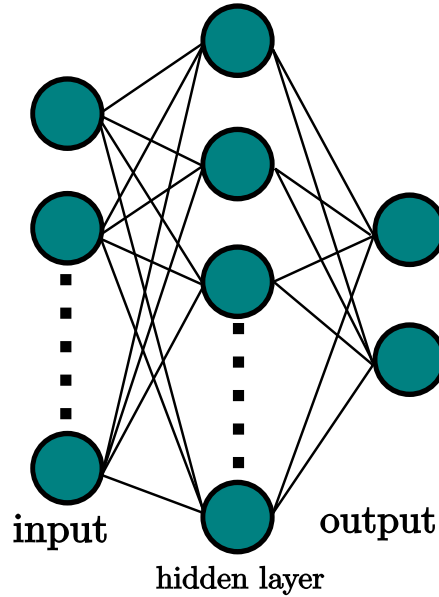


Figura 3.1: Esquema de una red neuronal completamente conectada. Cada círculo representa una neurona artificial, que tiene un término independiente o sesgo (*bias*) ajustable b_j^l y cada línea representa un peso (*weight*) w_{jk}^l . A izquierda, la capa de entrada o *input layer*. En el centro la capa oculta o *hidden layer*. A derecha la capa de salida o *output layer*.

Esquemáticamente, una red completamente conectada tiene el aspecto que muestra la figura 3.1. A la izquierda, el arreglo vertical de círculos representa la capa de entrada o *input layer*, indexada con el entero $l = 0$. Cada círculo representa a una neurona artificial, y contiene una variable o *feature* x_j , típicamente normalizada entre 0 y 1. En el centro se encuentra la capa oculta o *hidden layer*, indexada con $l = 1$, donde en la j -ésima neurona se construye una variable

$$a_j^{(l=1)} = \sigma \left(\sum_k w_{jk}^{(l=1)} x_k + b_j^{(l=1)} \right), \quad (3.1)$$

siendo x_k la k -ésima variable de entrada, b_j^l el término independiente o sesgo (*bias*) de la j -ésima neurona, w_{jk}^l el peso (*weight*) de la variable x_k en la j -ésima neurona y σ la función de activación. Esta función introduce las no-linearidades en el sistema, que de otra forma solo combinaría linealmente las variables de entrada. Las funciones de activación más populares son la ReLU (*Rectified Linear Unit*)

$$\text{ReLu}(x) = \text{máx}(0, x), \quad (3.2)$$

y la sigmoidea

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \quad (3.3)$$

³A esta forma de etiquetar en inglés se la denomina one-hot encoding.

En general para la l -ésima capa oculta, se siguen construyendo variables

$$a_j^l = \sigma \left(\sum_k w_{jk}^l a_k^{l-1} + b_j^l \right). \quad (3.4)$$

Finalmente, si en la tarea de clasificación hay d_y categorías, la capa final, indexada con $l = L$, tiene d_y neuronas. Cada neurona en esta capa final computa la probabilidad de pertenecer a cada categoría, típicamente usando la función de activación *softmax*

$$\sigma(a_j^L) = \frac{e^{a_j^L}}{\sum_{j'} e^{a_{j'}^L}}. \quad (3.5)$$

Los parámetros ajustables de la red, w_{jk}^l y b_j^l , se modifican con el objetivo de disminuir la función costo (a veces llamada función pérdida) de la red, vía el descenso por gradiente. A continuación definimos ambos conceptos.

Función costo o pérdida (cost or loss function):

Es conceptualmente una función que mide el error que se está cometiendo en la tarea de clasificación. La función costo comúnmente más utilizada por ventajas técnicas es la entropía cruzada, o *cross-entropy*, que para el caso de una clasificación binaria toma la forma

$$C = \frac{-1}{N} \sum_x \sum_j [y_j \ln(a_j^L) + (1 - y_j) \ln(1 - a_j^L)]. \quad (3.6)$$

Puede verse en (3.6) que $C \geq 0$, y que la igualdad vale solo si la red predice exactamente el etiquetado correcto. Para verlo, podemos considerar un dato de entrada x que pertenece a la clase $j_0 = 0$ (o bien, $j_0 = 1$), es decir, cuya etiqueta cumple $y_j = \delta_{j,j_0}$. Luego, si la red neuronal predice $a_j^L = \delta_{j,j_0}$ entonces el término correspondiente a x en (3.6) no contribuye. En el apéndice 5.3 derivamos a la entropía cruzada como función costo usando el principio de máxima similitud o *maximum likelihood principle*.

Descenso por gradiente (*gradient descend*):

Sabiendo que el gradiente de la función costo marca su dirección de máximo crecimiento, los parámetros ajustables se modifican en la dirección opuesta al gradiente. Para hacer a este cambio pequeño y controlado, se utiliza el parámetro llamado tasa de aprendizaje (*learning rate*), η , según

$$\begin{aligned} w_{jk} &\rightarrow w'_{jk} = w_{jk} - \eta \frac{\partial C}{\partial w_{jk}} \\ b_j &\rightarrow b'_j = b_j - \eta \frac{\partial C}{\partial b_j} \end{aligned} \quad (3.7)$$

Es importante remarcar aquí que η no es un parámetro entrenable, como los pesos y los sesgos. En cambio, η es definido por la persona que ejecuta al algoritmo y corresponde a lo que se denomina un hiperparámetro de la red neuronal, como también lo son el número de capas

en la arquitectura de la red, el número de neuronas en cada capa, etc. Una tasa de aprendizaje muy pequeña hace que el aprendizaje sea lento y pueda atorarse en mínimos locales con mayor facilidad, mientras que una muy grande puede impedir que el descenso por gradientes converja.

Aunque conceptualmente el proceso de minimización es muy simple, su implementación puede encontrar muchas dificultades, particularmente en relación con las altas dimensionalidades en los problemas de interés, la cantidad de datos y la cantidad de parámetros ajustables en la red.

En (3.6) puede verse que la función costo depende de la totalidad del conjunto de datos en los cuales se entrena. Tomar sucesivos gradientes sobre esta cantidad de datos resulta costoso, por lo que se introdujo la técnica conocida como descenso por gradiente estocástico (*stochastic gradient descend*), en la cual se toman m muestras aleatorias del conjunto de datos de entrenamiento, con las cuales se obtiene una aproximación del gradiente, agilizando así el proceso de ajuste. Este entero m se denomina ‘tamaño de lote’ (*batch size*). Al realizar el descenso por gradientes estocástico, tomando de a m muestras (*samples*) del conjunto de datos de entrenamiento, llega un momento en que se recorrió la totalidad de dicho conjunto. A esto se le denomina una época (*epoch*) de entrenamiento. El número de épocas a realizar durante el entrenamiento es un hiperparámetro más, que hay que fijar con algún criterio.

El famoso algoritmo utilizado para el cálculo de los gradientes en (3.7) para redes neuronales profundas se denomina algoritmo de propagación inversa (*back-propagation*), que describimos a continuación.

Algoritmo de propagación inversa

Es importante remarcar que el tipo de red neuronal presentada en la figura 3.1 tiene la siguiente propiedad, compartida con las redes neuronales profundas en general: las variables que se construyen en cada capa dependen de las variables de la capa previa, que está a su izquierda, con la obvia excepción de la capa de entrada. Luego, puede estudiarse como cambian las variables construidas en cada capa por la modificación de los parámetros de esa capa y de su capa anterior como muestra en detalle el capítulo 2 del libro de Nielsen[43], lo que permite definir las etapas del algoritmo de propagación inversa según:

1. Definir el input $x = a^1$
2. Alimentar hacia adelante (*feed forward*): para cada $l = 2, 3, \dots, L$ se calculan $z^l = w^l a^{l-1} + b^l$ y $a^l = \sigma(z^l)$
3. Propagación inversa: calcular la función costo y sus derivadas con respecto a los parámetros ajustables de la red capa a capa vía regla de la cadena, para realizar el descenso por gradiente estocástico.

En nuestros trabajos hemos usado los algoritmos de descenso por gradiente y propagación inversa definidos en las librerías de acceso libre de Tensorflow⁴. Propagación inversa es un caso especial de una herramienta más general denominada diferenciación automática (*automatic differentiation*). Para una pequeña demostración de cómo calcular derivadas utilizando auto-diferenciación en Tensorflow, escribimos un código disponible Github⁵.

⁴<https://www.tensorflow.org/>

⁵https://github.com/acevedo-s/AD_example

3.1.2 Sobreajuste y técnicas de regularización

Un problema usual de las redes neuronales es el de ‘sobreajuste’ (*overfitting*). Dada la cantidad de parámetros ajustables, puede ocurrir en la práctica que la red neuronal comience a aprender particularidades específicas del conjunto de datos de entrenamiento que no estarán presentes en nuevos conjuntos de datos. Por ejemplo, podría aprender un particular muestreo del ruido en los datos, que no es probable de volver a repetirse y que no es una característica física o relevante de nuestro sistema⁶. Para observar el sobreajuste en el entrenamiento de una red neuronal, se definen tres conjuntos de datos. En primer lugar, tenemos al conjunto de entrenamiento (*training set*). Este es el conjunto de datos con el que ajustamos los parámetros entrenables de la red con el algoritmo de propagación inversa. En segundo lugar, tenemos al conjunto de validación (*validation set*). Este conjunto de datos no interviene en el descenso por gradientes pero sí es parte del entrenamiento, ya que sobre él evaluamos la función costo o pérdida en cada época de entrenamiento. Si conforme avanza el entrenamiento obtenemos valores peores de nuestras métricas sobre el conjunto de validación relativos a los valores sobre el conjunto de entrenamiento, significa que nuestra red está sobreajustando al segundo conjunto. Finalmente, para afirmar que una red neuronal predice correctamente o no, hay que evaluarla sobre un conjunto de examen (*test set*) nunca antes visto por la red. A continuación presentaremos las dos técnicas más comunes para combatir el sobreajuste, conocidas como técnicas de regularización.

Regularización de pesos

La técnica de regularización más elemental se denomina regularización de pesos. Consiste en agregar un término a la función costo o pérdida que penalice la magnitud de los parámetros ajustables. A la pregunta del por qué no es deseable que la red tenga pesos ‘grandes’ pueden darse algunas respuestas conceptuales. En primer lugar, podemos pensar que cuanto más entrenamos a nuestro modelo (nuestra red neuronal), los pesos que reducen la función costo en el conjunto de entrenamiento aumentan más y más en una dada dirección específica de ese conjunto, lo que lleva al sobreajuste. En segundo lugar, pesos grandes hacen a nuestro modelo inestable, ya que una pequeña variación en los datos de entrada puede generar grandes variaciones en las variables de salida. Como tercer comentario podemos mencionar la posibilidad de anular pesos que no hayan aprendido características relevantes de nuestro conjunto de datos. De manera informal podemos pensar que los pesos que sobrevivan a la regularización tienen alguna una buena razón para tener un valor finito. Las dos regularizaciones de pesos estándar son las regularizaciones L_1 y L_2 . La primera, consiste en agregar a la función costo un término de la forma

$$L_1 = \frac{\lambda}{n} \sum_w |w|, \quad (3.8)$$

donde la suma es sobre los n pesos w en nuestro modelo, y $\lambda > 0$ es el parámetro de regularización, cuyo valor es arbitrario y corresponde a un hiperparámetro adicional del modelo.

⁶La interpretación más simple posible es que la red está aprendiendo ‘de memoria’ características que aparecen en el conjunto de datos de entrenamiento, que no tienen por qué repetirse en datos nuevos.

La regularización L_2 , en cambio, agrega un término de la forma

$$L_2 = \frac{\lambda}{2n} \sum_w w^2. \quad (3.9)$$

La principal diferencia entre ambas aparece en la derivada de la ecuación (3.7), donde podemos ver que la derivada es constante para la regularización L_1 mientras que es lineal con los pesos para la regularización L_2 . En la práctica, la comunidad utiliza con más frecuencia la regularización de tipo L_2 . Una discusión extensa sobre las diferencias entre ambas puede encontrarse en el capítulo 7 de la parte II del libro de Goodfellow[44].

Dropout

El *dropout* es una técnica de regularización introducida en la referencia [45], que consiste en anular aleatoriamente neuronas, junto con todas sus conexiones, durante el entrenamiento. La tasa de dropout (*dropout rate*) es la fracción de las neuronas que es anulada. Las fracciones típicamente utilizadas están entre 0.2 y 0.5. Durante la etapa de examen, en cambio, no se anula ninguna neurona([46] capítulo *overfitting and underfitting*). En cambio, si una neurona es retenida con probabilidad p durante el entrenamiento, el peso saliente de esa neurona es multiplicado por p en la etapa de examen para balancear el hecho de que hay más neuronas activas con respecto al entrenamiento. Los autores de este método de regularización interpretan a la implementación de dropout como la capacidad de promediar los resultados de entrenamientos con muchas redes neuronales diferentes, una por cada realización del muestreo que fija qué neuronas se conservan.

3.1.3 Redes neuronales convolucionales

Una red convolucional consiste en una red completamente conectada que es alimentada por características específicas que se extraen directamente de la imagen de entrada con un preprocesamiento vía filtros. Estos filtros aprenden sus parámetros en simultáneo con aquellos de la red completamente conectada en el entrenamiento, usando el algoritmo de propagación inversa. Las redes convolucionales toman como entrada directamente una imagen bidimensional (en general, un tensor arbitrario⁷), y están especialmente diseñadas para aprender patrones locales en imágenes, por lo que son ideales para estudiar sistemas con interacciones locales en arreglos geométricos periódicos.

Filtro: Consiste en una operación local que se efectúa sobre la totalidad de la imagen de entrada, generando así una nueva imagen procesada, o ‘mapa de características’ (*feature map*). Esta operación local consiste en tomar una combinación lineal sobre un conjunto pequeño de píxeles vecinos, típicamente de 3×3 , cuyos coeficientes se aprenden durante el entrenamiento, a la cual se le suma un término independiente o sesgo (*bias*) también entrenable. Se aplica entonces una función de activación y se genera un número, que corresponde a un píxel del mapa de características. Al barrer el filtro a lo largo de toda la imagen de entrada, se construye el

⁷Por ejemplo, para una imagen ‘a color’ tenemos un tensor de rango 3 donde dos índices corresponden a una imagen bidimensional mientras que el tercer índice tiene tres componentes o ‘canales’ denominadas R,G,B por las siglas en inglés de *red*, *green*, *blue*.

mapa de características. Los detalles de la aplicación de filtros en tensorflow pueden encontrarse en la documentación oficial⁸.

Una ventaja importante de la implementación de filtros, en comparación con una red completamente conectada, consiste en la cantidad de parámetros. Mientras que cada neurona en una red completamente conectada se conecta con todas las neuronas de la capa siguiente, un filtro es una única operación local que se efectúa sobre toda la imagen. Más aún, típicamente el tamaño de un mapa de características es menor al de la imagen original, debido al tamaño del filtro, salvo que uno explícitamente elija agregar ceros en los bordes (*padding*) para evitarlo.

Para disminuir aún más el número de parámetros ajustables de la red típicamente se aplican capas de agrupamiento (*pooling layers*). Estas capas de agrupamiento se aplican inmediatamente después de cada filtro. Hacer un agrupamiento (*pooling*), consiste en tomar una zona de la imagen generada tras aplicar un filtro, típicamente de 2×2 , y quedarse con alguna función escalar de esas 4 variables. El agrupamiento o *pooling* más simple y por ello el más popular es el *max-pooling*, que consiste en quedarse solo con el máximo de dicho subconjunto de datos. También puede optarse por tomar el promedio (*average-pooling*), la norma $L2$ (*L2-pooling*) del subconjunto de datos, etc. Finalmente, luego de aplicar filtros y hacer agrupamientos, se disponen a todas las variables en un vector (*flattening*) que se da como entrada a una red neuronal completamente conectada.

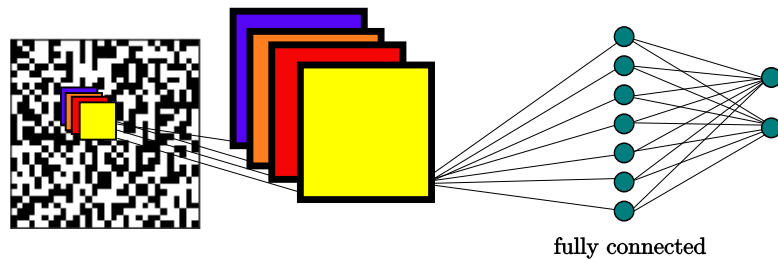


Figura 3.2: Caricatura de la aplicación de filtros y pooling layers en una red convolucional.

Gráficamente, una red convolucional elemental tiene el aspecto que muestra la figura 3.2. Allí la entrada es una imagen blanco y negro, sobre la cual se aplican varios filtros, generando un conjunto mapas de características. Por supuesto, pueden continuarse aplicando filtros a los mapas de características para generar mapas de características cada vez más abstractos. Luego se toman todas las variables resultantes y se las disponen en un vector 1-dimensional que se introduce en una red neuronal completamente conectada.

La imagen de entrada en la caricatura de red neuronal de la figura 3.2 es realmente una imagen de uno de nuestro conjunto de datos. Consiste en un arreglo bidimensional de 30×30 píxeles, donde cada píxel corresponde a una variable de Ising. Espines hacia arriba o abajo corresponden a píxeles blancos o negros respectivamente. Vemos que en la imagen la configuración de espines está desordenada, ya que es una configuración que corresponde a la fase paramagnética.

3.2 Modelos

A continuación utilizaremos el modelo de Ising antiferromagnético en las redes cuadrada, triangular y hexagonal (panal de abejas) para generar nuestros conjuntos de datos. Para el caso

⁸https://www.tensorflow.org/api_docs/python/tf/nn/conv2d

de las dos redes bipartitas introducimos frustración usando interacciones antiferromagnéticas a primeros y segundos vecinos, mientras que para el caso de la red triangular la frustración ya es inherente a la red. Cada espín es representado por una variable $\sigma = \pm 1$. El Hamiltoniano para las redes bipartitas elegidas se escribe entonces

$$H = J_1 \sum_{\langle i,j \rangle} \sigma_i \sigma_j + J_2 \sum_{\langle\langle i,j \rangle\rangle} \sigma_i \sigma_j, \quad (3.10)$$

con $J_1, J_2 > 0$, mientras que para la red triangular, $J_2 = 0$. Las notaciones $\sum_{\langle i,j \rangle}$ y $\sum_{\langle\langle i,j \rangle\rangle}$ corresponden a una suma sobre primeros y segundos vecinos, respectivamente. En adelante tomamos $J_1 = 1$, fijando la escala de energías y el sistema de unidades. Luego, $J_2 > 0$ regula la frustración del sistema. Además, la constante de Boltzmann es absorbida en la temperatura T .

La frustración en sistemas magnéticos corresponde a la imposibilidad de minimizar simultáneamente todas las interacciones, introduciendo degeneración en el sistema. Esto puede producir una variedad de comportamientos, incluso en modelos simples como el de Ising. Para el caso de la red cuadrada, para $J_2 > 1/2$ el estado fundamental está cuatro veces degenerado y presenta orden de largo alcance de tipo ‘tiras’. Además, la naturaleza de la transición de fase entre las fases ordenadas y desordenada depende del valor de J_2 y corresponde a un debate de muchos años[47, 48, 49, 50].

Para el antiferromagneto de Ising en la red triangular con interacciones a primeros vecinos, la frustración geométrica induce una degeneración macroscópica que destruye el orden de largo alcance incluso a $T = 0$. El sistema presenta entropía de punto cero y no tiene punto de Curie. Sin embargo, incluso siendo un paramagneto, a bajas temperaturas el sistema presenta correlaciones de corto alcance. En este sentido, este modelo puede verse como un líquido de espín clásico[51].

Para la red hexagonal, dual de la red triangular, hay dos subredes en las que el orden de Néel, o antiparalelo ‘encaja’ correctamente. Para $J_2 = 0$ el modelo ferromagnético y el antiferromagnético tienen una temperatura crítica en el límite termodinámico que puede ser determinada exactamente[52], y vale $T_c = 2/(\ln(2 + \sqrt{3})) \approx 1.519$. Al agregar interacciones a segundos vecinos, cuando J_2 alcanza el valor $1/4$ el modelo presenta una transición de fase de baja temperatura, desde la fase de Néel a una fase con alta degeneración y sin orden de largo alcance[53].

3.3 Aprendizaje supervisado en antiferromagnetos de Ising

Comenzamos estudiando el modelo de Ising antiferromagnético a primeros vecinos en la red hexagonal. Las configuraciones de espín a cada temperatura T que componen nuestro conjunto de datos fueron generadas con simulaciones Monte Carlo como se explica en la sección 3.3.4. Además, relegamos a esa sección los detalles de las arquitecturas utilizadas en los resultados que mostramos a continuación. La figura 3.3 muestra las probabilidades orden-desorden predichas por una red neuronal convolucional, como función de la temperatura. Para cada valor de T , estas probabilidades se obtienen promediando las predicciones sobre todas las configuraciones del conjunto de examen con esa temperatura. La zona sombreada representa el conjunto de temperaturas que fue excluido de la etapa de entrenamiento, en el siguiente sentido: Entrenamos

a la red neuronal solo con datos a temperatura T tal que $|T - T_c| > w$, es decir, excluyendo la región de transición. El valor de w es un hiperparámetro más, que en este caso fijamos en $w = 0.3$. La cantidad $2w$ es el ancho de la ventana sombreada. La línea a trazos vertical marca la temperatura crítica analítica del sistema en el límite termodinámico $T_c = \frac{2}{\log(2+\sqrt{3})} \approx 1.519J$. Este resultado es notable, ya que se entrena a la red neuronal con configuraciones ordenadas o desordenadas y ella es capaz de predecir la temperatura crítica de transición con precisión.

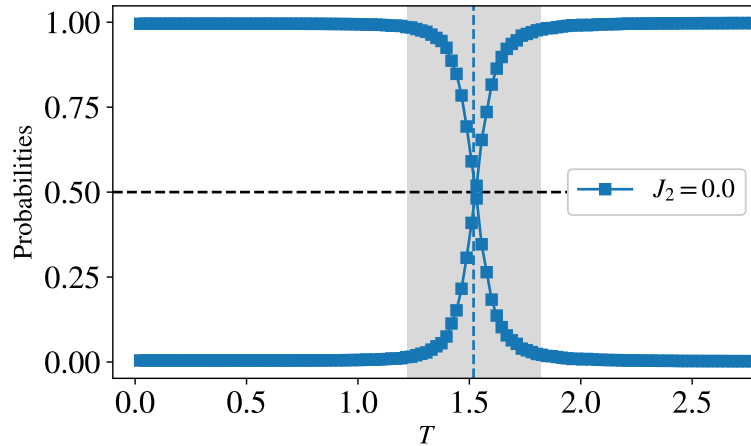


Figura 3.3: Probabilidades de orden-desorden (complementarias) como función de la temperatura para el antiferromagneto de Ising a primeros vecinos en la red hexagonal, calculadas por una red neuronal convolucional. La zona sombreada marca el rango de temperaturas removidas durante el entrenamiento. La línea vertical a trazos marca la temperatura crítica de transición del sistema en el límite termodinámico. La línea horizontal a trazos marca la probabilidad $p = 1/2$, como ayuda visual.

La figura 3.4 muestra tres configuraciones de espines utilizadas en la predicción de la figura 3.3. Por comodidad, mapeamos a las configuraciones de espines en la red hexagonal a un arreglo cuadrado, como detallamos en la sección 3.3.4. A izquierda, una configuración de baja temperatura, $T = 0.02$, con orden de Néel. Nuestra red neuronal entrenada predice que esta configuración tiene una probabilidad de estar ordenada mayor a 0.999. En el centro, una configuración con temperatura $T = 1.59$, que cae dentro de la ventana gris extraída durante el entrenamiento, para la cual la probabilidad de estar ordenada según la red es aproximadamente 0.75. Finalmente, a la derecha, una configuración a temperatura $T = 4.53$, donde la probabilidad de estar ordenada según la red es despreciable.

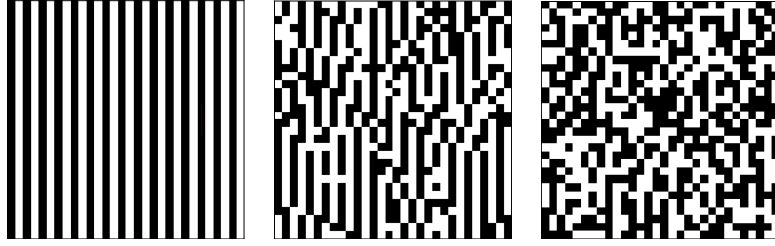


Figura 3.4: De izquierda a derecha, tres configuraciones de espines con temperaturas $T = 0.02$, $T = 1.59$ y $T = 4.53$, correspondientes al modelo de Ising antiferromagnético a primeros vecinos en la red hexagonal, mapeadas a un arreglo cuadrado (ver sección 3.3.4).

Por otra parte, el antiferromagneto de Ising en la red triangular en el límite termodinámico presenta una entropía de punto cero, está desordenado a toda temperatura y no tiene punto de Curie[54]. Sin embargo, en la simulación Monte Carlo del modelo el calor específico presenta un máximo a una dada temperatura que denotamos T^* , por debajo de la cual emergen correlaciones de corto alcance. La tarea de la red neuronal es entonces en este caso la de diferenciar regímenes de alta y baja temperatura, o sea, a ambos lados de T^* .

Para el caso de la red triangular, que es una red de Bravais, podemos mapear cada sitio de la red a una matriz cuadrada simplemente usando los vectores de la base. Cada sitio en la red está definido por una combinación lineal $\mathbf{r} = i_1\mathbf{e}_1 + i_2\mathbf{e}_2$, con $\mathbf{e}_{1,2}$ los vectores base de la red triangular. Luego, los índices $i_{1,2} = 1, \dots, 30$ pueden tomarse como los índices de una matriz cuadrada de $L \times L$ que corresponde a una configuración de espín en la red triangular.

El panel izquierdo de la figura 3.5 muestra las predicciones de una red completamente conectada sobre el modelo de Ising antiferromagnético en la red triangular. La zona sombreada representa como antes la región de temperaturas excluida del entrenamiento, y la línea a trazos vertical representa la temperatura T^* . En cruces azules, los resultados habiendo entrenado a la red neuronal con configuraciones ‘crudas’ de espines, donde la predicción de la temperatura de transición no es satisfactoria. En círculos naranjas, los resultados habiendo entrenado a la red neuronal con las correlaciones $C_{x,y} = \sigma_{x,y}\sigma_{L/2,L/2}$, donde x, y denota la posición de un espín en el sistema de L celdas. Usar las correlaciones, en este caso definidas con respecto al espín central, elimina la simetría Z_2 del sistema por lo que esperábamos a priori que la predicción mejorase. Si bien la predicción de la temperatura crítica mejora usando las correlaciones, el resultado tampoco es satisfactorio.

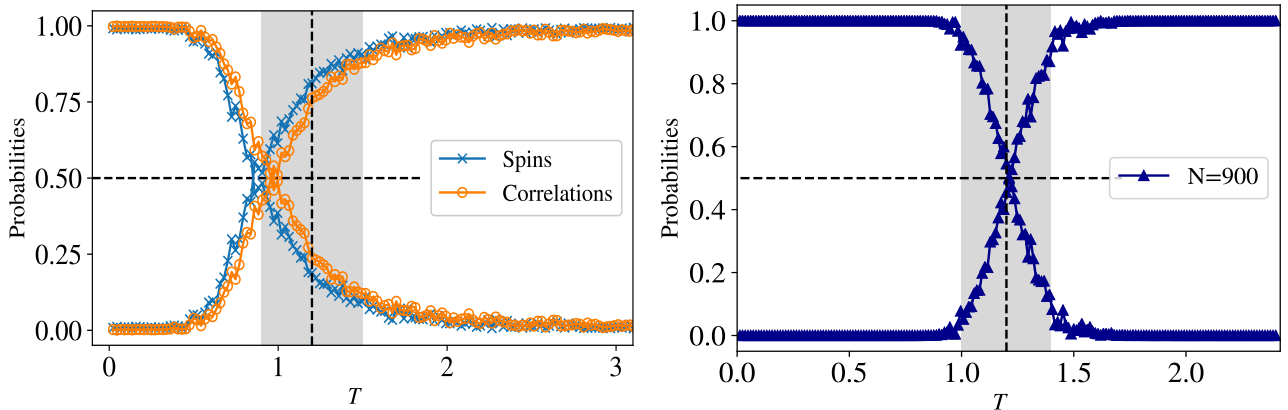


Figura 3.5: Panel izquierdo: En cruces azules (círculos naranjas), las probabilidades orden-desorden como funciones de la temperatura predichas por una red completamente conectada entrenada con configuraciones de espín (correlaciones). Panel derecho: probabilidades predichas por una red convolucional entrenada con configuraciones de espín. En ambos casos la línea a trazos vertical marca la temperatura T^* (ver texto), y la línea a trazos horizontal marca la probabilidad $p = 1/2$. La zona sombreada corresponde a la región de temperaturas excluida durante el entrenamiento.

El problema de la degeneración macroscópica debida a la frustración geométrica resulta superado al utilizar una red convolucional, como muestra el panel derecho de la figura 3.5.

Es intuitivo entender por qué para este caso falla la red completamente conectada y por qué la red convolucional puede diferenciar exitosamente las configuraciones de espín de altas y bajas temperaturas. En la red completamente conectada, la foto o configuración de entrada es vectorizada, es decir, colocada en un vector cuya longitud es la cantidad total de píxeles. En este paso píxeles que eran vecinos en la imagen inicial pueden quedar muy lejanos entre sí en la vectorización. Más aún, la presencia simultánea de diferentes patrones locales de correlaciones en diferentes regiones de la imagen de bajas temperaturas se hace difícil de reconocer en el vector. La red convolucional en cambio toma la imagen bidimensional y aplica filtros localmente a lo largo de toda la imagen, pudiendo reconocer diferentes patrones locales de correlaciones.

3.3.1 Transferencia de aprendizaje

El panel izquierdo de la figura 3.6 presenta las curvas de probabilidades obtenidas con una red convolucional entrenada en la red hexagonal con interacciones a primeros vecinos, con temperaturas fuera de la zona sombreada. Una vez entrenada, hacemos predecir a la red las probabilidades de orden-desorden en la red hexagonal con interacciones a primeros y segundos vecinos, es decir, la hacemos predecir sobre un sistema con la misma geometría pero con un Hamiltoniano o modelo modificado. Puede verse que el acuerdo con la estimación Monte Carlo para la temperatura de transición es excelente hasta $J_2 = 0.15$. Hacia $J_2 = 0.2$ el acuerdo empeora y la curva de probabilidad a baja temperatura ya no se aproxima a 1. Para $J_2 = 0.23$, alrededor de la temperatura de transición estimada vía Monte Carlo hay todavía un leve cambio en la probabilidad, es decir que la red distingue a esa temperatura algún cambio en el ordenamiento del sistema. Finalmente, para $J_2 = 0.25$ la probabilidad es constante en

todo el rango de temperaturas, la red no distingue ningún orden conocido. Este cambio en las probabilidades como función de J_2 sugiere una transición de fase, en la que se abandona el orden de Néel. Esta interpretación es respaldada con el cómputo del parámetro de orden del sistema en la fase de Néel, la magnetización alternada. Esta cantidad se construye sumando la magnetización de la subred A y restando la magnetización de la subred B y está presentada en el panel derecho de la figura 3.6. Allí puede verse que el parámetro de orden decrece continuamente como función de J_2 y se anula para $J_2 = 0.25$.

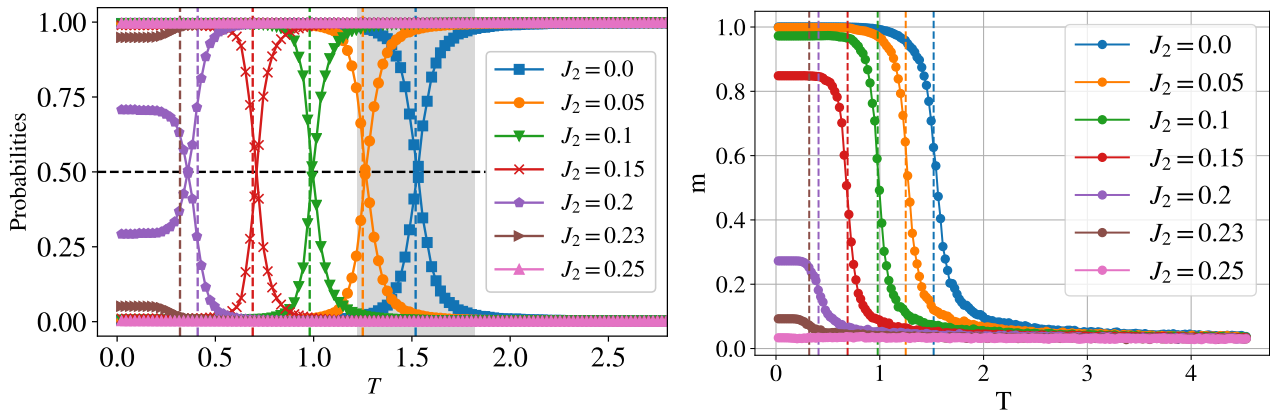


Figura 3.6: Panel izquierdo: Probabilidades de orden-desorden para conjuntos de datos con diferentes valores de J_2 , predichas por una red neuronal convolucional entrenada con configuraciones de espín con $J_2 = 0$ y temperaturas fuera de la región sombreada. Panel derecho: magnetización alternada como función de la temperatura para diferentes valores de J_2 . En ambos paneles las líneas a trazos verticales señalan las temperaturas a las cuales cada sistema presenta un máximo en el calor específico. El color de cada línea a trazos es igual al color de sus curvas de probabilidad correspondientes.

Continuando con la misma línea de estudio, analizamos el modelo de Ising con primeros y segundos vecinos en la red cuadrada, entrenando en dos casos límites. Por un lado, entrenamos una red convolucional con el modelo de Ising a primeros vecinos excluyendo como antes una ventana alrededor de la zona de transición. Luego, usamos esta red para predecir sobre datos generados con un modelo a primeros y segundos vecinos en la red cuadrada. La figura 3.7 muestra a izquierda los resultados de las probabilidades predichas por esta red para $J_2 \leq 1/2$, donde las líneas verticales indican las temperaturas de transición estimadas con el máximo del calor específico. Como puede observarse la red predice apropiadamente el cambio en la temperatura de transición como función de J_2 . Para $J_2 \approx 1/2$ el cambio en las probabilidades sugiere, como en el caso anterior, un cambio de fase donde el sistema deja de presentar orden de Néel. Si bien la transición de baja temperatura en la red cuadrada se encuentra en $J_2 = 0.5$ y las curvas de probabilidades no son chatas en ese caso, esto se explica a causa de que para $J_2 = 0.5$ las simulaciones Monte Carlo para la red cuadrada arrojaron un valor finito del parámetro de orden. Para cualquier $J_2 > 0.5$ el parámetro de orden de Néel es cero a toda temperatura y como consecuencia las curvas de probabilidades son completamente chatas, en $p = \pm 1$.

Por otra parte, entrenamos una red neuronal en $J_2 = 1$, excluyendo la zona de transición como antes. La transición entre la fase colineal y la paramagnética exhibe una naturaleza diferente,

primer o segundo orden dependiendo de J_2 , y ha sido estudiada en detalle[55]. Procediendo como antes, tomamos esta red y la usamos para predecir las temperaturas de transición sobre datos con $1/2 < J_2 < 1$. La figura 3.7 muestra a derecha los resultados donde el comportamiento es análogo al anterior pero del lado opuesto a la transición.

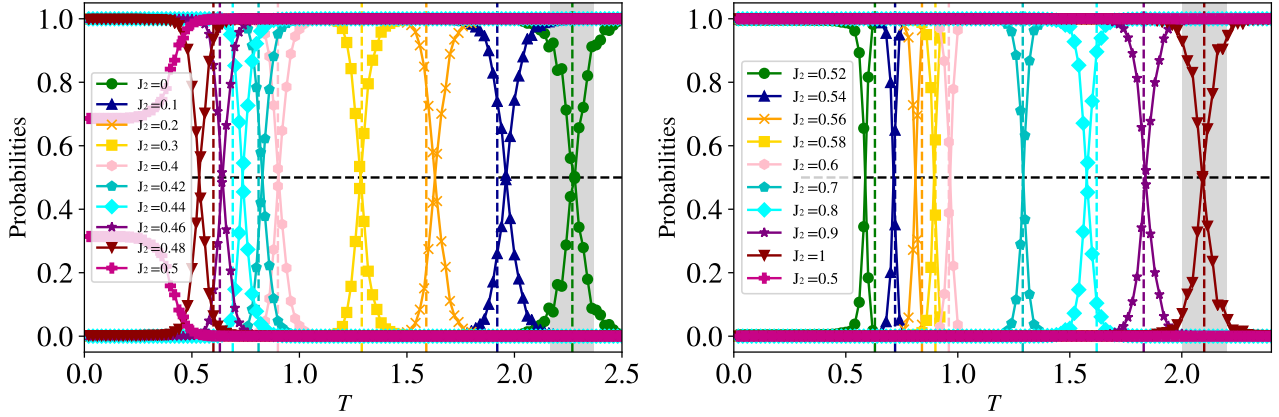


Figura 3.7: Panel izquierdo (derecho): Probabilidades de orden-desorden para conjuntos de datos con diferentes valores de J_2 , predichas por una red neuronal convolucional entrenada con configuraciones de espín con $J_2 = 0$ ($J_2 = 1$) y temperaturas fuera de la región sombreada. En ambos paneles las líneas a trazos verticales señalan las temperaturas a las cuales cada sistema presenta un máximo en el calor específico. El color de cada línea a trazos es igual al color de sus curvas de probabilidad correspondientes.

En la figura 3.8 construimos el diagrama de fases del antiferromagneto de Ising en la red cuadrada usando los cruces en las probabilidades de ambos paneles de la figura 3.7. En el panel incrustado en la figura 3.8 presentamos los resultados para el diagrama de fases de la referencia [47] en superposición con nuestros resultados como comparación. Destacamos que nuestro diagrama de fases está construido con solo dos entrenamientos, uno en $J_2 = 0$ y otro en $J_2 = 1$.

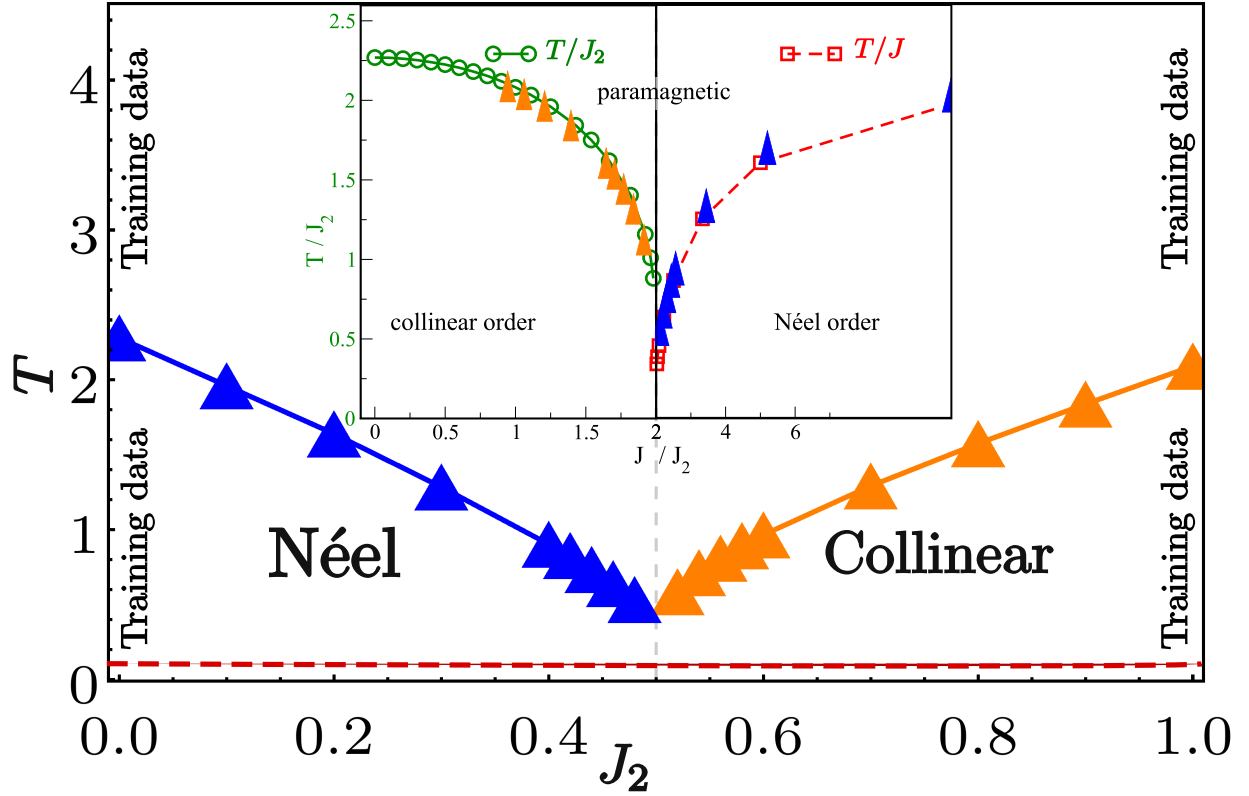


Figura 3.8: Panel principal: Diagrama de fases del antiferromagneto de Ising a primeros y segundos vecinos en la red cuadrada, predicho por dos redes neuronales convolucionales. Los triángulos azules (naranjas) indican los puntos de transición entre la fase de Néel (fase colineal) y la fase paramagnética, predichos por la red convolucional entrenada con datos con $J_2 = 0$ ($J_2 = 1$) del panel izquierdo (derecho) de la figura 3.7. La línea horizontal a trazos indica la temperatura usada en el método de confusión de la sección 3.3.2. Las líneas gruesas rojas verticales marcan las dos regiones de entrenamiento. Panel interno: Los círculos verdes y cuadrados rojos corresponden a las temperaturas de transición obtenidas en la figura 2 de la referencia [47]. Los triángulos naranjas y azules fueron agregados para comparar.

3.3.2 Aprendizaje por confusión

Este método puede considerarse un híbrido entre aprendizaje supervisado y no supervisado, pues utiliza herramientas del primero para predecir el etiquetado correcto, en vez de necesitarlo para el entrenamiento de la red[56].

Supongamos que nuestros datos dependen de un parámetro c que se encuentra en el intervalo $[a, b]$, y que el sistema tiene una transición de fase en un punto $a < c^* < b$ que no conocemos. Esto significa que los datos con $c < c^*$ pertenecen a una cierta fase A y los datos con $c > c^*$ pertenecen a otra cierta fase B. La idea central del método consiste en monitorear el rendimiento (*accuracy*) de la red realizando sucesivos entrenamientos, en donde se propone que la transición ocurre en un punto arbitrario c' . Si evaluamos la precisión, $P(c')$, esta función tiene una forma de W en el intervalo $[a, b]$, con un pico central justamente en el etiquetado correcto $c' = c^*$, como muestra esquemáticamente la figura 3.9.

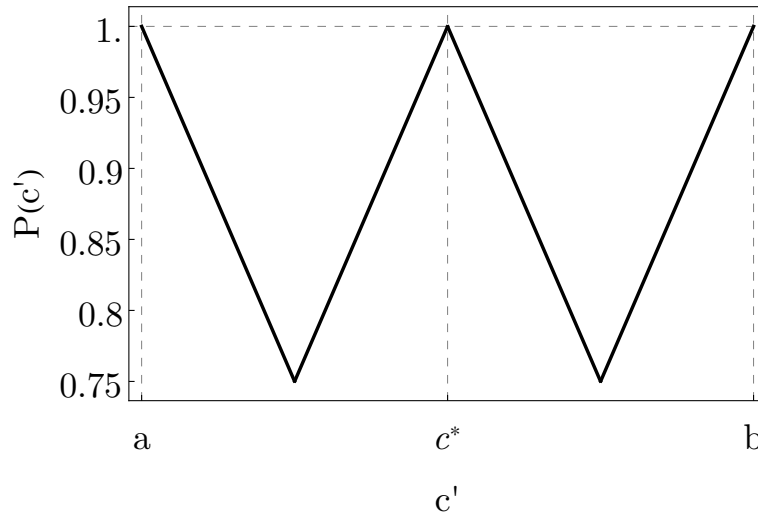


Figura 3.9: Esquema teórico del método de aprendizaje por confusión. Precisión de la red en función del etiquetado propuesto c' . Las líneas a trazos marcan las abscisas a , c^* y b , y la ordenada $P = 1$.

La forma de la curva puede entenderse de la siguiente manera: Si elegimos $c' = a$ ($c' = b$), estamos etiquetando a todos nuestros datos como pertenecientes a la fase B(A), por lo que la red aprende a decir que cualquier dato está en B(A), consiguiendo así un rendimiento trivialmente perfecto en la tarea de clasificación, $P(a) = 1$ ($P(b) = 1$). Esto da lugar al pico izquierdo (derecho) de la función $P(c')$ con forma de W en $[a, b]$. Cuando el parámetro propuesto coincide con el punto real de transición, $c' = c^*$, el rendimiento de la red neuronal en la tarea de la clasificación de fases es máximo, pues la estamos entrenando con el etiquetado correcto. De aquí viene el término confusión, ya que si etiquetamos mal a nuestros datos, la red se confunde durante el entrenamiento y la clasificación empeora. Es importante observar que un etiquetado incorrecto genera que algunas configuraciones con las mismas características sean presentadas a la red como pertenecientes a la fase A o como pertenecientes a la fase B, generando necesariamente un problema en el entrenamiento.

En nuestro trabajo usamos el aprendizaje por confusión para determinar el valor de J_2 en la red cuadrada para el cual a muy baja temperatura el sistema transiciona del orden de Néel al orden colinear. Si bien sabemos que la transición está en $J_2 = 0.5$, reobtenemos el resultado usando redes neuronales, habiéndolo supuesto desconocido. En la figura 3.10 graficamos la precisión (*accuracy*) como función de J_2 , donde cada punto corresponde a un entrenamiento independiente, con datos a temperatura fija $T = 0.02$. Puede verse una W distorsionada con centro alrededor de $J_2 = 0.5$, en concordancia con lo esperado.

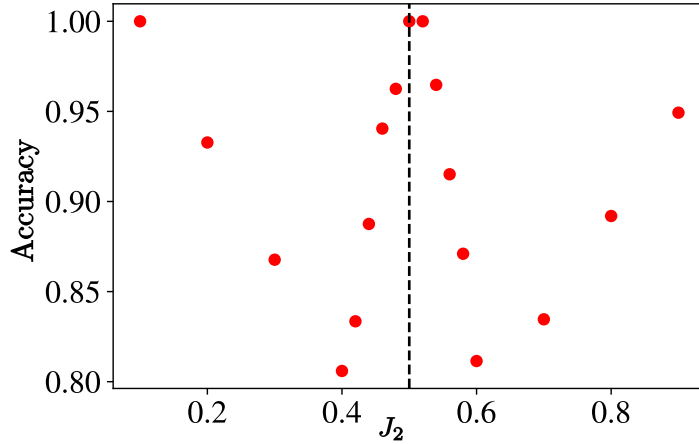


Figura 3.10: Precisión (*accuracy*) como función del J_2 propuesto como valor de transición para una red convolucional entrenada con el modelo de Ising a primeros y segundos vecinos en la red cuadrada.

3.3.3 Conclusiones

Utilizamos redes neuronales artificiales para que reconozcan las fases de los antiferromagnetos de Ising en las redes cuadrada, triangular, y hexagonal, con interacciones a primeros vecinos, entrenándolas con configuraciones de espines obtenidas de correspondientes simulaciones Monte Carlo.

Comenzamos entrenando de manera supervisada una red neuronal convolucional para que distinga entre configuraciones ordenadas y desordenadas en la red hexagonal, cuyas temperaturas están intencionalmente alejadas de la temperatura crítica, T_c , del sistema. Las probabilidades orden/desorden de la red evaluadas para toda temperatura $T \in [0.02, 4.53]$ se cruzan en T_c , señalando la transición de fase.

Luego observamos que en la red triangular, donde no hay una transición de fase, sino dos regímenes con fuertes correlaciones locales o sin ellas, una red densa no es capaz de hallar la temperatura de transición, T^* , con precisión. Sin embargo, una red convolucional, especialmente diseñada para detectar y procesar patrones locales, sí puede predecir el valor de T^* correctamente.

Más adelante nos enfocamos en el concepto de transferencia de aprendizaje, para reconocer la transición de fase de baja temperatura del modelo antiferromagnético frustrado $J_1 - J_2$ en las redes cuadrada y hexagonal. Las redes neuronales son entrenadas exclusivamente en los modelos a primeros vecinos, donde $J_2 = 0$, y se usan para predecir las probabilidades orden/desorden en datos con J_2 finito. Observamos que la probabilidad de orden predicha a bajas temperaturas cae a cero cuando J_2 excede el valor que corresponde a la transición de fase de bajas temperaturas. Realizando un proceso análogo, entrenando una red únicamente para diferenciar las dos fases del modelo de Ising con $J_2 = 1$ en la red cuadrada, podemos generar el diagrama de fases completo de este sistema, que comparamos con resultados conocidos.

Nuestra técnica de clasificación se enfoca en discriminar dos tipos de datos que corresponden a una fase de alta temperatura y otra de baja temperatura, permitiendo predecir la temperatura

de transición entre ambas, pero no proporciona información sobre cuál es el tipo de transición de fase dada. La naturaleza de la transición de fase es una temática muy interesante para abordar, aunque escapa el objetivo del presente estudio. Hay, sin embargo, enfoques que utilizan inteligencia artificial para poder diferenciar transiciones de primer orden, segundo orden, y de tipo Berezinskii-Kosterlitz-Thouless, como el estudio de la llamada dimensión intrínseca[57] de las configuraciones.

Finalmente, usamos la técnica denominada aprendizaje por confusión para mostrar cómo es posible encontrar el etiquetado correcto con sucesivos entrenamientos en los que se propone un etiquetado arbitrario.

3.3.4 Tecnicismos

Mapeo de la red hexagonal a un arreglo cuadrado

Indexamos cada celda unidad en la red hexagonal con dos índices enteros (i, j) , donde $0 \leq i < N_1 = 30$ y $0 \leq j < N_2 = 15$ como muestra la figura 3.11.

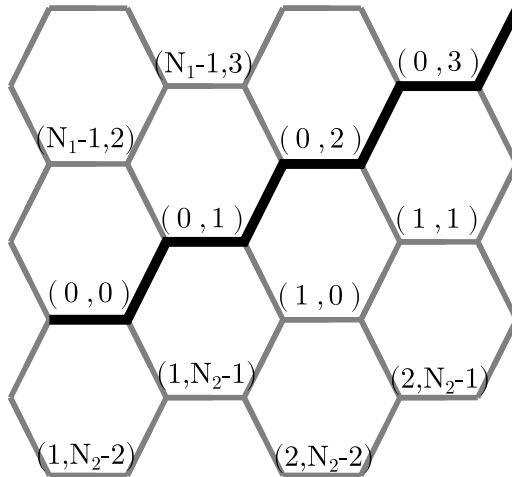


Figura 3.11: Red hexagonal. Cada par ordenado corresponde a la indexación elegida para cada celda unidad. La línea negra gruesa señala el conjunto de sitios que conforma la primer fila de la matriz cuadrada A en el mapeo (3.11).

Mapeamos la red hexagonal a un arreglo cuadrado de 30×30 de la siguiente manera: cada espín en la red hexagonal es indexado por tres índices enteros en un tensor $S_{i,j}^k$, donde i y j determinan la celda unidad (i, j) , y $k = 0$ o $k = 1$ corresponde al espín izquierdo o derecho en la celda unidad respectivamente. Luego, podemos construir un arreglo cuadrado A de espines según

$$A_{mn} = S_{m \lfloor n/2 \rfloor}^{\text{mod}(n,2)} \quad (3.11)$$

con $0 \leq m, n < 30$, $\text{mod}(n, 2)$ el resto en la división de n por 2, y $\lfloor n/2 \rfloor$ la función piso, que da la parte entera del resultado de la división. Como ejemplo, podemos tomar el conjunto de espines señalado por la línea negra gruesa en la figura 3.11, es decir, fijamos $i = 0$ en el tensor S_{ij}^k . Así, la primer fila de A_{mn} ($m = 0$) está dada por

$$A_{00} = S_{00}^0, A_{01} = S_{00}^1, A_{02} = S_{01}^0, A_{03} = S_{01}^1, A_{04} = S_{02}^0, A_{05} = S_{02}^1, \dots \quad (3.12)$$

Simulación Monte Carlo

Nuestros conjuntos de datos fueron generados con una simulación Monte Carlo, usando el algoritmo de metrópolis y una dinámica de *single spin-flip*. Para cada valor de J_2 realizamos 400 simulaciones independientes, comenzando a alta temperatura ($T_0 = 4.53$). Consideramos un conjunto de 200 temperaturas equiespaciadas en el rango $[0.02, T_0]$. Para cada temperatura se exporta una configuración de espín una vez alcanzado el equilibrio. Así nuestro conjunto de datos para cada valor de J_2 consiste en 80000 configuraciones o imágenes. Todas las simulaciones se realizaron con 900 sitios, con condiciones de contorno periódicas.

Las simulaciones para las redes triangular y cuadrada fueron realizadas por Inés Corte, junto con las figuras 3.10, 3.7 y el panel derecho de la figura 3.5, en el marco del trabajo [58], mientras que las simulaciones para la red hexagonal fueron realizadas por mí, escritas en C .

Entrenamiento

Los datos generados en la simulación Monte Carlo son separados 70 % como conjunto de entrenamiento (10 % del cual es tomado como conjunto de validación) y 30 % como conjunto de examen o predicción. Datos con $T < T_c$ son etiquetados con 0 y datos con $T > T_c$ son etiquetados con 1. Para la red triangular, cambiamos T_c por T^* . Los valores específicos de la temperatura son utilizados solamente en la etapa de examen, donde analizamos las predicciones de la red como función de la temperatura para encontrar la temperatura de transición. El ancho de ventana w para excluir temperaturas alrededor de la temperatura de transición fue variable entre 0.1 y 0.3. No encontramos diferencias en las predicciones moviendo a w entre estos valores.

Redes Completamente conectadas

En el panel izquierdo de la figura 3.5 usamos como entrada un vector con las configuraciones de espín que exportamos de la simulación Monte Carlo, o bien las correlaciones correspondientes con respecto al espín central. Tanto configuraciones como correlaciones están normalizadas donde los dos posibles valores se fijan a 0 y 1. La red neuronal tiene una capa oculta de 32 neuronas con funciones de activación ReLU. La capa final tiene dos neuronas con funciones de activación Softmax.

El método de optimización es *Adam*⁹, elegido por su amplio uso en la bibliografía, y la función costo elegida es la entropía cruzada categórica¹⁰. Para el entrenamiento usamos entre 50000 y 70000 configuraciones, 10 % de las cuales tomamos como conjunto de validación. La tasa de aprendizaje es de orden 10^{-4} , el tamaño de lote es de 128, el regularizador es de tipo *L2*, de orden 10^{-5} y el número de épocas es de 20. La precisión en validación final es mayor a 0.98 y el error en validación final es de orden 10^{-1} .

⁹https://www.tensorflow.org/api_docs/python/tf/keras/optimizers/Adam

¹⁰https://www.tensorflow.org/api_docs/python/tf/keras/losses/CategoricalCrossentropy. El término *categorical* está asociado con el uso del formato ‘one-hot encoding’ para etiquetas.

Redes Convolucionales

La entrada es una matriz de 30×30 con configuraciones de espín normalizadas a los valores binarios 0 y 1. La segmentación de los datos en entrenamiento validación y evaluación es tomada al igual que para la red completamente conectada. Muchas arquitecturas pueden ser adecuadas para la tarea de clasificación. Las redes convolucionales usadas en este trabajo consisten primero en una o dos capas convolucionales de 3 a 10 filtros de tamaño 3×3 , cada una seguida por una capa de *max* o *average-pooling*. Luego los tensores son vectorizados y conectados a una capa densa de 3 a 16 neuronas con funciones de activación ReLU. El método de optimización es *Adam* y la función costo o pérdida es la entropía cruzada categórica. Los tamaños de lote tomados están entre 128 y 512. El número de épocas está entre 3 y 5 y la tasa de aprendizaje es de orden $10^{-3} - 10^{-4}$. La capa final tiene nuevamente dos neuronas con funciones de activación Softmax. La precisión en validación es mayor a 0.99 y la pérdida correspondiente menor a 10^{-2} .

3.4 Aprendizaje no supervisado: AutoEncoders

Los autoencoders[44] son redes neuronales que se entrenan para generar una copia aproximada de un dado dato de entrada, x . Están compuestos por dos partes: el encoder, que codifica el dato de entrada, $h = f(x)$, y el decoder, que a partir del código producido por el encoder genera una réplica aproximada del dato original, $r = g(h)$.

Si la reconstrucción del dato de entrada fuese siempre perfecta el autoencoder tendría poca utilidad. Sin embargo, si el autoencoder realiza una copia aproximada del dato de entrada, entonces está forzado a aprender las características más relevantes del mismo. Por esta razón, el código producido por el encoder, $h = f(x)$, es forzado a pertenecer a un denominado espacio latente, generalmente de dimensión mucho menor a la dimensión del dato original, x .

Al igual que en el caso de las redes neuronales completamente conectadas o convolucionales descriptas anteriormente, los autoencoders aprenden o ajustan sus parámetros minimizando una función costo, o pérdida. Esta función a minimizar debe definir una distancia entre el dato de entrada x y el dato reconstruido por el autoencoder $r = g(f(x))$. La función pérdida estándar para autoencoders es el error cuadrático medio (*MSE* por sus siglas en inglés).

La diferencia más importante con las redes neuronales antes mencionadas reside en que en la tarea de reconstrucción del autoencoder no necesita de una etiqueta con información del sistema (a qué clase pertenece, etc) pues el autoencoder aprende únicamente a crear una reproducción aproximada del dato de entrada, independiente de lo que represente.

Entre las aplicaciones de los autoencoders se encuentran¹¹ la reducción de dimensionalidad y el aprendizaje de características relevantes en conjuntos y eliminar ruido de imágenes (*denoising*), por ejemplo, para optimizar tareas de clasificación (*feature learning*). En particular, los denominados autoencoders variacionales (VAE's) pueden utilizarse como modelos generativos, por ejemplo para la generación de rostros humanos sintéticos, debido a la regularidad de su espacio latente. Por último, y de especial relevancia para nuestro trabajo, los autoencoders pueden también utilizarse para la detección de anomalías, como operaciones fraudulentas, ecocardiogramas de corazones enfermos o transiciones de fase.

¹¹Sobre reducción de dimensionalidad y feature learning:[44] capítulo 14 sección 9. Sobre denoising y anomaly detection: <https://www.tensorflow.org/tutorials/generative/autoencoder>

En un autoencoder convolucional, el encoder está compuesto por capas convolucionales, donde se realiza un proceso de *down-sampling*, es decir, que la dimensión de la imagen de entrada de la capa es mayor que la dimensión de la imagen de salida. Esta reducción de dimensionalidad puede hacerse con *pooling layers*, o bien usando un paso o *stride* mayor que uno como hiperparámetro de la capa convolucional. El paso de la capa convolucional es el número de píxeles que se mueve cada filtro al barrer la imagen de entrada cada vez que va a generar un píxel en la imagen de salida. Una vez codificada la imagen de entrada, el decoder debe realizar un proceso de *up-sampling* para construir una imagen con la dimensión de la imagen original. Este proceso se realiza via *UpSampling layers* o *Transposed convolutional layers*.

UpSampling layers

Aumentan la dimensión de la imagen de entrada realizando copias de sus filas y columnas, y están seguidas de una capa convolucional.

Capas convolucionales traspuestas (Transposed convolutional layers)

Al introducir este tipo de capas, primero es usual mencionar que capa convolucional traspuesta y capa deconvolucional no son sinónimos. Una deconvolución se entiende como la transformación inversa a una convolución. En cambio, una capa convolucional traspuesta toma la imagen de entrada, agrega zeros entre sus filas y columnas y luego aplica una convolución estándar. De esta manera la dimensión de la imagen de salida es mayor a la de entrada, pero numéricamente el proceso no tiene semejanza alguna con la operación inversa a una convolución. Para más información, puede consultarse la página oficial de tensorflow¹².

3.4.1 Entrenamientos a baja temperatura

Entrenamiento 1: Datos de entrenamiento no frustrados a baja temperatura

Entrenamos un autoencoder convolucional para reconstruir configuraciones de espín que presentan orden de Néel, correspondientes a $J_2 = 0$ y $0.02 \leq T \leq 0.2$. Denotamos a este entrenamiento como el entrenamiento I, y al autoencoder entrenado resultante, el autoencoder I.

La figura (3.12) presenta en sus paneles superiores configuraciones de espines correspondientes a $T = 0.02$ a izquierda y $T = 4.53$ a derecha, ambas con interacciones únicamente a primeros vecinos ($J_2 = 0$). El orden de Néel en la red hexagonal es mapeado a un orden de tiras en el arreglo cuadrado mostrado, como explicamos en la sección 3.3.4. Los paneles inferiores de la figura (3.12) muestran las reconstrucciones correspondientes. El autoencoder reproduce correctamente la configuración ordenada e intenta construir una configuración ordenada incluso a partir de la configuración desordenada.

¹²https://www.tensorflow.org/api_docs/python/tf/keras/layers/Conv2DTranspose

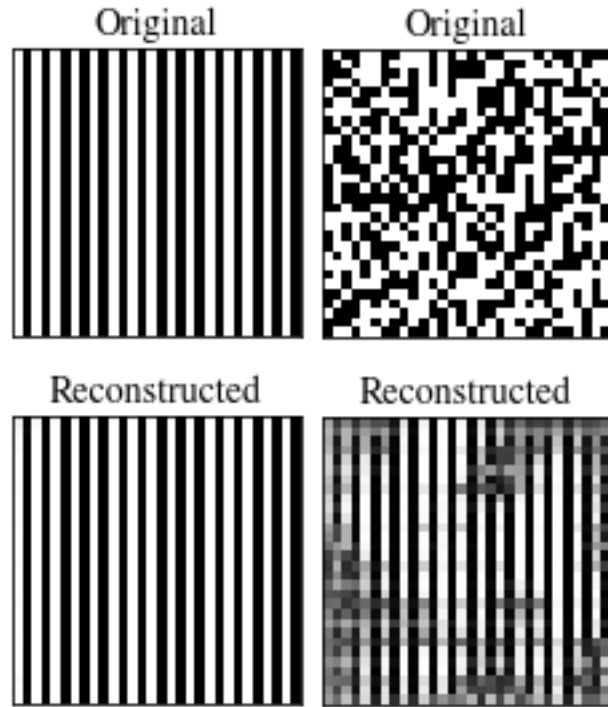


Figura 3.12: Fila superior: configuraciones de espines originales sin frustración ($J_2 = 0$). A izquierda, una configuración ordenada tomada a $T = 0.02$. A derecha, una tomada a $T = 4.53$. En la fila inferior, las reconstrucciones correspondientes realizadas por el autoencoder I.

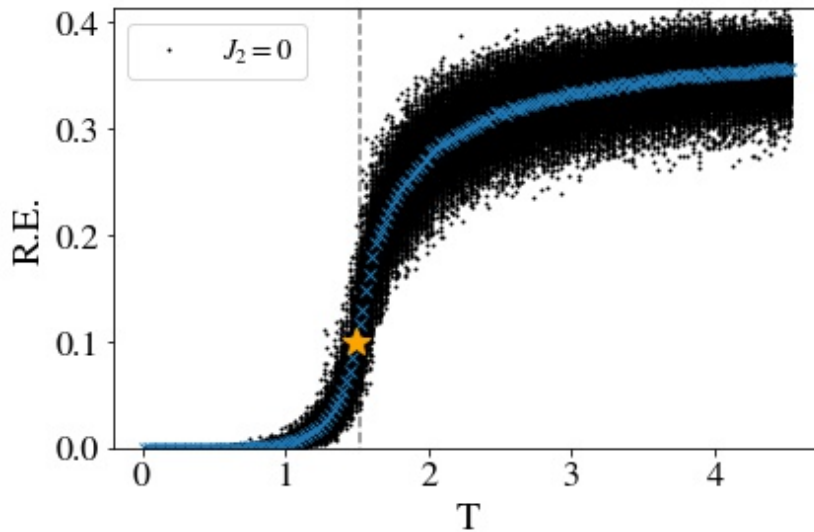


Figura 3.13: En puntos negros el error de reconstrucción versus temperatura para 400 realizaciones del sistema no frustrado ($J_2 = 0$), calculadas por el autoencoder I. La línea punteada gris marca la temperatura crítica analítica del sistema en el límite termodinámico. La estrella amarilla marca la estimación numérica del punto de inflexión de la curva de error de reconstrucción promedio, marcado con cruces azules.

La figura 3.13 muestra el error de reconstrucción como función de la temperatura para 400 realizaciones del sistema. A baja temperatura, cuando el sistema converge a uno de los dos estados fundamentales con orden de Néel, el autoencoder es capaz de reproducir estas configuraciones y el error de reconstrucción llega a un mínimo (orden 10^{-4}). A altas temperaturas el autoencoder no puede reproducir las configuraciones desordenadas, como mostramos en la figura 3.12, y el error de reconstrucción aumenta. El error de reconstrucción máximo posible usando nuestras elecciones de métrica y normalización es de 1. Como puede observarse en la figura, hay un buen acuerdo entre el punto de inflexión en la curva de error de reconstrucción medio y la temperatura crítica del sistema en el límite termodinámico. La estimación numérica del punto de inflexión se desarrolla en la sección 3.4.5

La figura 3.14 muestra el error de reconstrucción medio como función de la temperatura, para diferentes valores de J_2 , computados con el autoencoder I. A alta temperatura, todas las curvas tienden a converger a un valor constante, ya que el sistema está en su fase paramagnética. A baja temperatura el error de reconstrucción decrece con respecto al valor de alta temperatura para sistemas con $J_2 < 1/4$, que están en la fase de Néel. Este no es el caso para sistemas con $J_2 > 1/4$, en acuerdo con la transición de fase de baja temperatura en $J_2 = 1/4$.

Aunque en la figura 3.14 es posible distinguir tres regímenes que corresponden a las 3 fases del sistema, para las curvas con $J_2 > 1/4$ el error de reconstrucción no permite discriminar apropiadamente altas y bajas temperaturas. Por ejemplo, para la curva gris en la figura 3.14 los errores de reconstrucción de bajas y de altas temperaturas coinciden dentro de la dispersión. Esto puede entenderse como una consecuencia de la simpleza del entrenamiento. El autoencoder tuvo la tarea de reconstruir configuraciones ordenadas, y con ese entrenamiento está siendo usado para intentar distinguir configuraciones altamente frustradas de configuraciones desordenadas, donde ninguna de las dos presenta orden de Néel de largo alcance. Esta dificultad será superada en la siguiente sección.

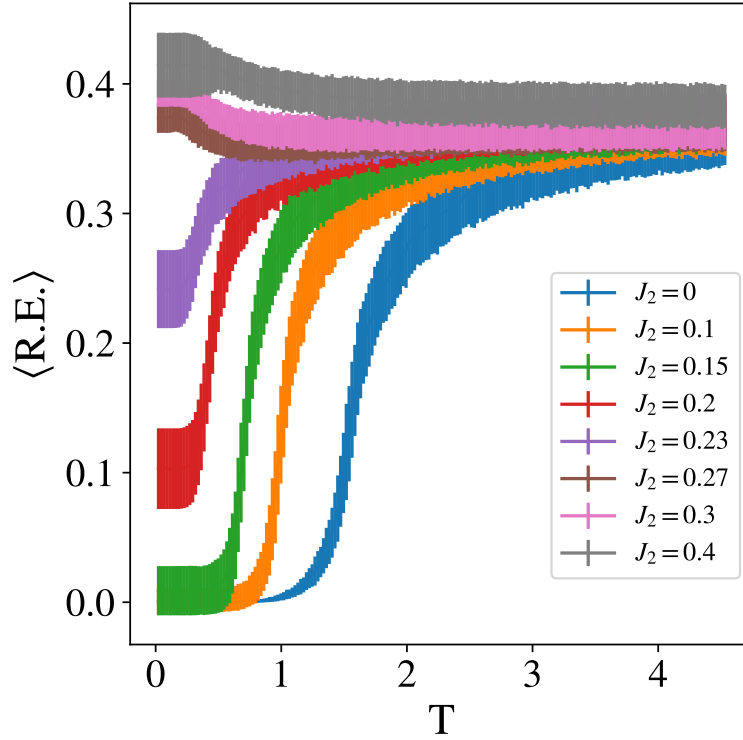


Figura 3.14: Error de reconstrucción medio versus temperatura para difentes valores de J_2 , calculados con el autoencoder I. Las barras de error muestran la desviación estándar calculada usando las 400 realizaciones del sistema.

En la figura 3.14 puede observarse que la dispersión es también función de la temperatura. La figura 3.15 muestra la desviación estándar σ en el error de reconstrucción como función de la temperatura, para diferentes valores de J_2 . Para $J_2 < 1/4$ se observan picos en la dispersión que coinciden aproximadamente con los puntos de inflexión respectivos, marcados con las líneas verticales a trazos. Además se observa que todas las dispersiones convergen aproximadamente al mismo valor a altas temperaturas. A bajas temperaturas, en cambio, la dispersión aumenta conforme aumenta la frustración.

Para $J_2 > 1/4$ puede verse que la dispersión aumenta con respecto al valor de alta temperatura alrededor de la temperatura de transición, pero no se observan picos como en la figura 3.15. Una figura que muestra este comportamiento será presentada en la sección 3.4.3 para una elección diferente de entrenamiento.

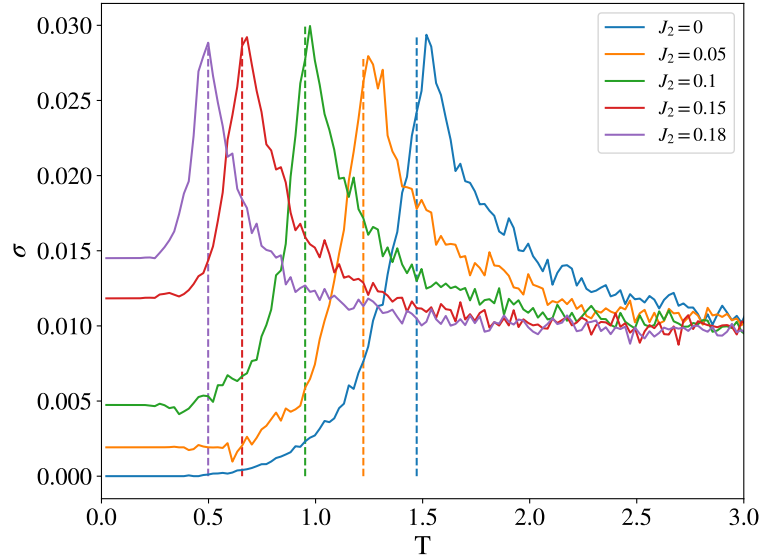


Figura 3.15: Desviación estándar σ en el error de reconstrucción para diferentes valores de J_2 en la fase de Néel, calculados con el autoencoder I. En líneas a trazos las estimaciones de cada punto de inflexión en cada curva de error de reconstrucción medio correspondiente.

Entrenamiento II: Datos de entrenamiento frustrados a baja temperatura

Entrenamos un autoencoder con configuraciones con $0.02 \leq T \leq 0.2$ y $J_2 = 1/2$, donde el sistema no presenta orden de largo alcance y tiene un estado fundamental altamente degenerado debido a la frustración. Denotamos este como el entrenamiento II, y al autoencoder resultante como autoencoder II.

La figura 3.16 muestra configuraciones tomadas a $T = 0.02$ para tres realizaciones diferentes del sistema con $J_2 = 1/2$ en la fila superior, y sus respectivas reconstrucciones realizadas por el autoencoder II en la fila inferior. Puede observarse que hay diferentes zonas en cada imagen con diferentes órdenes locales. El error de reconstrucción final en el conjunto de datos de validación es de orden 10^{-3} .

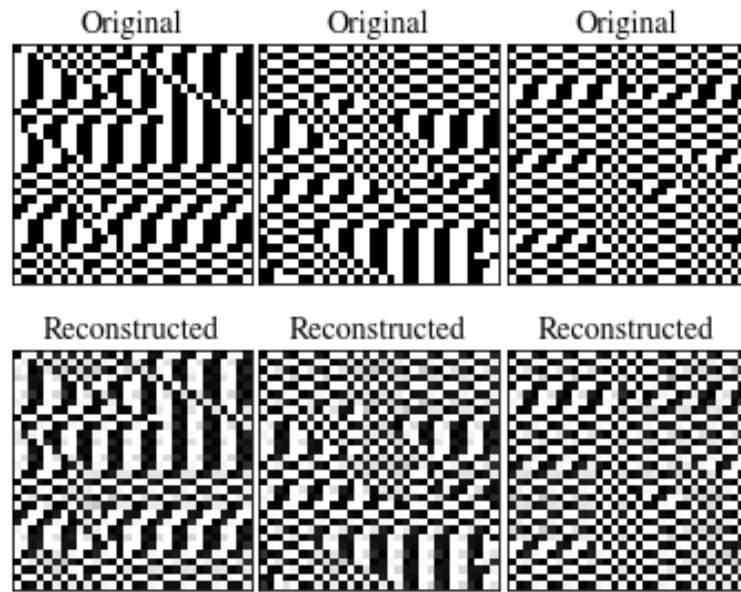


Figura 3.16: Fila superior: tres configuraciones correspondientes a $T = 0.02$ para tres realizaciones diferentes del sistema con $J_2 = 1/2$. Fila inferior: reconstrucciones respectivas realizadas por el autoencoder II.

La figura 3.17 muestra el error de reconstrucción medio computado con el autoencoder II como función de la temperatura para diferentes valores de J_2 . El comportamiento es similar a aquél de la figura 3.14. Mientras el sistema se encuentre en la fase en la cual el autoencoder fue entrenado, el error de reconstrucción disminuye a baja temperatura, relativo al valor compartido de altas temperaturas.

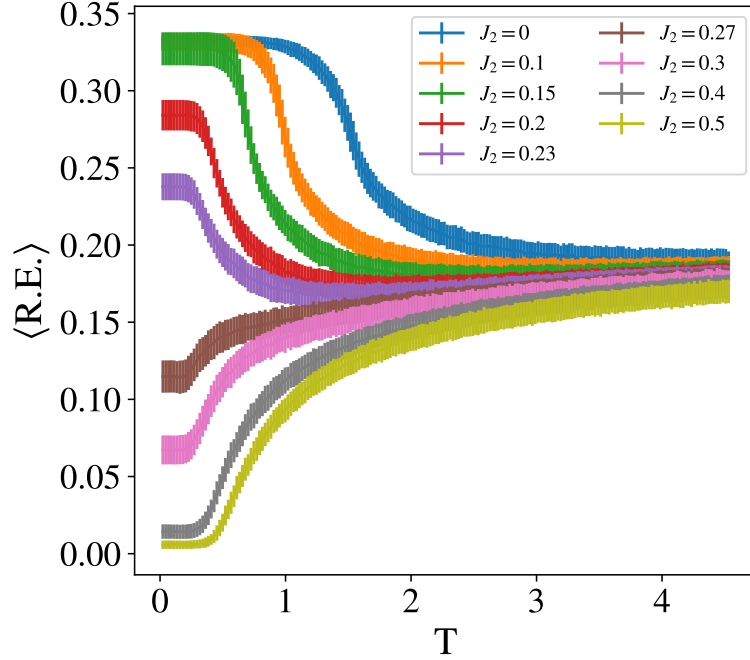


Figura 3.17: Error de reconstrucción medio versus temperatura para diferentes valores de J_2 , calculado con el autoencoder II. Las barras de error muestran la desviación estándar calculada con 400 realizaciones del sistema.

La figura 3.18 muestra el error de reconstrucción para las 400 realizaciones del sistema con $J_2 = 0$. La escala de color en este caso corresponde a la magnetización alternada por sitio m , es decir, el parámetro de orden en la fase de Néel. Cada panel corresponde a un entrenamiento equivalente pero diferente. La línea vertical a trazos corresponde a la temperatura crítica analítica del sistema en el límite termodinámico.

En este caso el error de reconstrucción no solo aumenta, señalando la transición de fase, sino que se divide en dos partes. La información adicional del parámetro de orden permite probar que la bifurcación de los datos corresponde a la simetría Z_2 del hamiltoniano, rota en el estado fundamental. Sin embargo, no debería haber ninguna preferencia *a priori* por ninguna de las dos configuraciones de Néel para tener el error de reconstrucción más bajo. Por esta razón mostramos que entrenamientos equivalentes pero independientes pueden arrojar los dos resultados posibles. Por equivalentes pero independientes nos referimos a que todos los hiperparámetros de la red son iguales en ambos entrenamientos, pero en cada uno la inicialización aleatoria de los parámetros ajustables del autoencoder son diferentes. Además, en cada entrenamiento los datos son mezclados aleatoriamente antes de ser separados en lotes.

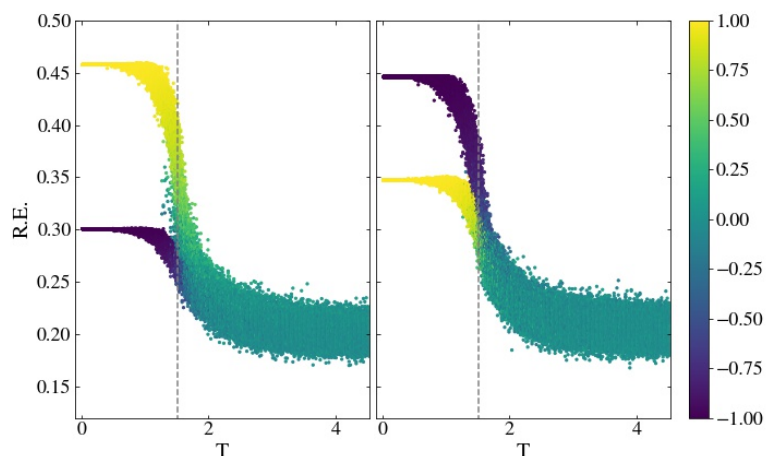


Figura 3.18: Error de reconstrucción versus temperatura para las 400 simulaciones del sistema no frustrado ($J_2 = 0$), calculado con el autoencoder II. La escala de color marca la magnetización alternada por sitio de cada configuración, es decir, el parámetro de orden de la fase de Néel. Cada panel corresponde a un entrenamiento equivalente pero independiente.

3.4.2 Diagrama de fases

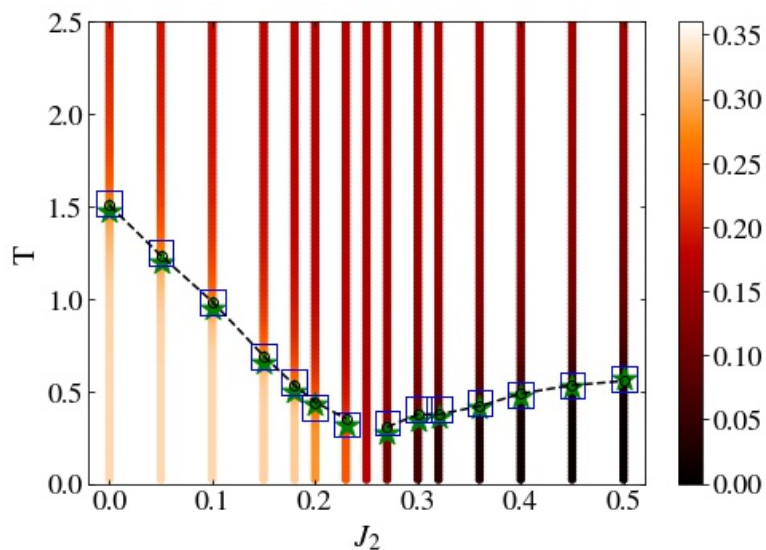


Figura 3.19: Diagrama de fases del sistema. En la escala de color, el error de reconstrucción computado con el autoencoder II. En estrellas verdes, los puntos de inflexión en las curvas de error de reconstrucción medio de la figura 3.17. en círculos negros vacíos unidos por líneas a trazos, las temperaturas de transición obtenidas con una CNN que usa a las estrellas verdes para su etapa de etiquetado (ver texto). En cuadrados azules vacíos, los máximos del calor específico en cada simulación Monte Carlo.

La figura 3.19 muestra el diagrama de fases $T - J_2$ del sistema, construido con el autoencoder II, donde la escala de color marca el error de reconstrucción. La parte superior del diagrama

corresponde a la fase paramagnética, donde el error de reconstrucción es aproximadamente igual para todos los valores de J_2 , como también muestra la figura 3.17. La parte inferior derecha (izquierda) del diagrama corresponde a las curvas en la figura 3.17 con $J_2 > 1/4$ ($J_2 < 1/4$) que tienen una disminución (un aumento) en su error de reconstrucción con respecto al valor de alta temperatura. Las estrellas verdes corresponden al punto de inflexión de la curva de error de reconstrucción medio para cada valor de J_2 en la figura 3.17, que son tomadas como el límite entre fases, generalizando el resultado de la figura 3.13.

Para corroborar que los puntos de inflexión estiman correctamente la ubicación de la transición de fase, realizamos una clasificación usando una red neuronal convolucional (CNN) en la cual etiquetamos los datos usando al punto de inflexión de la siguiente manera:

Dado un valor de J_2 , tomamos las configuraciones de espín alejadas de la temperatura que marca el punto de inflexión, T^* , es decir que estén a una temperatura T tal que $|T - T_c| > w$, con $w = 0.2$. Etiquetamos entonces las configuraciones de alta temperatura con 1 y las de baja temperatura con 0 y entrenamos a la red neuronal convolucional con estas configuraciones. Como la *accuracy* resultante en la etapa de validación es mayor a 0.999, podemos afirmar que la transición se encuentra dentro de la ventana de temperaturas removidas, es decir, cerca de T^* . Sabemos que un etiquetado incorrecto llevaría a una disminución en la *accuracy*[56].

Luego tomamos la red entrenada y la hacemos predecir las probabilidades de orden-desorden para configuraciones de espín en todo el rango de temperaturas, y encontramos la temperatura a la cual ambas curvas se cortan con probabilidad $p = 1/2$. Las temperaturas de transición correspondientes están graficadas en la figura 3.19 en círculos negros vacíos unidos por la línea negra a trazos. Aunque en estas clasificaciones estamos usando una arquitectura de aprendizaje supervisado, el etiquetado fue obtenido de los resultados previos no supervisados de la figura 3.17. Por esta razón, esta segunda predicción de las temperaturas de transición también puede considerarse como aprendizaje no supervisado.

Finalmente, para comparar, los cuadrados azules vacíos en la figura 3.19 corresponden al máximo en la curva de calor específico del sistema, que señala la transición para cada valor de J_2 . La arquitectura de la red neuronal convolucional que realiza la clasificación y sus hiperparámetros son presentados en la sección 3.4.5.

3.4.3 Entrenamientos a alta temperatura

Entrenamiento III: Datos de entrenamiento de alta temperatura

Aquí entrenamos un autoencoder con configuraciones de espín con $J_2 = 0$ y temperaturas $4 \leq T \leq 4.5$, que corresponden a la fase paramagnética.

El panel izquierdo de la figura 3.20 muestra el error de reconstrucción versus temperatura para configuraciones de espín con $J_2 = 0$, calculadas con el autoencoder III. Aunque el autoencoder fue entrenado en la fase paramagnética, puede reconstruir correctamente el orden de Néel de baja temperatura, e incluso muestra la bifurcación, como en la figura 3.18. Es importante remarcar que la situación recíproca no ocurre en los entrenamientos I y II, es decir, cuando el autoencoder aprende la fase de baja temperatura no es capaz de reconstruir configuraciones desordenadas, lo que constituye una tarea más compleja.

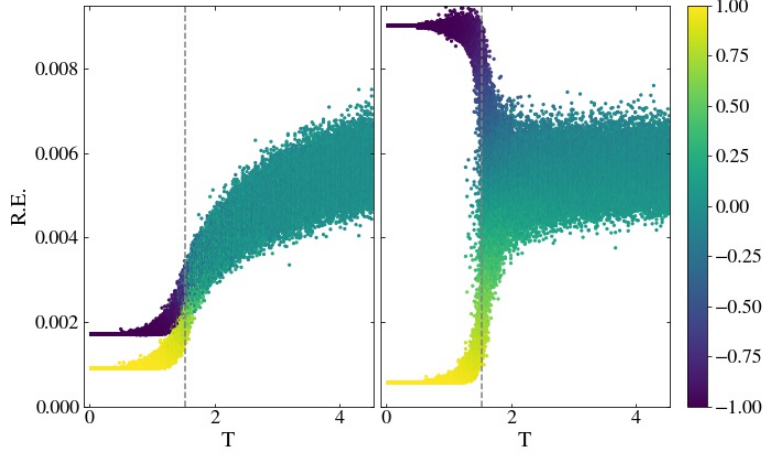


Figura 3.20: Error de reconstrucción versus temperatura para 400 realizaciones del sistema con $J_2 = 0$. El panel izquierdo (derecho) fue computado con el autoencoder III (IV). La línea vertical a trazos marca la temperatura crítica de transición en el límite termodinámico. La escala de color marca la magnetización alternada por sitio, (parámetro de orden en la fase de Néel).

Entrenamiento IV: Datos de entrenamiento aleatorios

Para verificar si el aprendizaje del orden de baja temperatura es una consecuencia de efectos de tamaño finito en los datos, entrenamos un autoencoder con arreglos de zeros y unos (pseudo-) aleatorios, que simulan estados a temperatura infinita en el sistema de espines de Ising. El panel derecho de la figura 3.20 muestra el error de reconstrucción como función de la temperatura para configuraciones de espín con $J_2 = 0$, computado con el autoencoder IV. El AutoEncoder puede reconstruir aproximadamente configuraciones de espín ordenadas y desordenadas (error máximo de aproximadamente de 0.009). Sin embargo la transición es todavía visible, y la bifurcación está nuevamente presente.

En el caso de sistemas con un parámetro de orden simple, como los modelos de Ising no frustrados, puede mostrarse que PCA o autoencoders con una única variable en su espacio latente pueden codificar las configuraciones de espín en una variable Z que está altamente correlacionada con el parámetro de orden del sistema[59]. Sin embargo, a altas temperaturas o haciendo escaleo de tamaños finitos estas dos cantidades pueden diferir[60].

El autoencoder del entrenamiento IV es capaz de reproducir en buena aproximación todas las configuraciones de espines del modelo de Ising $J_1 - J_2$ en la red hexagonal, como muestra la figura 3.21.

La figura 3.22 muestra la desviación estandar en el error de reconstrucción medio como función de la temperatura para diferentes valores de J_2 , usando el autoencoder IV. Para $J_2 < 1/4$ el error de reconstrucción se bifurca en la temperatura a la cual se ordena como en la figura 3.20 (no lo mostramos aquí), lo que corresponde al aumento en la dispersión a bajas temperaturas en la figura 3.22. Para $J_2 > 1/4$ no hay cambios en el error de reconstrucción medio que superen la desviación estándar a ninguna temperatura. La desviación estándar aumenta en la temperatura de transición (líneas a trazos) para todo J_2 , y todas las dispersiones convergen a un valor común a alta temperatura.

A bajas temperaturas, cada curva de dispersión tiene un plateau bien definido, con un

valor más alto que el valor de altas temperaturas. Esto permite separar datos de alta y baja temperatura alejados de cada transición de fase y etiquetarlos con esta información para realizar una clasificación como aquella de la figura 3.19, si fuese necesario.

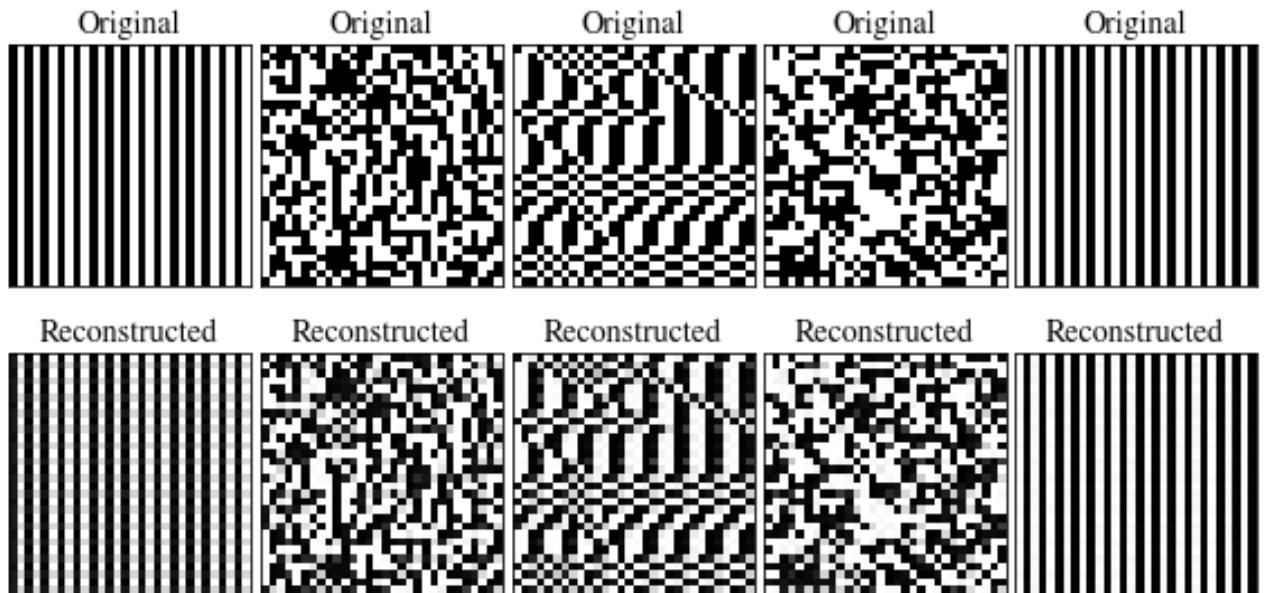


Figura 3.21: Reconstrucciones calculadas con el autoencoder IV. De izquierda a derecha, una configuración con orden de Néel con $m = -1$, $J_2 = 0$ y $T = 0.02$; una configuración paramagnética con $J_2 = 0$ y $T = 4.53$; una configuración de baja temperatura en presencia de frustración, con $J_2 = 1/2$ y $T = 0.02$; una configuración sintética pseudo aleatoria; y una configuración con orden de Néel con $m = +1$, $J_2 = 0$ y $T = 0.02$.

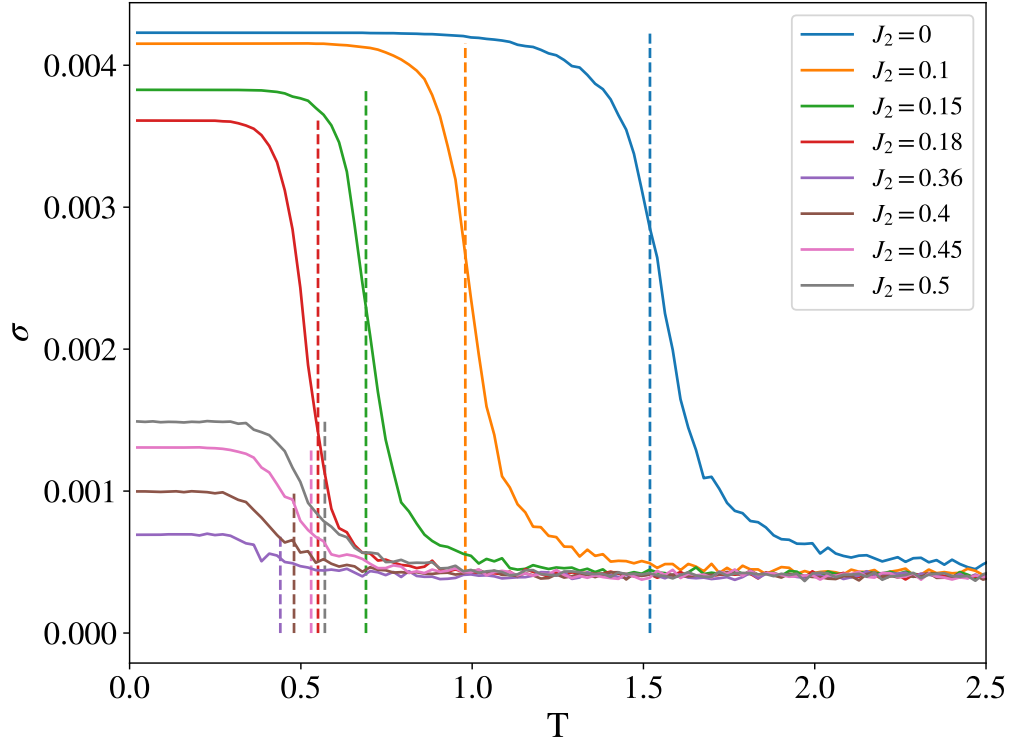


Figura 3.22: Desviación estándar en el error de reconstrucción como función de la temperatura para diferentes valores de J_2 , calculado con el autoencoder IV. En línea a trazos, las temperaturas de transición correspondientes a los máximos de calor específico en cada simulación Monte Carlo.

Mayor degeneración

Como muestran las figuras 3.18 y 3.20, el error de reconstrucción se divide en dos en la temperatura de transición, debido a la doble degeneración del estado fundamental de Néel el modelo de Ising antiferromagnético a primeros vecinos sobre la red hexagonal. Aquí naturalmente surge la pregunta de si el error de reconstrucción calculado por el autoencoder puede dividirse en n ramas en el caso de un estado fundamental con degeneración n . A partir de esta pregunta, decidimos explorar brevemente un sistema con un estado fundamental con degeneración 4. La figura 3.23 muestra el error de reconstrucción calculado con el autoencoder IV sobre datos del modelo de Ising antiferromagnético $J_1 - J_2$ sobre la red cuadrada, con $J_2/J_1 = 0.9$. Para este valor del cociente ente acoplamientos el estado fundamental tiene degeneración 4 y muestra orden de largo alcance de tipo tiras (o *stripes*, en inglés).

Es posible que en otro entrenamiento equivalente pero diferente el error de reconstrucción se divida en m ramas, con $m < n$. Por esta razón es importante repetir el proceso de entrenamiento y estudiar los respectivos resultados. Además, puede contarse la cantidad de configuraciones dentro de cada rama contruyendo un histograma en el error de reconstrucción para la temperatura mínima de trabajo. En este histograma (no mostrado aquí) si el estado fundamental tiene degeneración n y el conjunto de datos está balanceado correctamente, las barras tendrán alturas aproximadamente iguales para $m = n$ y alturas diferentes para $m < n$.

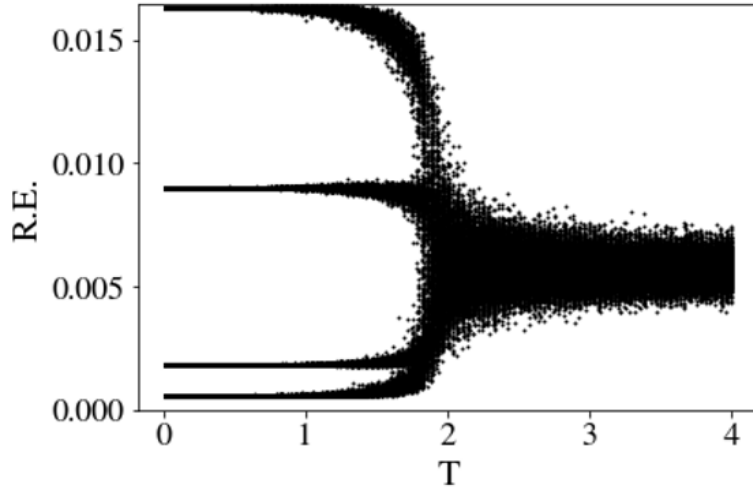


Figura 3.23: Error de reconstrucción como función de la temperatura para 400 realizaciones del modelo de Ising $J_1 - J_2$ antiferromagnético con $J_2/J_1 = 0.9$, calculado con el autoencoder IV.

3.4.4 Conclusiones

Hemos usado aprendizaje no supervisado para hallar transiciones de fase en el antiferromagneto de Ising $J_1 - J_2$ sobre la red hexagonal, a través de la técnica conocida como detección de anomalías.

Entrenamos AutoEncoders convolucionales para discriminar entre las diferentes fases magnéticas, eligiendo distintos conjuntos de entrenamiento en el espacio de parámetros del sistema. Monitoreamos el error de reconstrucción para localizar dónde ocurren las transiciones de fase. Empezamos entrenando un autoencoder para que reconstruya configuraciones completamente ordenadas (Entrenamiento I), correspondientes al modelo sin frustración, $J_2 = 0$, y observamos que la temperatura crítica del sistema coincide con el punto de inflexión en la curva de error de reconstrucción medio versus temperatura. Este criterio para separar las dos fases puede extenderse para conjuntos de datos con J_2 finito y permite construir el diagrama de fases $T - J_2$. En particular, calculamos el diagrama de fases entrenando exclusivamente con configuraciones de baja temperatura y alta frustración (Entrenamiento II). Además, validamos este resultado usándolo para generar las etiquetas empleadas para entrenar supervisadamente una red neuronal convolucional clasificadora, mostrando una interacción interesante entre diferentes redes neuronales. Adicionalmente, comparamos los resultados determinados por ambos métodos con los máximos del calor específico del sistema, encontrando un excelente acuerdo.

Luego, mostramos cómo el error de reconstrucción puede “levantar” la doble degeneración del estado fundamental de Néel, separándose en dos ramas para diferentes elecciones de entrenamiento. Destacamos que para el caso del entrenamiento IV, el AutoEncoder fue entrenado para reproducir matrices aleatorias cuyos elementos son ceros y unos, construidas con el generador de números aleatorios de Python, y que corresponden a una temperatura infinita en nuestro modelo. En este caso, el AutoEncoder puede reproducir en buena aproximación configuraciones de todas las fases del modelo, y es posible monitorear la desviación estándar, en lugar del error de reconstrucción, para detectar transiciones de fase. Finalmente, el AutoEncoder IV fue usado para reproducir configuraciones de espines del modelo de Ising antiferromagnético $J_1 - J_2$ sobre

la red cuadrada, con $J_2/J_1 = 0.9$ para mostrar que en este caso el error de reconstrucción a bajas temperaturas puede separarse en cuatro ramas, que es la degeneración del estado fundamental en este sistema.

3.4.5 Tecnicismos

AutoEncoders Convolucionales

El autoencoder está formado por dos capas convolucionales con funciones de activación ReLU, $\text{strides}=2$, tamaño de filtro=3, $\text{padding}=\text{'same'}$, regularización de *kernels* y *biases* de tipo *l2* y de magnitud 10^{-4} . La tasa de aprendizaje es de 10^{-3} y el tamaño de lote es de 256. Del conjunto de datos de entrenamiento tomamos el diez por ciento como conjunto de datos de validación. Cada capa convolucional tiene a continuación una capa de *dropout* con una tasa de *dropout* de 0.2. Las imágenes iniciales tienen un tamaño de 30×30 píxeles, y los mapas de características generados por el encoder son de 8×8 píxeles.

El decoder esta formado primero por dos capas convolucionales transpuestas con las mismas características que las capas convolucionales mencionadas. Luego, la capa final es convolucional, con una función de activación sigmoidea y un único filtro, que combina todos los mapas de características en una imagen final de salida.

Denotamos N_1 y N_2 al número de filtros en la primera y segunda capa de el encoder, respectivamente. La primera capa convolucional del decoder tiene, en cambio, N_2 filtros, y la segunda, N_1 .

En los entrenamientos I y II $N_1 = 16$ y $N_2 = 8$, mientras que en los entrenamientos III y IV $N_1 = 32$ y $N_2 = 16$. Aquí enfatizamos que el autoencoder no tiene un espacio latente de dimension N_2 , porque el encoder produce un conjunto de N_2 mapas de características, cada uno de dimensión 8×8 .

En los entrenamientos I, II y III el número de épocas usado está entre 100 y 200. El entrenamiento IV consistió en 2 épocas sobre 250000 configuraciones pseudo-aleatorias generadas en python. Es interesante destacar aquí que un número menor de configuraciones para el entrenamiento IV lleva a un sobreajuste de la red neuronal, la cual es incapaz de reconstruir nuevas imágenes aleatorias en la etapa de examen.

La entrada de la red neuronal está normalizada. Espines hacia arriba corresponden a un +1 mientras que espines hacia abajo corresponden a un 0. Luego, cuando computamos el error cuadrático medio entre la entrada y la salida del autoencoder, el máximo error cuadrático medio posible es 1. Recordamos aquí que para un conjunto X de N imágenes bidimensionales de tamaño $L \times L$, el error cuadrático medio de reconstrucción del conjunto es

$$MSE(X) = \frac{1}{NL^2} \sum_{n=1}^N \sum_{i,j=1}^L |x_{ij}^n - \hat{x}_{ij}^n|^2, \quad (3.13)$$

donde $x^n \in X$ y \hat{x}^n es la reconstrucción de x^n hecha por el autoencoder.

Classificador convolucional (CNN)

La red neuronal está formada por dos capas convolucionales con funciones de activación ReLU, con tamaño de filtro 3, sin agregar ceros e los bordes (*padding*) y sin regularizadores. La

tasa de aprendizaje es de 10^{-4} y el tamaño de lote es de 256. Cada capa convolucional tiene a continuación una capa *max-pooling* con *pool-size*=2. Luego, hay una capa de *flattening*, conectada a una capa densa de 16 neuronas con funciones de activación ReLU. La capa final tiene dos neuronas con funciones de activación softmax.

Como la temperatura máxima en nuestro Monte Carlo es 4.53 (en unidades de J_1) y algunas temperaturas de transición pueden ser tan pequeñas como algunas pocas décimas, los datos de entrenamiento están altamente desbalanceados. Para balancear nuestros datasets en la clasificación simplemente descartamos configuraciones de altas temperaturas. Para $J_2 > 1/4$ los conjuntos de datos contienen temperaturas $0.02 \leq T \leq 1$ y para $J_2 < 1/4$ temperaturas $0.02 \leq T \leq 3$. De cada conjunto de datos de entrenamiento tomamos el 10 por ciento como conjunto de validación. En cada entrenamiento, la precisión en validación es mayor a 0.99.

Estimación numérica del punto de inflexión

Para obtener numéricamente el punto de inflexión en la curva de error de reconstrucción medio aplicamos un filtro de Savitzky-Golay de la librería *scipy*, que permite suavizar los datos ajustando localmente sucesivos polinomios de bajo orden. Esta librería tiene dos hiperparámetros esenciales que son el tamaño de ventana (cantidad de puntos que se usan para cada ajuste) y el grado del polinomio a ajustar. En este trabajo usamos tamaño de ventana de 21 y polinomios de orden 3. Finalmente tomamos la derivada numérica y buscamos su máximo.

3.4.6 Análisis de componentes principales (PCA)

PCA es un algoritmo cuyo objetivo principal es reducir la dimensionalidad de un conjunto de datos. El método original, usado en este trabajo, se basa en una transformación lineal, por lo que se lo puede encontrar citado como *linear PCA* para diferenciarlo de sus generalizaciones no lineales, como *kernel PCA*.

Supongamos tener un conjunto de n variables aleatorias reales X_i en un vector $\mathbf{X} = (X_1, \dots, X_n)^T$. La matriz de covarianza $K_{\mathbf{X}\mathbf{X}}$ está definida elemento a elemento según

$$(K_{\mathbf{X}\mathbf{X}})_{ij} = \langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle \quad (3.14)$$

donde la notación de brackets $\langle \cdot \rangle$ indica el valor medio, o matricialmente, según

$$K_{\mathbf{X}\mathbf{X}} = \langle (\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})^T \rangle = \langle \mathbf{X}\mathbf{X}^T \rangle - \mu_{\mathbf{X}}\mu_{\mathbf{X}}^T \quad (3.15)$$

donde $\mu_{\mathbf{X}} = \langle \mathbf{X} \rangle$. En la diagonal se encuentran las varianzas de cada variable aleatoria por separado, mientras que fuera de la diagonal se encuentran las covarianzas entre variables aleatorias diferentes. Las covarianzas miden la correlación entre las fluctuaciones de dos variables aleatorias.

$K_{\mathbf{X}\mathbf{X}}$ es simétrica y real, por lo que puede diagonalizarse mediante una transformación ortogonal \mathbf{V} , que tiene por columnas a los autovectores de $K_{\mathbf{X}\mathbf{X}}$. Cada autovector de $K_{\mathbf{X}\mathbf{X}}$ se denomina una dirección principal del conjunto de datos, mientras que cada autovalor asociado se denomina una componente principal del conjunto de datos. La reducción de dimensionalidad puede obtenerse entonces al retener sólo algunas direcciones principales, aquellas con las mayores componentes principales. Este proceso corresponde entonces a describir un conjunto de datos

dado con las direcciones que contienen la mayor dispersión, descartando las direcciones en que la dispersión es más pequeña.

En nuestro trabajo, cada configuración de espines x_n , $1 \leq n \leq 80000$ representada por una matriz de 30×30 es vectorizada, es decir, transformada en un vector concatenando las filas de la matriz original. Este vector tiene entonces 900 componentes.

La figura 3.24 muestra representaciones 2-dimensionales obtenidas con PCA, para $J_2 = 0$ a izquierda y $J_2 = 1/2$ a derecha, donde la barra de color marca la temperatura. Retuvimos las dos direcciones principales Z_1 y Z_2 con mayores componentes principales. Puede verse en ambas figuras que a alta temperatura las configuraciones están distribuidas uniformemente en un disco centrado en cero, ya que están descorrelacionadas. A baja temperatura en cambio puede verse que para $J_2 = 0$ las configuraciones convergen a dos puntos situados en $Z_2 = 0$ y $Z_1 \approx \pm 30$. Para $J_2 = 1/2$ no se observa que a baja temperatura el algoritmo pueda agrupar los puntos en ningún lugar del plano, como consecuencia de la alta degeneración.

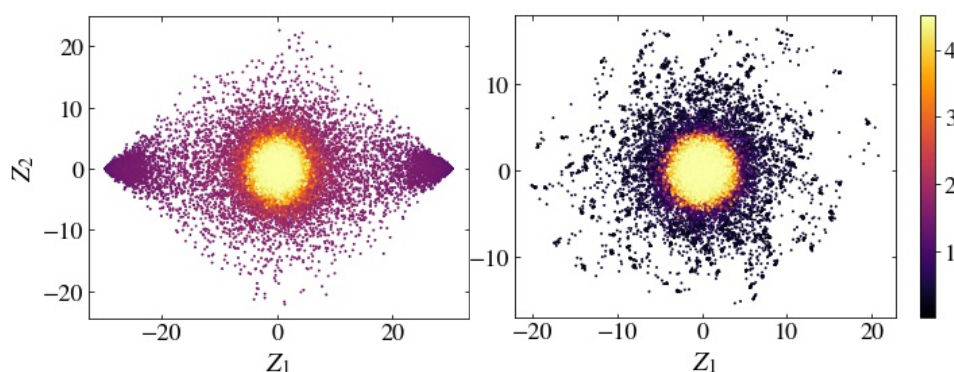


Figura 3.24: Representación de dimensión 2 del sistema con $J_2 = 0$ a la izquierda y $J_2 = 1/2$ a la derecha. Cada punto corresponde con una configuración de espines, y la escala de colores marca la temperatura.

La figura 14 muestra los 900 autovalores de PCA para diferentes valores de frustración. Cada lista de autovalores está normalizada dividiendo por su máximo valor, y ordenada en orden decreciente. Para $J_2 = 0$ (o $J_2 \ll 1/4$) solo hay una componente principal relevante, que corresponde al parámetro de orden de la fase ordenada, el cual es una combinación lineal de los espines. Puede verse que aumentar la frustración aumenta el número de componentes principales relevantes. Es interesante remarcar que el área bajo la curva es máxima alrededor de $J_2 = 1/4$, donde el sistema tiene su transición de fase de baja temperatura. La tendencia hacia un mayor número de componentes principales tiene lugar donde el sistema está máximamente frustrado. Esta figura muestra que si $J_2 \ll 1/4$ cada configuración de espines puede estar representada en buena aproximación con un único valor real, su parámetro de orden. Enfatizamos aquí que el parámetro de orden puede ser directamente observado en la dirección principal obtenida sin ningún tipo de información previa sobre su existencia o forma.

Al aumentar J_2 no es posible representar a nuestro conjunto de datos con unas pocas direcciones principales obtenidas mediante la transformación lineal que diagonaliza la matriz de covarianza, lo que lleva a la necesidad de utilizar un algoritmo no lineal, como *kernel PCA* o *autoencoders*.

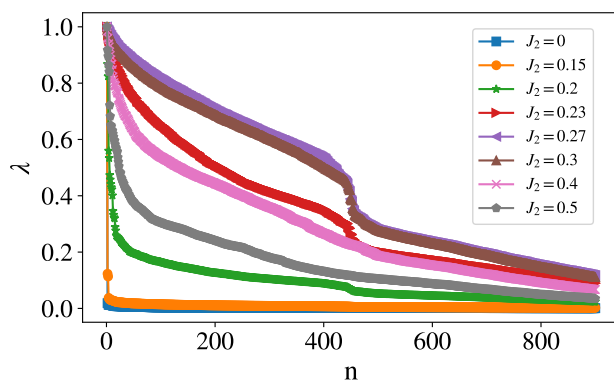


Figura 3.25: Componentes principales λ de cada conjunto de datos, para diferentes valores de J_2 . Para cada caso las 900 componentes principales están normalizadas dividiendo por la mayor, y ordenadas en orden decreciente.

4 Conclusiones generales

En esta tesis doctoral hemos estudiado analítica y numéricamente sistemas magnéticos de muchos cuerpos en baja dimensión. Comenzamos estudiando la llamada Cadena Kagomé, un material uno-dimensional frustrado geoméricamente, ampliando los resultados obtenidos durante mi trabajo de tesina de grado. Esta geometría es de interés porque modela un material existente[9], y porque el modelo de Heisenberg asociado puede desarrollar un plateau de magnetización con magnones localizados espacialmente, producto de una relación de dispersión nula para estas excitaciones, o bien, su “banda chata”[11]. En particular, nos centramos en el cómputo analítico de un Hamiltoniano perturbativo en el límite de plaquetas débilmente interactuantes, el cual nos permitió hallar una segunda solución exacta antes desconocida para un estado fundamental exacto, factorizado, correspondiente a un cristal de magnones, generalizando los resultados de la referencia [11]. Además, combinamos esta técnica analítica con el cálculo numérico de curvas de magnetización y correlaciones cuánticas usando DMRG. Dado que cada técnica tiene sus ventajas y desventajas, la combinación de diferentes técnicas para abordar el mismo problema es en general deseable. En este caso, si bien DMRG es considerada básicamente una técnica exacta en 1D, no es útil para buscar puntos específicos del sistema donde el mismo presente singularidades, dado que el espacio de parámetros es demasiado grande. Para esta tarea, típicamente es deseable algún cálculo analítico. Una vez encontrado el punto o la región de interés, este/a puede ser estudiado/a en detalle numéricamente con DMRG, como en nuestro trabajo.

Cadenas de espín con cristales de magnones como estados fundamentales proporcionan un escenario interesante para agregar complejidades adicionales, como interacciones a vecinos más lejanos, desorden, un campo magnético transversal, o interacciones antisimétricas. En la sección 2.2 estudiamos estas últimas como perturbación al cristal de magnones presente en la cadena diente de sierra. Comenzamos desarrollando una teoría de campos efectiva de baja energía para describir el cristal de magnones presente en el estado fundamental de la cadena diente de sierra en ausencia de interacciones antisimétricas. Observamos que la banda chata producto de la frustración geométrica del sistema se manifiesta en la teoría efectiva dando un coeficiente de rigidez nulo. Nos enfocamos luego en el estudio numérico de la inclusión de interacciones antisimétricas sobre el sistema, usando DMRG. Identificamos dos fases conocidas, que son las de Líquido de Luttinger y la de Meissner, y observamos una fase que denominamos de ‘baja rigidez’ donde las corrientes de espín tienen un comportamiento marcadamente diferente a los convencionales a causa de la banda chata. En particular, el aumento en la magnitud de las corrientes de espín en la fase de baja rigidez puede interpretarse gracias a la teoría de campos efectiva, que le dio nombre a la fase.

Más adelante, hacia el final del capítulo 2, abordamos el estudio de un modelo de Heisenberg dos-dimensional sobre la geometría de bicapa hexagonal. A partir de una teoría de campos efectiva describimos un proceso de orden por desorden, en donde el orden colineal de Néel se ve favorecido energéticamente por la presencia tanto de fluctuaciones térmicas como cuánticas.

Este resultado es relevante para el material $\text{Bi}_3\text{Mn}_4\text{O}_{12}(\text{NO}_3)$ en donde hay experimentos que sugieren la existencia de orden por desorden[29].

Por otra parte, nuestra teoría efectiva permite describir el conocido estado fundamental compuesto por un arreglo de singletes entre capas hexagonales, presente en nuestro modelo cuando el acoplamiento entre capas es el dominante, a partir del anulamiento del coeficiente de rigidez de la teoría, de manera análoga a lo que ocurría en los cristales de magnones en una dimensión espacial.

Nuestra teoría de campos efectiva es especialmente útil en el caso de nuestra geometría bidimensional frustrada, ya que muchas técnicas populares de estudio fallan. Por ejemplo, la frustración impide usar Monte Carlo cuántico, mientras que la bidimensionalidad impide usar DMRG, diagonalización exacta¹ o bosonización.

En el capítulo 3 nos dedicamos a la implementación de técnicas de aprendizaje automático en el estudio de sistemas magnéticos de Ising en dos dimensiones. Nos enfocamos en modelos clásicos debido a la posibilidad de generar grandes volúmenes de datos simulando sistemas magnéticos frustrados a temperatura finita via métodos Monte Carlo.

En la sección 3.3 nos enfocamos en la implementación de técnicas de aprendizaje supervisado para clasificar las fases de antiferromagnetos de Ising frustrados. Nos centramos en el concepto de transferencia de aprendizaje, entrenando redes neuronales profundas en modelos con un valor fijo de J_2 , y monitoreando su rendimiento al modificar J_2 . Así construimos el diagrama de fases del modelo $J_1 - J_2$ antiferromagnético de Ising en la red cuadrada con dos entrenamientos diferentes en $J_2 = 0$ y $J_2 = 1$ mostrando un buen acuerdo con resultados conocidos. Luego, aplicamos la técnica denominada aprendizaje por confusión para reobtener el conocido valor de J_2 en el cual el modelo tiene una transición de fases de baja temperatura. Este enfoque es interesante porque, aún usando arquitecturas de aprendizaje supervisado, no necesita de información previa sobre el sistema para lograr clasificar los datos.

El último trabajo presentado en esta tesis doctoral consiste en el uso de AutoEncoders convolucionales para determinar transiciones de fase en el antiferromagneto de Ising $J_1 - J_2$ sobre la red hexagonal, a través de la técnica de detección de anomalías. Construimos el diagrama de fases del modelo a partir del monitoreo del error de reconstrucción de la red neuronal, que presenta puntos de inflexión en los puntos de transición. Aquí cabe destacar que nuestros AutoEncoders completamente convolucionales no tienen un espacio latente de baja dimensionalidad, como es usual, ya que no es necesario (e incluso puede resultar problemático) para la detección de anomalías.

Por fuera de este manuscrito queda un trabajo² en donde usamos AutoEncoders completamente conectados para el estudio del llamado *Neural Network Flow* de configuraciones de espines. Este flujo fue introducido recientemente en las referencias [62, 63], y se genera a partir de consecutivas reconstrucciones de una configuración de espines usando AutoEncoders o *Restricted Boltzmann Machines* (RBMs). Nuestro trabajo se centra en el estudio del flujo de configuraciones de espines del ferromagneto de Ising en dos dimensiones usando un AutoEncoder completamente conectado. Estudiamos la dependencia del flujo con los hiperparámetros del modelo y refutamos la declaración en [62] de que el flujo de configuraciones fluye al punto crítico del

¹Aquí decimos que la técnica falla porque no pueden alcanzarse tamaños mayores a unas pocas celdas unidad. Para nuestro material de interés el problema es todavía más grave que lo usual, dado que tiene espines $S = 3/2$.

²Actualmente se encuentra en proceso de evaluación, y su versión preliminar puede encontrarse en arxiv [61].

sistema.

También queda por fuera de este manuscrito el uso de redes neuronales como *ansatz* variacional. Recientemente se ha visto que los llamados modelos autorregresivos, como las redes neuronales recurrentes, pueden usarse como modelos generativos en física estadística. Esto significa que pueden ser utilizados para generar datos, que en nuestro contexto corresponden típicamente a configuraciones de espines. Una de las principales ventajas de estos algoritmos es que no necesitan ser entrenados con conjuntos de datos conocidos, ya sea experimentales o sintéticos. En cambio, su entrenamiento puede efectuarse únicamente con información sobre el Hamiltoniano del sistema y sus simetrías. Estos modelos se han utilizado tanto en sistemas clásicos a temperatura finita [35, 64] como en sistemas cuánticos a temperatura cero [36, 65] y constituyen nuestro objeto de estudio actual. En este enlace³ puede verse y descargarse un código escrito en tensorflow 2, que usa redes neuronales recurrentes como *ansatz* variacional para generar configuraciones de espines del modelo de Ising ferromagnético a primeros vecinos en una dimensión espacial, siguiendo las referencias [35, 64]. Los parámetros entrenables de la red son ajustados minimizando una energía libre variacional parametrizada por la red neuronal, de manera tal que luego del entrenamiento la red es capaz de generar configuraciones de espines a partir de las cuales uno puede calcular valores de expectación. La modificación del algoritmo para el estudio de sistemas cuánticos siguiendo a [36] no se encuentra lista al día de la fecha de presentación de esta tesis doctoral, y constituye un objetivo para el futuro cercano.

Para concluir, podemos decir que hemos estudiado sistemas magnéticos con dos enfoques marcadamente diferentes. Por un lado tenemos el enfoque tradicional, con técnicas establecidas hace decenas de años, numéricas como DMRG y analíticas como teorías de campos efectivas, y por otro lado tenemos las herramientas de aprendizaje automático y ciencia de datos, altamente maleables y actualmente en proceso de expansión continuo.

³<https://github.com/acevedo-s/VNA-Ising-chain>

5 Apéndice

5.1 Density Matrix Renormalization Group (DMRG)

DMRG es una técnica numérica introducida en los 90'[66] centrada en un truncado eficiente del espacio de Hilbert de sistemas cuánticos fuertemente correlacionados, basado un proceso de decimación. El algoritmo ha alcanzado los mayores niveles de precisión de la historia en el estudio de sistemas en una dimensión espacial, donde suele considerarse básicamente exacto. Permite calcular cantidades estáticas, dinámicas, entre otras, en sistemas muy diversos. Una revisión de DMRG puede hallarse en la referencia [67].

En la actualidad, DMRG tiene bases teóricas fuertemente establecidas, que permiten saber sobre qué sistemas el algoritmo funciona correctamente y sobre qué sistemas no. DMRG construye estados que se denominan estados de productos de matrices (*matrix product states*) que representan eficientemente a los estados de sistemas cuánticos en una dimensión espacial, pero en general no a sistemas cuánticos en dimensión mayor. Un desarrollo exhaustivo de los fundamentos teóricos de esta técnica numérica escapa a los fines de esta tesis doctoral. Para implementarla en nuestros trabajos, utilizamos las librerías de ALPS. Sin embargo, a continuación presentamos por completitud los conceptos centrales del método con una cadena de espín 1/2. Luego, finalizamos este apéndice con una breve introducción a los estados de productos de matrices y su relación con los estados fundamentales de sistemas cuánticos en una dimensión.

5.1.1 Líneas centrales del algoritmo

Consideremos una cadena de espines 1/2 de longitud infinita, donde separamos la cadena en un conjunto de bloques A idénticos. Escribimos a la cadena entera como la concatenación de todos los bloques que la conforman, $A \cdots A \cdots$ y llamamos l a la longitud de cada bloque. El Hamiltoniano que describe a un bloque A de longitud l que puede estar en m estados linealmente independientes es $H_A^{(l,m)}$, de dimensiones $m \times m$, con $m = 2^l$.

En su forma más básica, los cinco pasos centrales en una iteración del algoritmo ‘de tamaño infinito’ son los siguientes. (i) Diagonalizar el Hamiltoniano de un ‘superbloque’, un sistema formado por dos bloques y dos sitios, como muestra la figura 5.1, al que denotamos $A \cdot A$, extrayendo el estado fundamental. Este superbloque tiene longitud $2l + 2$, y su Hamiltoniano tiene dimensión 2^{2l+2} . (ii) Construir la matriz densidad ρ del estado fundamental del superbloque, y a partir de ella obtener la matriz densidad reducida del bloque A' , definido por un bloque A al que se le agrega un sitio adicional (ver figura 5.1). Suponiendo que el estado fundamental del superbloque es el estado puro¹ $|\psi\rangle$, la matriz densidad del estado fundamental es $\rho = |\psi\rangle\langle\psi|$,

¹Para diferenciar entre estados puros, mixtos, separables y entrelazados, pueden consultarse <https://www.quantiki.org/wiki/mixed-states> <https://www.quantiki.org/wiki/pure-states>
<https://www.quantiki.org/wiki/separable-and-entangled-states>

y la matriz densidad reducida del subsistema A' es la traza parcial $\rho_{A'} = \text{Tr}_{B'}\rho$, con B' el complemento de A' en el superbloque. (iii) Diagonalizar la matriz densidad reducida, para hallar sus m autovalores más grandes y sus respectivos autovectores. (iv) Cambiar de base con un cambio de base no cuadrado O de dimensiones $m \times n$, con $n = 2^{l+1}$, según

$$H_A^{(l+1,m)} = OH_{A'}^{(l,m)}O^\dagger, \quad (5.1)$$

donde las filas de O son los m autovectores seleccionados de la matriz densidad reducida $\rho_{A'}$. El Hamiltoniano $H_A^{(l+1,m)}$ permite calcular el estado fundamental de un sistema de longitud $l+1$ sin aumentar la dimensión m del espacio de Hilbert del bloque A original. (v) Volver a iterar desde (i) hasta llegar al tamaño deseado, donde en cada paso se actualiza según (5.1) el Hamiltoniano que describe a un bloque por un Hamiltoniano que tiene las mismas dimensiones pero describe al estado fundamental de un bloque de mayor longitud.

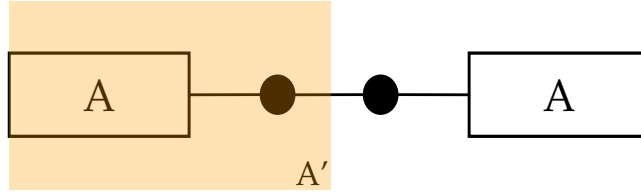


Figura 5.1: Superbloque. Cada rectángulo corresponde a un bloque A , cada círculo corresponde a un sitio, y la zona sombreada corresponde a un bloque agrandado en un sitio, A' . La parte no sombreada del superbloque corresponde B' , el complemento de A' .

Es importante remarcar que el método genera m estados para un bloque específicamente optimizados para describir únicamente al ‘estado objetivo’, el estado fundamental, aunque el algoritmo puede generalizarse para poder describir también algunos de los primeros estados excitados. Si se desea estudiar el estado fundamental de un sistema de longitud L fija, el algoritmo es adaptable[68]. Puede demostrarse[66] que la elección de los m autoestados de $\rho_{A'}$ con mayores autovalores corresponde a la aproximación del estado fundamental con m estados más eficiente posible, en el sentido que minimiza la distancia entre el estado fundamental exacto y su aproximación con m estados.

5.1.2 Estados de productos de matrices

Los estados de productos de matrices (*matrix product state*) son importantes en información cuántica y en materia condensada cuántica, tanto desde un punto de vista teórico como práctico en la simulación numérica de sistemas cuánticos. Antes de introducirlos, mencionamos los siguientes teoremas[69].

Teorema 1 (Lieb-Robinson): Existencia de una velocidad de grupo máxima

Supongamos tener un modelo uno-dimensional de N sitios, donde en cada sitio ubicamos un sistema cuántico de d niveles. Supongamos tener un Hamiltoniano \mathcal{H} con interacciones a primeros vecinos con acoplamientos finitos. Es decir, suponemos que $\mathcal{H} = \sum_{j=1}^N \mathcal{H}_{j,j+1}$, tal que $\|\mathcal{H}_{j,j+1}\| \leq J$ para algún $J > 0$. Entonces existe una velocidad de grupo máxima $v > 0$

y existen $\mu, c > 0$ tales que para cualesquiera operadores A y B con soporte en conjuntos disjuntos X e Y , se tiene que

$$|[A(t), B]| \leq c \|A\| \|B\| \exp(-\mu[\text{dist}(X, Y) - vt]), \quad (5.2)$$

donde la distancia entre conjuntos se toma como $\text{dist}(X, Y) = \min_{i \in X, j \in Y} (|i - j|)$, y $A(t) = e^{i\mathcal{H}t} A e^{-i\mathcal{H}t}$. Además, la velocidad v , denominada de Lieb-Robinson, es de orden J .

Teorema 2 (Hastings): Ley de áreas para cadenas de espín gapeadas

Supongamos tener un modelo descrito por un Hamiltoniano como el del teorema 1. Supongamos también que \mathcal{H} tiene un estado fundamental único, con un gap $\Delta E > 0$ al primer estado excitado. Si se considera un bloque $I = \{1, \dots, n\}$, entonces la entropía de entrelazamiento² $S(\rho_I)$ cumple

$$S(\rho_I) \leq S_{max} = c_0 \xi \log(6\xi) \log(d) 2^{6\xi \log(d)}, \quad (5.3)$$

para una constante numérica $c_0 > 0$ de orden 1, $\xi = \max(2v/\Delta E, \xi_C)$, v es la velocidad de Lieb-Robinson y ξ_C es otra constante de orden 1.

Lo importante del teorema 2 es que nos muestra que los estados fundamentales de sistemas cuánticos con gap en una dimensión espacial cumplen una ley de áreas para su entropía de entrelazamiento. Si en cambio el estado fundamental no tiene gap, la entropía de entrelazamiento crece logarítmicamente con el tamaño del subsistema. Esta característica es central para probar que cadenas cuánticas de longitud N con interacciones locales y simetría de traslación³, tanto críticas como gapeadas tienen estados fundamentales aproximables por estados de productos de matrices con un número de parámetros que crece polinómicamente con N [71, 72].

Para un sistema usual en mecánica estadística, la entropía debería en cambio cumplir una ley de volumen, es decir, debería ser proporcional a N . Ley de áreas en una dimensión espacial para la entropía de entrelazamiento significa que esta está acotada por una constante independiente de N . Si en vez de observar el estado fundamental tomásemos un estado genérico, su entropía de entrelazamiento cumple una ley de volumen.

A continuación presentamos a los estados de productos de matrices⁴, definidos en un arreglo uno-dimensional de N sitios ocupados por sistemas cuánticos de d niveles, con condiciones de contorno periódicas. Para cada sitio se introducen dos partículas virtuales de D estados, donde D puede llamarse la dimensión de enlace (*bond dimension*), auxiliar, o virtual. Cada uno de estos subsistemas virtuales D -dimensionales se inicializan o preparan en un estado máximamente entrelazado con uno de sus vecinos, habiendo dos partículas y dos vecinos por sitio. El estado inicial es entonces un estado producto de N estados máximamente entrelazados de dos partículas,

²Muchos conceptos de información cuántica y computación cuántica pueden consultarse en el libro de Nielsen[70]

³También existe un criterio de aproximabilidad por MPSs que incluye sistemas sin invarianza de traslación basado en el escaleo de las entropías de Renyi[69].

⁴Bibliografía adicional sobre *tensor networks* puede hallarse en <https://tensornetwork.org>

$$|\psi_D^0\rangle = \frac{1}{D^{N/2}} \sum_{\gamma_1, \dots, \gamma_N=1}^D |\gamma_0\gamma_1\rangle |\gamma_1\gamma_2\rangle |\gamma_2\gamma_3\rangle \dots |\gamma_{N-1}\gamma_N\rangle, \quad (5.4)$$

donde tomamos condiciones de contorno periódicas $N \equiv 0$. Luego, se aplican N operadores lineales locales a cada uno de los estados $\frac{1}{\sqrt{D}} \sum_{k=1}^D |\gamma_i\gamma_{i+1}\rangle$ con los operadores

$$A^{(i)} = \sum_{j_i=1}^d \sum_{\alpha_i, \beta_i=1}^D A_{j_i\alpha_i\beta_i}^{(i)} |j_i\rangle \langle \alpha_i, \beta_i|, \quad (5.5)$$

con $i = 1, \dots, N$, lo que da lugar al estado⁵ de producto de matrices

$$|\psi\rangle = \sum_{j_1, \dots, j_N=1}^d Tr \left[A^{(1)}[j_1] \dots A^{(N)}[j_N] \right] |j_1, \dots, j_N\rangle. \quad (5.6)$$

A continuación mencionamos tres propiedades de los estados de productos de matrices, en formato de lemas.

Lema 1

Los estados de productos de matrices están parametrizados por una cantidad de parámetros polinómica con respecto al número de sitios del sistema.

Esta propiedad puede verse a partir de que aplicamos N operadores locales de $d \times D \times D$ para construir un MPS genérico, por lo que el número de parámetros en un MPS es de orden NdD^2 .

Lema 2

La entropía de entrelazamiento para estados de productos de matrices en una dimensión espacial está acotada por una constante independiente de la cantidad de sitios N del sistema.

Lema 3

Todo estado puro puede descomponerse como un estado de producto de matrices.

La prueba de este Lema puede encontrarse en la sección 4 del review [67], y se basa en usar descomposición en valores singulares repetidas veces. Para estados que no tienen ley de áreas, D debe crecer exponencialmente con el tamaño del sistema para obtener una descomposición exacta.

5.2 Simulación Monte Carlo

Siguiendo al libro de Newman y Barkema[73], describimos a continuación las principales bases teóricas de una simulación Monte Carlo pensada para simular computacionalmente el modelo

⁵La comunidad de información cuántica tiende a no escribir los factores de normalización, dado que siempre tienen que estar ahí para que $|\psi\rangle$ sea un estado. Aquí adoptamos la misma convención.

de Ising.

Supongamos que nuestro sistema de estudio está en el estado μ . Definimos $R(\mu \rightarrow \nu)dt$ como la probabilidad de que el sistema esté en el estado ν un tiempo dt después. $R(\mu \rightarrow \nu)$ es la tasa de transición desde el estado μ al ν , y la asumimos independiente del tiempo. Definimos un conjunto de pesos $w_\mu(t)$ que representan la probabilidad de que el sistema se encuentre en el estado μ al tiempo t . La denominada ecuación maestra para la evolución de las probabilidades de ocupación en términos de las tasas de transición es

$$\frac{dw_\mu(t)}{dt} = \sum_\nu \left(w_\nu(t)R(\nu \rightarrow \mu) - w_\mu(t)R(\mu \rightarrow \nu) \right). \quad (5.7)$$

Por supuesto, las probabilidades w_μ cumplen

$$\sum_\mu w_\mu(t) = 1, \quad \forall t. \quad (5.8)$$

Luego, dado un observable Q , su valor de expectación al tiempo t es

$$\langle Q \rangle = \sum_\mu w_\mu(t)Q_\mu, \quad (5.9)$$

donde Q_μ es el valor que toma Q en el estado μ .

La interpretación física detrás del valor de expectación anterior es la de tener un conjunto muy grande de sistemas idénticos, cada uno en su baño térmico, y medir en cada uno de ellos instantáneamente el valor de Q , para luego tomar la media de todos los valores.

Otro punto de vista del valor de expectación anterior es uno donde pensamos a $\langle Q \rangle$ como un promedio temporal del operador Q . Este enfoque asume que en un determinado período de tiempo en el que estamos midiendo Q , como en un experimento o en una simulación Monte Carlo, el sistema recorre un conjunto de estados representativo de la distribución $w_\mu(t)$ subyacente. Si bien esta suposición puede fallar en escenarios especiales, este es el enfoque que toman las simulaciones Monte Carlo para calcular valores de expectación de observables físicos.

Equilibrio

Volviendo a la ecuación maestra (5.7), un estado de equilibrio es aquel donde los pesos o probabilidades $w_\mu(t)$ resultan independientes del tiempo, es decir, uno donde $dw_\mu(t)/dt = 0$.

Puede probarse que, eligiendo las tasas de transición $R(\mu \rightarrow \nu)$ correctamente, las probabilidades de ocupación $w_\mu(t)$ siempre alcanzan el equilibrio a tiempos grandes. Es decir,

$$\lim_{t \rightarrow \infty} w_\mu(t) = p_\mu, \quad (5.10)$$

con p_μ las probabilidades de ocupación en equilibrio, independientes del tiempo. Para sistemas en equilibrio térmico con un reservorio a temperatura T , las probabilidades de ocupación en equilibrio son las de Boltzmann,

$$p_\mu = \frac{1}{\mathcal{Z}} e^{-\beta E_\mu} = \frac{e^{-\beta E_\mu}}{\sum_\nu e^{-\beta E_\nu}}, \quad (5.11)$$

con $\beta = 1/T$ y \mathcal{Z} la función de partición.

Estimador

Formalmente, el valor de expectativa térmico de un observable Q en el ensamble canónico es

$$\langle Q \rangle = \frac{\sum_{\mu} Q_{\mu} e^{-\beta E_{\mu}}}{\sum_{\mu} e^{-\beta E_{\mu}}}. \quad (5.12)$$

Sin embargo, a causa de la cantidad exponencial de estados, la suma no es posible de computar exactamente. Las técnicas Monte Carlo funcionan eligiendo un subconjunto aleatorio de estados de una distribución de probabilidad q_{μ} que uno especifica. Si elegimos M estados μ_1, \dots, μ_M , entonces nuestro estimador es

$$Q_M = \frac{\sum_{i=1}^M Q_{\mu_i} q_{\mu_i}^{-1} e^{-\beta E_{\mu_i}}}{\sum_{i=1}^M q_{\mu_i}^{-1} e^{-\beta E_{\mu_i}}} \quad (5.13)$$

donde $Q_M \rightarrow \langle Q \rangle$ cuando $M \rightarrow \infty$. La obvia pregunta a continuación es quién es q_{μ} . La opción más simple es elegir cualquier estado con igual probabilidad, de manera tal que

$$Q_M = \frac{\sum_{i=1}^M Q_{\mu_i} e^{-\beta E_{\mu_i}}}{\sum_{i=1}^M e^{-\beta E_{\mu_i}}}. \quad (5.14)$$

Sin embargo, la cantidad de estados es absurdamente grande, y un escenario usual es aquel donde la suma del numerador (o ambas) está dominada por un conjunto reducido de estados, mientras que el resto de los estados contribuyen de manera despreciable. Por ejemplo, sabemos que a temperaturas cercanas a cero el sistema estará en su estado fundamental, o en sus primeros estados excitados, y no tiene entonces sentido calcular el estimador a esas temperaturas con cualquier estado. La técnica para elegir solo los estados relevantes para calcular estimadores de entre todos los estados posibles se denomina Muestreo de Importancia (Importance Sampling), y la desarrollamos a continuación.

Muestreo de importancia (Importance Sampling)

Asumiendo que el sistema real que queremos simular ocupa sus estados disponibles según la distribución de Boltzmann, la correcta elección de q_{μ} es precisamente una probabilidad de Boltzmann; $q_{\mu} = \frac{1}{Z} e^{-\beta E_{\mu}}$. Luego, obtenemos para el estimador de $\langle Q \rangle$,

$$Q_M = \frac{1}{M} \sum_{i=1}^M Q_{\mu_i}. \quad (5.15)$$

Procesos de Markov

Para calcular observables es necesario entonces generar un conjunto aleatorio de estados muestreados de una distribución de Boltzmann. Sin embargo, no podemos elegir un estado aleatorio μ y tomarlo con probabilidad proporcional a $e^{-\beta E_{\mu}}$ porque simplemente hay demasiados estados y el procedimiento es totalmente ineficiente. La manera estándar de hacerlo es a través de un proceso de Markov, es decir, un mecanismo que dado un estado μ genera un estado ν de manera no determinista o estocástica, con una dada probabilidad de transición $P(\mu \rightarrow \nu)$.

Un proceso de Markov tiene dos características. (i) Las probabilidades de transición $P(\mu \rightarrow \nu)$ no varían con el tiempo y (ii) dependen exclusivamente de los dos estados μ y ν , y no dependen de ningún otro estado. Además, se satisface la condición

$$\sum_{\nu} P(\mu \rightarrow \nu) = 1, \quad (5.16)$$

ya que el proceso de Markov debe generar a algún estado ν . Nada impide que $\nu = \mu$, es decir, $P(\mu \rightarrow \mu)$ no tiene por qué ser cero.

Una cadena de Markov es un conjunto de estados generados por un proceso de Markov en donde cada estado generado es usado como subsecuente estado inicial del proceso. El proceso de Markov es elegido de tal manera que al correr por un tiempo suficientemente largo, comenzando desde cualquier estado, eventualmente produce una secuencia de estados muestreada con probabilidad de Boltzmann. El proceso de llegar a la distribución de Boltzmann se denomina llegada al equilibrio, ya que es el proceso que realizaría el sistema real mientras llega al equilibrio termodinámico a una dada temperatura.

Como en la distribución de Boltzmann todos los estados del sistema tienen una probabilidad finita, aunque quizás exponencialmente pequeña, de ocurrir, es necesario que la cadena de Markov sea capaz de acceder a todo estado μ del sistema, esperando suficiente tiempo. Esta condición se denomina Ergodicidad.

Balance detallado

A través de la condición denominada de ‘balance detallado’ podemos probar que una cadena de Markov genera estados según la distribución de Boltzmann una vez alcanzado el equilibrio. Primero, escribimos la condición de llegar al equilibrio en términos de las probabilidades de ocupación p_{μ} y las probabilidades de transición $P(\mu \rightarrow \nu)$. En el equilibrio, la tasa con la cual el sistema transiciona desde y hacia un estado cualquiera μ es la misma (ver ecuación (5.7)); es decir,

$$\sum_{\nu} p_{\nu} P(\nu \rightarrow \mu) = \sum_{\nu} p_{\mu} P(\mu \rightarrow \nu) = p_{\mu}. \quad (5.17)$$

Para cualquier conjunto de probabilidades de transición que satisfaga la ecuación anterior, la distribución de probabilidad p_{μ} es de equilibrio.

A continuación, mostraremos cómo asegurar que la distribución de probabilidad de un proceso de Markov converja a p_{μ} para tiempos suficientemente largos.

Las probabilidades de transición $P(\mu \rightarrow \nu)$ pueden pensarse conformando una matriz, denominada Matriz de Markov o Matriz Estocástica del proceso de Markov. Luego, si la probabilidad de que el sistema esté en el estado μ al tiempo t es $w_{\mu}(t)$, y medimos el tiempo en pasos del proceso de Markov, la probabilidad de ocupación del estado ν al tiempo $t + 1$ resulta

$$w_{\nu}(t + 1) = \sum_{\mu} P(\mu \rightarrow \nu) w_{\mu}(t). \quad (5.18)$$

O bien, en notación matricial,

$$\mathbf{w}(t + 1) = \mathbf{P} \cdot \mathbf{w}(t), \quad (5.19)$$

donde las componentes del vector $\mathbf{w}(t)$ son las probabilidades $w_\mu(t)$. Si el proceso de Markov alcanza el estado de equilibrio $w(\infty)$ cuando $t \rightarrow \infty$, ese estado satisface

$$\mathbf{w}(\infty) = \mathbf{P} \cdot \mathbf{w}(\infty). \quad (5.20)$$

Sin embargo, también es posible que el sistema alcance equilibrios dinámicos, en donde

$$\mathbf{w}(\infty) = \mathbf{P}^n \cdot \mathbf{w}(\infty), \quad (5.21)$$

para $n > 1$. Para evitar los equilibrios dinámicos se introduce la condición de balance detallado,

$$p_\mu P(\mu \rightarrow \nu) = p_\nu P(\nu \rightarrow \mu), \quad (5.22)$$

la cual asegura que p_μ es una distribución de equilibrio (pues implica la ecuación (5.17)), y prohíbe los equilibrios dinámicos.

Finalmente, podemos probar que el proceso de Markov tenderá a la distribución p_μ a tiempos grandes sabiendo que cuando $t \rightarrow \infty$, $\mathbf{w}(t)$ tiende exponencialmente hacia el autovector de \mathbf{P} con máximo autovalor, debido a que \mathbf{P} es una matriz estocástica. En efecto, podemos reescribir la ecuación (5.17) como

$$\mathbf{p} = \mathbf{P} \cdot \mathbf{p}. \quad (5.23)$$

Es decir que si se cumple balance detallado (que implica equilibrio), entonces \mathbf{p} es un autovector normalizado de \mathbf{P} , con autovalor 1. Si ahora tenemos en cuenta la ecuación (5.20), tenemos la igualdad

$$\mathbf{w}(\infty) = \mathbf{p}, \quad (5.24)$$

o bien, $\mathbf{w}(t)$ tiende a \mathbf{p} para tiempos grandes⁶. Aquí estamos asumiendo que el autovector de autovalor 1 es único. Puede probarse que la ergodicidad del proceso de Markov garantiza la unicidad de este autovector.

Dadas las condiciones de balance detallado y ergodicidad, podemos entonces hacer que el proceso de Markov converja a una distribución de probabilidad p_μ arbitraria. En nuestro caso de interés, esta distribución es la de Boltzmann. Por lo tanto, la condición de balance detallado para la distribución de Boltzmann se lee

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{p_\nu}{p_\mu} = e^{-\beta(E_\nu - E_\mu)}. \quad (5.25)$$

Tasa de aceptación

Para realizar una simulación Monte Carlo de un sistema que sigue la distribución de Boltzmann hay que definir las probabilidades de transición $P(\mu \rightarrow \nu)$, las cuales tienen que satisfacer (i) ergodicidad, (ii) la ecuación (5.16) y (iii) balance detallado (ecuación (5.25)). No hay una manera canónica de elegir las probabilidades de transición, por lo que a priori en cada problema pueden elegirse diferentes opciones más o menos eficientes numéricamente.

⁶No estamos mostrando aquí el carácter exponencial de la convergencia.

Para empezar, notamos que si tomamos $\nu = \mu$ en la ecuación (5.25) la ecuación se cumple independientemente del valor de $P(\mu \rightarrow \mu)$, por lo que existe una libertad para elegir este valor dentro del rango de valores permitidos, $[0, 1]$, y las restricciones previas mencionadas. Esta libertad es explotada introduciendo la denominada tasa de aceptación, escribiendo a la probabilidad de transición como

$$P(\mu \rightarrow \nu) = g(\mu \rightarrow \nu)A(\mu \rightarrow \nu), \quad (5.26)$$

donde $g(\mu \rightarrow \nu)$ se denomina la probabilidad de selección y $A(\mu \rightarrow \nu)$ la tasa de aceptación. La primera es la probabilidad de que nuestro algoritmo intente pasar desde el estado μ al ν , y la segunda la fracción de veces que nuestro algoritmo tomará esa propuesta de movimiento. Si la propuesta es rechazada, entonces el siguiente paso en la cadena de Markov es nuevamente μ , y si la propuesta es aceptada, el siguiente paso de la cadena es el estado ν . Cuantos más movimientos sean aceptados, más eficiente será el algoritmo en muestrear el espacio de fases del sistema.

Algoritmo de Metropolis

Hay muchas maneras de elegir las probabilidades de selección $g(\mu \rightarrow \nu)$. Aquí hacemos la distinción entre la dinámica del algoritmo y lo que se denomina el algoritmo de Metropolis⁷. La dinámica de *single spin flip* consiste en considerar exclusivamente transiciones donde se propone invertir un único espín por paso. En el algoritmo de metrópolis, las probabilidades de selección $g(\mu \rightarrow \nu)$ para todos los estados considerados por la dinámica se toman iguales, y las probabilidades de selección para el resto de los estados se toman nulas. Si nuestro sistema tiene N espines, entonces, para el algoritmo de Metropolis

$$g(\mu \rightarrow \nu) = 1/N. \quad (5.27)$$

Luego, usando la ecuación anterior, la condición de balance detallado resulta

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = e^{-\beta(E_\nu - E_\mu)}. \quad (5.28)$$

Una forma eficiente de elegir las tasas de aceptación, propuesta por Metropolis y sus colaboradores, es fijar arbitrariamente la mayor de ellas al valor 1 (máximo valor posible), y despejar la menor de la ecuación de balance detallado, es decir

$$A(\mu \rightarrow \nu) = \begin{cases} e^{-\beta(E_\nu - E_\mu)} & \text{if } E_\nu - E_\mu > 0 \\ 1 & \text{en caso contrario.} \end{cases} \quad (5.29)$$

5.3 Maximum likelihood estimation

Mencionamos en el texto general que la salida de una red neuronal que lleva a cabo una clasificación se interpreta como una distribución de probabilidades, usando una función de activación *Softmax*. Interpretamos aquí a la red neuronal como un modelo para la probabilidad

⁷Metropolis no lleva tilde (~~Metrópolis~~) porque es el apellido del físico estadounidense Nicholas Constantine Metropolis.

condicional $P_{\theta}(y|\sigma)$ de asignarle una fase, es decir, una etiqueta $y = 0, 1$, al resultado de una medida σ del sistema. Aquí θ corresponde al conjunto de parámetros de la red neuronal.

Para estimar los parámetros de la red neuronal usamos el principio de máxima verosimilitud, o en inglés, *the maximum likelihood principle*, en donde los parámetros de un modelo estadístico se eligen como aquellos que maximizan la probabilidad de los datos observados en un experimento. Para un conjunto de datos con observaciones $\mathcal{O} = \{\sigma_n, y_n\}_{n=1}^M$, donde $y_n = 0, 1$ etiqueta la fase en la cual la medida σ_n fue realizada, la función de verosimilitud o *likelihood function* de los parámetros θ dadas las observaciones es

$$\mathcal{L}(\theta|\mathcal{O}) \equiv p(\mathcal{O}|\theta) = \prod_{n=1}^M P_{\theta}(y_n|\sigma_n)^{y_n} (1 - P_{\theta}(y_n|\sigma_n))^{1-y_n}, \quad (5.30)$$

donde $\mathbf{y} = (y_1, \dots, y_n)$. En vez de maximizar la función de verosimilitud, es típicamente más conveniente desde un punto de vista práctico minimizar el logaritmo de esta función con el signo cambiado,

$$-\log(p(\mathcal{O}|\theta)) = -\sum_{n=1}^M \left[y_n P_{\theta}(y_n|\sigma_n) + (1 - y_n)(1 - P_{\theta}(y_n|\sigma_n)) \right], \quad (5.31)$$

que es equivalente a la entropía cruzada presentada en (3.6). Dos ejemplos simples con solución analítica sobre *maximum likelihood estimation* pueden encontrarse en este enlace del MIT⁸.

Para cada conjunto de parámetros θ , la función de máxima verosimilitud coincide con una distribución de probabilidad de un conjunto de observaciones. Es inmediato que nuestra $\mathcal{L}(\theta|\mathcal{O})$ es positiva y es fácil ver que está normalizada, ya que

$$\begin{aligned} \sum_{\{\mathbf{y}\}} p(\mathcal{O}|\theta) &= \sum_{y_1, \dots, y_M=0}^1 P_{\theta}(y_1|\sigma_1)^{y_1} (1 - P_{\theta}(y_1|\sigma_1))^{1-y_1} \dots P_{\theta}(y_M|\sigma_M)^{y_M} (1 - P_{\theta}(y_M|\sigma_M))^{1-y_M} \\ &= \left[\sum_{y_1=0}^1 P_{\theta}(y_1|\sigma_1)^{y_1} (1 - P_{\theta}(y_1|\sigma_1))^{1-y_1} \right]^M = 1^M = 1, \end{aligned} \quad (5.32)$$

donde para pasar de la primera a la segunda línea sacamos factores comunes $M - 1$ veces y cada corchete puede evaluarse explícitamente dado que cada uno tiene dos términos.

En general, la función de verosimilitud es conocida, de manera exacta o aproximada, pues puede haber diferencias entre la verosimilitud verdadera y la asumida. Un punto sutil señalado por Jean Barbier en su apunte sobre inferencia bayesiana⁹ es el siguiente. La verosimilitud no debe ser considerada como una distribución de probabilidad. Es una función de los parámetros a inferir. Esto es así porque los datos u observaciones \mathcal{O} están fijos (ya ocurrieron). Decir ‘la verosimilitud de los datos dados los parámetros’ es incorrecto. En cambio, lo correcto es decir ‘la verosimilitud de los parámetros dados los datos’. En la verosimilitud, \mathcal{O} en realidad juega el rol de los parámetros, y θ el de los argumentos, por eso la notación $\mathcal{L}(\theta|\mathcal{O})$.

⁸<https://www.youtube.com/watch?v=00krscK7iBA>

⁹<https://www.sissa.it/sites/default/files/CoursePisa.pdf>

Bibliografía

- [1] Assa Auerbach. Interacting electrons and quantum magnetism. Springer Science & Business Media, 2012.
- [2] Mehran Kardar. Statistical physics of fields. Cambridge University Press, 2007.
- [3] John Cardy. Scaling and renormalization in statistical physics, volume 5. Cambridge university press, 1996.
- [4] Michel Le Bellac et al. Quantum and statistical field theory. Oxford University Press, 1991.
- [5] Thierry Giamarchi. Quantum physics in one dimension, volume 121. Clarendon press, 2003.
- [6] Masaki Oshikawa, Masanori Yamanaka, and Ian Affleck. Magnetization plateaus in spin chains: “haldane gap” for half-integer spins. Phys. Rev. Lett., 78:1984–1987, Mar 1997.
- [7] C. A. Lamas, S. Capponi, and P. Pujol. Combined analytical and numerical approach to study magnetization plateaux in doped quasi-one-dimensional antiferromagnets. Phys. Rev. B, 84:115125, Sep 2011.
- [8] Akihiro Tanaka, Keisuke Totsuka, and Xiao Hu. Geometric phases and the magnetization process in quantum antiferromagnets. Phys. Rev. B, 79:064412, Feb 2009.
- [9] Yingying Tang, Wenbin Guo, Hongping Xiang, Suyun Zhang, Ming Yang, Meiyan Cui, Nannan Wang, and Zhangzhen He. Synthesis, structure, and magnetic properties of $\text{a}_2\text{cu}_5(\text{teo}_3)(\text{so}_4)_3(\text{oh})_4$ ($a = \text{na}, \text{k}$): The first compounds with a 1d kagomé strip lattice. Inorganic Chemistry, 55(2):644–648, 2016. PMID: 26741164.
- [10] S. Acevedo, C. A. Lamas, M. Arlego, and P. Pujol. Magnon crystals and magnetic phases in a kagome-stripe antiferromagnet. Phys. Rev. B, 100:195145, Nov 2019.
- [11] J Schulenburg, Andreas Honecker, Jürgen Schnack, J Richter, and Heinz-Jürgen Schmidt. Macroscopic magnetization jumps due to independent magnons in frustrated quantum spin lattices. Physical review letters, 88:167207, 05 2002.
- [12] Kunj Tandon, Siddhartha Lal, Swapan Pati, S Ramasesha, and Diptiman Sen. Magnetization properties of some quantum spin ladders. Physical Review B, 59, 06 1998.
- [13] Michael Karbach, Gerhard Müller, Harvey Gould, and Jan Tobochnik. Introduction to the bethe ansatz i. Computers in Physics, 11:36, 01 1997.

- [14] Michael Karbach, Kun Hu, and Gerhard Muller. Introduction to the bethe ansatz ii. Computers in Physics, 12, 09 1998.
- [15] F. Elias, M. Arlego, and C. A. Lamas. Magnetization process in a frustrated plaquette dimerized ladder. Phys. Rev. B, 95:214426, Jun 2017.
- [16] I. Dzyaloshinsky. A thermodynamic theory of “weak” ferromagnetism of antiferromagnetics. Journal of Physics and Chemistry of Solids, 4(4):241–255, 1958.
- [17] Tôru Moriya. Anisotropic superexchange interaction and weak ferromagnetism. Phys. Rev., 120:91–98, Oct 1960.
- [18] S. Miyahara, J.-B. Fouet, S. R. Manmana, R. M. Noack, H. Mayaffre, I. Sheikin, C. Berthier, and F. Mila. Uniform and staggered magnetizations induced by dzyaloshinskii-moriya interactions in isolated and coupled spin-1/2 dimers in a magnetic field. Phys. Rev. B, 75:184402, May 2007.
- [19] Zhihao Hao, Yuan Wan, Ioannis Rousochatzakis, Julia Wildeboer, A. Seidel, F. Mila, and O. Tchernyshyov. Destruction of valence-bond order in a $s = \frac{1}{2}$ sawtooth chain with a dzyaloshinskii-moriya term. Phys. Rev. B, 84:094452, Sep 2011.
- [20] Sebastian D Huber and Ehud Altman. Bose condensation in flat bands. Physical Review B, 82(18):184502, 2010.
- [21] Sebastian Greschner, M Piraud, F Heidrich-Meisner, IP McCulloch, U Schollwöck, and T Vekua. Symmetry-broken states in a system of interacting bosons on a two-leg ladder with a uniform abelian gauge field. Physical Review A, 94(6):063628, 2016.
- [22] S. Greschner, M. Piraud, F. Heidrich-Meisner, I. P. McCulloch, U. Schollwöck, and T. Vekua. Spontaneous increase of magnetic flux and chiral-current reversal in bosonic ladders: Swimming against the tide. Phys. Rev. Lett., 115:190402, Nov 2015.
- [23] François Crépin, Nicolas Laflorencie, Guillaume Roux, and Pascal Simon. Phase diagram of hard-core bosons on clean and disordered two-leg ladders: Mott insulator–luttinger liquid–bose glass. Phys. Rev. B, 84:054517, Aug 2011.
- [24] E Orignac, R Citro, M Di Dio, S De Palo, and M-L Chiofalo. Incommensurate phases of a bosonic two-leg ladder under a flux. New Journal of Physics, 18(5):055017, may 2016.
- [25] J.H.P. Colpa. Diagonalization of the quadratic boson hamiltonian. Physica A: Statistical Mechanics and its Applications, 93(3):327 – 353, 1978.
- [26] C.A et al. Lamas. Path integral approach to order by disorder selection in partially polarized quantum spin systems. The European Physical Journal B, 2015.
- [27] E. Orignac and T. Giamarchi. Meissner effect in a bosonic ladder. Phys. Rev. B, 64:144515, Sep 2001.
- [28] G. Cuniberti H.J. Schulz and P. Pieri. Fermi liquids and luttinger liquids. Arxiv, 1998.

- [29] M. Matsuda, M. Azuma, M. Tokunaga, Y. Shimakawa, and N. Kumada. Disordered ground state and magnetic field-induced long-range order in an $s = 3/2$ antiferromagnetic honeycomb lattice compound $\text{bi}_3\text{mn}_4\text{o}_{12}(\text{no}_3)$. Phys. Rev. Lett., 105:187201, Oct 2010.
- [30] M. Matsuda, S. E. Dissanayake, D. L. Abernathy, Y. Qiu, J. R. D. Copley, N. Kumada, and M. Azuma. Frustrated magnetic interactions in an $s = 3/2$ bilayer honeycomb lattice compound $\text{bi}_3\text{mn}_4\text{o}_{12}(\text{no}_3)$. Phys. Rev. B, 100:134430, Oct 2019.
- [31] Mojtaba Alaei, Hamid Mosadeq, Ismaeil Abdolhosseini Sarsari, and Farhad Shahbazi. Origin of magnetic frustration in $\text{bi}_3\text{mn}_4\text{o}_{12}(\text{no}_3)$. Phys. Rev. B, 96:140404, Oct 2017.
- [32] F. A. Gómez Albarracín and H. D. Rosales. Field-induced multiple order-by-disorder state selection in an antiferromagnetic honeycomb bilayer lattice. Phys. Rev. B, 93:144413, Apr 2016.
- [33] Hao Zhang, Carlos A. Lamas, Marcelo Arlego, and Wolfram Brenig. Quantum phase diagram of a frustrated antiferromagnet on the bilayer honeycomb lattice. Phys. Rev. B, 93:235150, Jun 2016.
- [34] Juan Carrasquilla and Roger G Melko. Machine learning phases of matter. Nature Physics, 13(5):431–434, 2017.
- [35] Dian Wu, Lei Wang, and Pan Zhang. Solving statistical mechanics using variational autoregressive networks. Phys. Rev. Lett., 122:080602, Feb 2019.
- [36] Mohamed Hibat-Allah, Martin Ganahl, Lauren E. Hayward, Roger G. Melko, and Juan Carrasquilla. Recurrent neural network wave functions. Phys. Rev. Research, 2:023358, Jun 2020.
- [37] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. nature, 521(7553):436–444, 2015.
- [38] Stefano Curtarolo, Dane Morgan, Kristin Persson, John Rodgers, and Gerbrand Ceder. Predicting crystal structures with data mining of quantum calculations. Phys. Rev. Lett., 91:135503, Sep 2003.
- [39] Christopher C Fischer, Kevin J Tibbetts, Dane Morgan, and Gerbrand Ceder. Predicting crystal structure by merging data mining with quantum mechanics. Nature materials, 5(8):641–646, 2006.
- [40] Giacomo Torlai, Guglielmo Mazzola, Juan Carrasquilla, Matthias Troyer, Roger Melko, and Giuseppe Carleo. Neural-network quantum state tomography. Nature Physics, 14(5):447–450, 2018.
- [41] Li Huang and Lei Wang. Accelerated monte carlo simulations with restricted boltzmann machines. Phys. Rev. B, 95:035105, Jan 2017.
- [42] Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie Vogt-Maranto, and Lenka Zdeborová. Machine learning and the physical sciences. Rev. Mod. Phys., 91:045002, Dec 2019.

- [43] Michael A. Nielsen. Neural Networks and Deep Learning. Determination Press, 2015.
- [44] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, 2016.
- [45] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1):1929–1958, 2014.
- [46] Francois Chollet. Deep Learning with Python. Manning Publications Co., USA, 1st edition, 2017.
- [47] Ansgar Kalz, Andreas Honecker, and Marion Moliner. Analysis of the phase transition for the ising model on the frustrated square lattice. Phys. Rev. B, 84:174407, Nov 2011.
- [48] Ansgar Kalz and Andreas Honecker. Location of the potts-critical end point in the frustrated ising model on the square lattice. Phys. Rev. B, 86:134410, Oct 2012.
- [49] Songbo Jin, Arnab Sen, and Anders W. Sandvik. Ashkin-teller criticality and pseudo-first-order behavior in a frustrated ising model on the square lattice. Phys. Rev. Lett., 108:045702, Jan 2012.
- [50] Songbo Jin, Arnab Sen, Wenan Guo, and Anders W. Sandvik. Phase transitions in the frustrated ising model on the square lattice. Phys. Rev. B, 87:144406, Apr 2013.
- [51] Leon Balents. Spin liquids in frustrated magnets. Nature, 464, March 2010.
- [52] Raymond Marie Ferdinand Houtappel. Order-disorder in hexagonal lattices. Physica, 16(5):425–455, 1950.
- [53] A. Bobák, T. Lučivjanský, M. Žukovič, M. Borovský, and T. Balcerzak. Tricritical behaviour of the frustrated ising antiferromagnet on the honeycomb lattice. Physics Letters A, 380(34):2693–2697, 2016.
- [54] GH Wannier. Antiferromagnetism. the triangular ising net. Physical Review, 79(2):357, 1950.
- [55] Ansgar Kalz, Andreas Honecker, and Marion Moliner. Analysis of the phase transition for the ising model on the frustrated square lattice. Physical Review B, 84(17):174407, 2011.
- [56] Evert PL Van Nieuwenburg, Ye-Hua Liu, and Sebastian D Huber. Learning phase transitions by confusion. Nature Physics, 13(5):435–439, 2017.
- [57] T. Mendes-Santos, X. Turkeshi, M. Dalmonte, and Alex Rodriguez. Unsupervised learning universal critical behavior via the intrinsic dimension. Phys. Rev. X, 11:011040, Feb 2021.
- [58] I. Corte, S. Acevedo, M. Arlego, and C.A. Lamas. Exploring neural network training strategies to determine phase transitions in frustrated magnetic models. Computational Materials Science, 198:110702, 2021.
- [59] Sebastian J Wetzels. Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders. Physical Review E, 96(2):022140, 2017.

- [60] Alexandrou et al. The critical temperature of the 2d-ising model through deep learning autoencoders. The European Physical Journal B, 2020.
- [61] Santiago Acevedo, Carlos A. Lamas, Alejo Costa Duran, Mauricio B. Sturla, and Tomás S. Grigera. On the neural network flow of spin configurations. 2022.
- [62] Dimitrios Giataganas, Ching-Yu Huang, and Feng-Li Lin. Neural network flows of low q-state potts and clock models. New Journal of Physics, 24(4):043040, apr 2022.
- [63] Shotaro Shiba Funai and Dimitrios Giataganas. Thermodynamics and feature extraction by machine learning. Phys. Rev. Research, 2:033415, Sep 2020.
- [64] Mohamed Hibat-Allah, Estelle M. Inack, Roeland Wiersema, Roger G. Melko, and Juan Carrasquilla. Variational neural annealing. 2021.
- [65] Christopher Roth. Iterative retraining of quantum spin models using recurrent neural networks, 2020.
- [66] Steven R. White. Density matrix formulation for quantum renormalization groups. Phys. Rev. Lett., 69:2863–2866, Nov 1992.
- [67] U. Schollwöck. The density-matrix renormalization group. Rev. Mod. Phys., 77:259–315, Apr 2005.
- [68] Steven R. White. Density-matrix algorithms for quantum renormalization groups. Phys. Rev. B, 48:10345–10356, Oct 1993.
- [69] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. Rev. Mod. Phys., 82:277–306, Feb 2010.
- [70] Michael A. Nielsen and Isaac L. Chuang. Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge University Press, 2010.
- [71] F. Verstraete and J. I. Cirac. Matrix product states represent ground states faithfully. Phys. Rev. B, 73:094423, Mar 2006.
- [72] J. Ignacio Cirac, David Pérez-García, Norbert Schuch, and Frank Verstraete. Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. Rev. Mod. Phys., 93:045003, Dec 2021.
- [73] M.E.J. Newman & G.T. Barkema. Monte Carlo Methods in Statistical Physics. Oxford University Press, 1999.