WSTĘP DO TEORII PRZEJŚĆ FAZOWYCH

Andrzej Pękalski

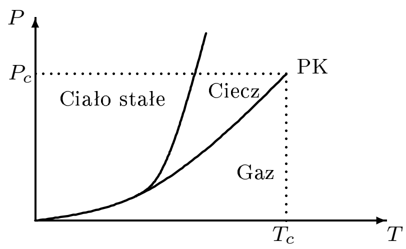
Spis treści

Dziękuję panu Kamilowi Trojanowi za wydatną pomoc w redakcji tego wykładu

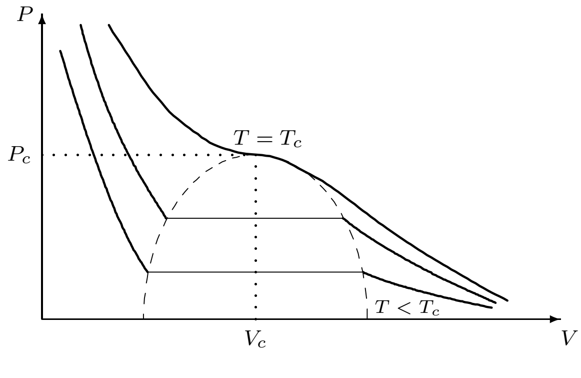
ROZDZIAŁ 1

Wstęp

Powszechnie znane są przykłady zmian fazy w układzie fizycznym, np. wrzenie i krzepnięcie wody, topnienie metali, pojawienie się spontanicznej magnetyzacji itd. Fazy te są na ogół niepodobne do siebie a często przejście między nimi jest nagłe. Możemy jednak czasami tak zmieniać parametr termodynamiczny wywołujący to przejście, że przy pewnej jego wartości różnica między fazami zniknie. Punkt taki nazwiemy **punktem krytycznym** (PK). Jeżeli parametrem tym będzie, jak to jest na ogół, temperatura, wówczas odpowiednią jej wartość nazwiemy **temperaturą krytyczną** (T_c). Powyżej tej temperatury mamy już jednorodną fazę i dalsze zmiany parametru nie przynoszą zmian w układzie. Najlepiej znanym przykładem jest układ ciecz-gaz, dla którego rzut powierzchni PVT (zmienne używane w takim układzie) na płaszczyznę PT ma postać jak na rysunku (1.1), a na płaszczyznę PV jak na rysunku (1.2).



RYSUNEK 1.1. Diagram fazowy w płaszczyźnie PT.



RYSUNEK 1.2. Diagram fazowy w płaszczyźnie PV.

Innymi przykładami punktów krytycznych są:

- magnetyczny punkt krytyczny (temperatura Curie lub Néela), powyżej której znika spontaniczne namagnesowanie ferro lub antyferromagnetyka,
- temperatura krytyczna stopu podwójnego, powyżej której atomy obu składników są wymieszane jednorodnie,
- punkt λ helu 4, poniżej którego istnieją makroskopowe obszary ruchu nadpłynnego,

- punkt krytyczny nadprzewodnika – temperatura poniżej której znika oporność i mogą krążyć nieznikające prądy.

Pierwotnie przejścia fazowe charakteryzowano przy pomocy klasyfikacji Ehrenfesta: jeżeli potencjał termodynamiczny i jego $(n-1)$ pochodnych jest ciągły, natomiast nieciągła jest n-ta pochodna – przejście jest $n-$tego rodzaju. Poznano jednak później przejścia fazowe, które nie poddawały się takiej klasyfikacji, m.in. nie uwzględniała ona faktu, że pewne wielkości mogą być rozbieżne w punkcie krytycznym. Obecnie stosowana jest raczej klasyfikacja podana przez Fishera: Przejścia fazowe są **nieciągłe** jeżeli pochodne potencjału termodynamicznego zmieniają się skokowo (odpowiada to z grubsza PF I rodzaju wg. Ehrenfesta). Np. dla cieczy zmiennymi są P i T, potencjałem potencjał Gibbsa $G = G(P,T)$. Przy przejściu przez krzywą podziału faz poniżej T_c nieciągłe są objętość właściwa $v = (\partial G/\partial P)_T$ i entropia $S = -(\partial G/\partial T)_p$. Dla ferromagnetyka zmiennymi są H i T, energią – energia swobodna Helmholtza $F = F(H,T)$. Poniżej T_c magnetyzacja $M = -(\partial F/\partial H)_T$ zmienia się skokowo gdy pole magnetyczne przechodzi przez zero.

Druga klasa przejść fazowych to przejścia **ciągłe**, których dokładniejszą definicję podamy później, a które teraz możemy scharakteryzować jako takie przejścia w których ciągłe są potencjał i jego pierwsza pochodna, natomiast wyższe pochodne są rozbieżne, bądź zmieniają się skokowo. Przejścia te zachodzą w PK. Cechą charakterystyczną zbliżania się do PK jest wzrost fluktuacji takich wielkości jak energia czy magnetyzacja. Fluktuacje te blisko PK osiągają wielkości makroskopowe, a więc odpowiednie pochodne termodynamiczne (ciepła właściwe, podatność) stają się bardzo duże, lub wręcz dążą do nieskończoności. Dlatego też jednym z centralnych problemów teorii przejść fazowych (PF) będzie badanie asymptotycznych praw rządzących zbliżaniem się do PK. Z ciągłymi przejściami fazowymi związane jest pojęcie **parametru porządku**. Jest to wielkość, która jest różna od zera poniżej PK i może tam przyjmować na ogół kilka wartości, jest natomiast równa zero powyżej PK. Dla układu ciecz-gaz parametrem porządku (PP) będzie różnica gęstości $(\rho_{ciecz} - \rho_{gaz})$, natomiast dla ferromagnetyka PP jest to po prostu spontaniczna magnetyzacja M. Znajdowanie PP nie zawsze jest tak proste i dla pewnych zjawisk (np. nadciekłości) parametr ten nie jest bezpośrednio mierzalny.

Ponieważ w dalszym ciągu przejścia fazowe opisywać będziemy na ogół poprzez układy ciecz-gaz i ferromagnetyk, przypomnimy teraz podstawowe jego własności. Źródłem ferromagnetyzmu jest kwantowe oddziaływanie wymienne wypadkowych momentów magnetycznych (spinów) pojawiających się na niezapełnionych wewnętrznych powłokach elektronowych (d lub f). Przyjmujemy, że momenty orbitalne nie dają istotnego wkładu do tego zjawiska. Oddziaływanie wymienne wynika z sił coulombowskich i zakazu Pauliego i prowadzi do tego, że energia pary elektronów jest mniejsza jeżeli spiny elektronów ustawione są równolegle (dla ferromagnetyka), lub przeciwnie do siebie (antyferromagnetyk). Jeżeli na skutek anizotropii spiny będą ustawiały się wzdłuż pewnej osi, mówimy o ferromagnetyku 1-osiowym, jeżeli preferują pewną płaszczyznę – o ferromagnetyku typu łatwa płaszczyzna. Jeżeli brak takich wyróżnień – magnetyk jest izotropowy. W $T = 0$ układ ma minimum energii i wszystkie spiny skierowane są zgodnie (gdy nie ma zewnętrznego pola $H = 0$), tworząc wypadkowy moment magnetyczny – magnetyzację (dla ferro) lub magnetyzacje podsieci (dla antyferro). Może ona mieć ustalony kierunek (lecz nie zwrot) dla ferromagnetyka 1-osiowego, lub kierunek ten może być dowolny dla ferromagnetyka izotropowego. W miarę podgrzewania układu dostarczamy energię, która zużywana jest na burzenie idealnego porządku – ruch termiczny powoduje oscylacje momentów magnetycznych jonów (spinów w węzłach sieci krystalicznej) wokół położeń równowagi jakimi są położenia dla $T = 0$. Na skutek tych oscylacji zmniejsza się wypadkowy rzut momentu magnetycznego na ów wybrany kierunek – magnetyzacja maleje. Ciągle jednak większość spinów ma ten sam zwrot. Gdy temperatura osiąga T_c, wypadkowy moment (magnetyzacja) równa się zero. Nie oznacza to, że wszystkie spiny są skierowane chaotycznie. Istnieją niewielkie obszary, w których orientacja spinów jest skorelowana, działają bowiem ciągle siły wymiany, jednak wypadkowa magnetyzacja, jako suma wkładów m.in. od takich skorelowanych obszarów, redukuje się. Miarą uporządkowania spinów w takim obszarze jest tzw. **parametr bliskiego porządku**, czyli funkcja korelacyjna spinów niezbyt odległych od siebie. Parametr (dalekiego) porządku, czyli magnetyzacja, znika w PK w sposób ciągły. Dla $T < T_c$ mamy fazę uporządkowaną (PP $\neq 0$), powyżej T_c – nieuporządkowaną (PP $= 0$) o wyższej symetrii. PP łamie symetrię fazy nieuporządkowanej. Powyżej T_c energia termiczna przeważa nad energią wiązania (porządkującą), poniżej T_c jest odwrotnie, natomiast w bezpośrednim otoczeniu T_c energie te są porównywalne. Oznacza to, że w pobliżu T_c po każdym zaburzeniu układ będzie bardzo długo powracał do stanu równowagi. Dlatego też doświadczenia nad procesami równowagowymi w otoczeniu PK, a takimi będziemy się tu zajmowali, są trudne i długotrwałe.

Od teorii PF oczekujemy nie tyle, że opisze ruch wszystkich cząstek tworzących układ, a raczej że powie nam coś o najistotniejszych cechach charakterystycznych naszego układu w otoczeniu PK. Podobnie jak w fizyce statystycznej, której metody będziemy tu stosować, operować będziemy modelami rzeczywistych układów fizycznych.

1. Warunki stabilności

Jak wiadomo, układ fizyczny w fizyce statystycznej opisać możemy poprzez tzw. **równanie podstawowe** łączące parametry ekstensywne (zależące od wielkości układu), np. entropię S i energię wewnętrzną U, objętość V oraz liczbę cząstek N

$$S = S(U, V, N) \qquad \text{lub} \qquad U = U(S, V, N). \tag{1.1}$$

Równoważny każdemu z tych równań jest układ równań dla pochodnych parametrów ekstensywnych (czyli dla parametrów intensywnych), np. temperatury T, ciśnienia P i potencjału chemicznego μ

$$T = T(S, V, N), \qquad P = P(S, V, N) \quad \text{i} \quad \mu_i = \mu_i(S, V, N). \tag{1.2}$$

Położenie równowagi dane jest przez maksimum S lub minimum U. Warunki konieczne równowagi dwu podukładów prowadzą do konieczności równości temperatur i ciśnień w podukładach, natomiast warunki dostateczne – do nieujemności współczynnika ściśliwości izotermicznej oraz ciepła właściwego. To z kolei wymusza pewne zachowanie potencjałów termodynamicznych, np. potencjał Gibbsa $G = G(T, P)$ musi być wypukłą funkcją obu argumentów, natomiast potencjał Helmholtza, $F = F(T, V)$ musi być wypukłą funkcją T i wklęsłą funkcją V. Jeżeli przy zmianie parametru potencjał przestaje spełniać te warunki, wówczas układ przechodzi do innej fazy, w której warunki te są znów spełnione. Przyjmijmy, że zewnętrznym parametrem który zmieniamy jest temperatura, natomiast utrzymujemy stałą objętość układu i stałą liczbę cząstek. Wówczas ponieważ $S = S(U, V)$, więc

$$dU = \left(\frac{\partial S}{\partial U}\right)_V,$$
$$dS = \left(\frac{-PdV + đQ}{T}\right) = \frac{đQ}{T}.$$

Warunek dodatniości ciepła właściwego ma postać

$$C_v = \left(\frac{\partial T}{\partial S}\right)_V > 0. \tag{1.3}$$

Dzieląc to przez $T(> 0)$ dostajemy

$$\frac{1}{T}\left(\frac{\partial T}{\partial S}\right)_V = \left(\frac{dT}{đQ}\right) > 0. \tag{1.4}$$

Ten warunek oznacza po prostu, że dodanie ciepła do układu musi podwyższyć jego temperaturę, jeżeli układ ma pozostać stabilny. W przeciwnym wypadku zostaje ono zużyte na przemianę fazową, np. na topnienie, i układ dzieli się na różne fazy.

Równoważnym warunkiem jest niedodatniość ściśliwości izotermicznej (lub nieujemność podatności magnetycznej)

$$\kappa_T < 0, \qquad \text{lub} \qquad \chi_T = \left(\frac{\partial M}{\partial H}\right)_V > 0, \tag{1.5}$$

co oznacza, że zwiększenie zewnętrznego pola magnetycznego musi prowadzić do wzrostu namagnesowania.

2. Wykładniki krytyczne

Wiemy już jakie muszą być spełnione warunki by układ pozostawał stabilny. Warto zbadać jak, przy dążeniu parametru zewnętrznego do wartości krytycznej, wielkości fizyczne, takie jak magnetyzacja czy podatność, dążą do zera bądź nieskończoności. Uzasadnienie celowości takich badań wkrótce poznamy.

Jeżeli nieujemna funkcja $f(x)$ zachowuje się, dla x $\downarrow 0$, jak x^λ, tzn.

$$f(x) \propto x^\lambda \quad \text{dla} \quad x \downarrow 0 \tag{1.6}$$

lub dokładniej, gdy

$$\lim_{x \downarrow 0} \left[\frac{\ln f(x)}{\ln x} \right] = \lambda, \tag{1.7}$$

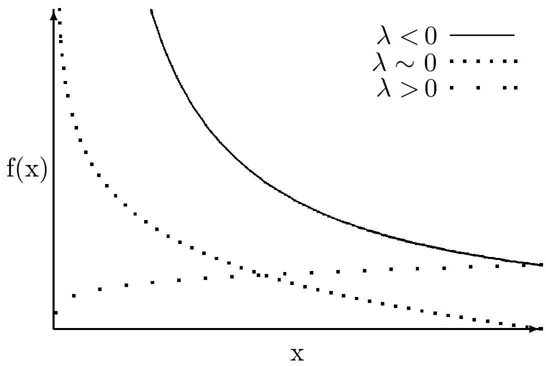
wówczas wielkość λ nazywamy **wykładnikiem (krytycznym)**. Nie oznacza to prostej proporcjonalności, lecz możemy spodziewać się, że na ogół

$$f(x) = Ax^\lambda \left[1 + ax \ldots\right] \qquad \text{dla} \quad x \downarrow 0, \tag{1.8}$$

gdzie A jest amplitudą osobliwości (w sensie fizycznym, a więc może to odpowiadać np. dążeniu do zera), natomiast a jest amplitudą wiodącej poprawki. Jeżeli z wyników doświadczalnych znana jest $f(x)$, to wykładnik krytyczny wielkości reprezentowanej przez $f(x)$ możemy znaleźć obliczając nachylenie krzywej doświadczalnej, łącząc bowiem definicję wykładnika z regułą de l'Hospitala dostajemy

$$\lambda = \frac{d \ln f(x)}{d \ln x}. \tag{1.9}$$

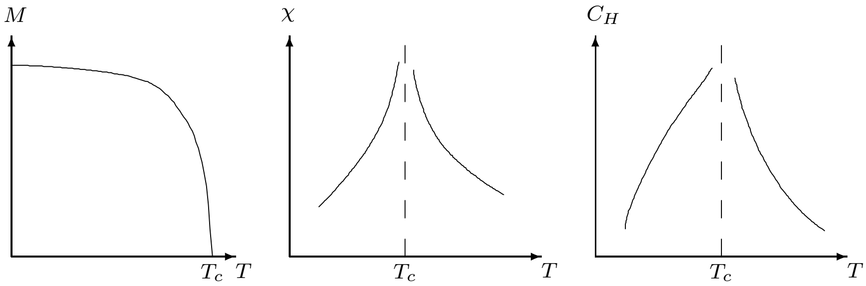
W zależności od znaku λ mamy trzy możliwości



Rysunek 1.3. Zachowanie funkcji $f(x) = x^\lambda$ w zależności od wartości λ. Środkowa wartość odpowiada zależności $f(x) \sim\mid \ln(x) \mid$.

Zwykle w teorii przejść fazowych zmienną jest odstępstwo rozpatrywanego parametru termodynamicznego od wartości krytycznej. Na ogół parametrem tym jest temperatura T, a wtedy zmienna x określona jest jako

$$x \equiv \varepsilon = \frac{T - T_c}{T_c}. \tag{1.10}$$



RYSUNEK 1.4. Schematyczne zależności namagnesowania, podatności i ciepła właściwego od temperatury.

2.1. Przykłady wykładników krytycznych. Podamy je dla układów ciecz-gaz i dla ferromagnetyka. Otrzymane doświadczalnie wykresy interesujących nas wielkości termodynamicznych mają postać (dane dla Ni) jak na rysunku 1.4.

(1) **Parametr porządku**
Dla układu ciecz-gaz

$$\frac{\rho_{ciecz}(T) - \rho_{gaz}(T)}{2\rho_c} \sim (-\varepsilon)^\beta \tag{1.11}$$

i dla magnetyka

$$\frac{M(T)}{M(0)} \sim (-\varepsilon)^\beta \tag{1.12}$$

Współczynniki ρ_c i M(0) wprowadza się aby prawa strona nie zależała od stałych materiałowych. Z danych doświadczalnych dla układów 3D wynika, że $\beta \approx \frac{1}{3}$, tzn. między 0.2 i 0.4.

(2) **Reakcja układu na zewnętrzne zaburzenie**
Ściśliwość izotermiczna (κ_T) i podatność magnetyczna (χ_0) są miarą reakcji układu na zewnętrzne zaburzenie. Wielkości te oblicza się z nachylenia izoterm P, V lub H, M. W odróżnieniu od parametru porządku, teraz możemy podejść do PK zarówno od strony wyższych jak i niższych temperatur, przy czym wykładniki nie muszą być sobie równe, jednakże we wszystkich znanych przypadkach tak jest. Obie wielkości są rozbieżne gdy $T \to T_c$, możemy więc napisać

$$\kappa_T = C(-\varepsilon)^{-\gamma}(1 + \cdots) \quad \text{dla} \qquad T > T_c. \tag{1.13}$$

Analogiczny wzór jest słuszny dla $T < T_c$, z tym, że zamiast γ mamy γ'.

$$\kappa_T \sim (-\varepsilon)^{-\gamma'}. \tag{1.14}$$

$$\chi_0 \sim (-\varepsilon)^{-\gamma'}. \tag{1.15}$$

Aby mieć dodatnie wykładniki wprowadziliśmy -γ. Typowe ich wartości, to 1.15 – 1.35 dla materiałów 3D.

(3) **Izoterma krytyczna**
Jest to zależność $(P - P_c)$ od różnicy gęstości (dla układu ciecz-gaz) lub H od M dla magnetyka

$$(P - P_c) \sim \varepsilon^\delta sign(\rho - \rho_c) \tag{1.16}$$

zmienną jest tutaj

$$\varepsilon = \frac{\rho_c - \rho}{\rho_c} \tag{1.17}$$

dla układu ciecz-gaz. Dla ferromagnetyka

$$H = |M|^\delta. \tag{1.18}$$

W obu przypadkach $T = T_c$. Dane doświadczalne dają $\delta = 4.2 - 4.6$. Im większa jest wartość δ, tym bardziej płaska jest izoterma krytyczna i tym trudniej jest doświadczalnie wyznaczyć δ.

(4) **Ciepło właściwe**
Może być mierzone przy stałej objętości (dla układu ciecz-gaz)

$$C_V \sim (-\varepsilon)^{-\alpha}, \tag{1.19}$$

gdzie tym razem

$$\varepsilon = \frac{T - T_c}{T_c}. \tag{1.20}$$

i dla ferromagnetyka (przy zewnętrznym polu równym zeru)

$$C_H \sim (-\varepsilon)^{-\alpha}. \tag{1.21}$$

Dane doświadczalne to $0 - 0.2$, co świadczyłoby o rozbieżności logarytmicznej.

(5) **Funkcja korelacyjna**
Funkcja ta związana jest z podatnością. Rozpatrzmy to na przykładzie ferromagnetyka. Dwupunktowa funkcja korelacyjna fluktuacji spinów zdefiniowana jest jako

$$G(r, r\prime) = \langle(\sigma(r) - \langle\sigma\rangle)(\sigma(r\prime) - \langle\sigma\prime\rangle)\rangle \tag{1.22}$$

Dla układów jednorodnych

$$\langle\sigma\prime\rangle = \langle\sigma\rangle, \tag{1.23}$$

gdzie $\sigma(r)$ jest wartością zredukowanej magnetyzacji w pewnym małym obszarze wokół punktu r. Jeżeli układ jest translacyjnie niezmienniczy, wówczas dwuspinowa funkcja korelacyjna zależy tylko od jednego parametru – różnicy wektorów $\vec{r\prime}$ i $\vec{r}$. Zerowa, tzn. dla $k = 0$, transformata Fouriera funkcji korelacyjnej ma, na podstawie twierdzenia fluktuacyjno - dysypacyjnego, postać

$$G(0) = \int d^3 r G(r) = \chi_T \tag{1.24}$$

jest więc proporcjonalna do podatności. Sama funkcja korelacyjna maleje (także dla $T = T_c$) do zera wraz ze wzrostem odległości r między spinami, jak

$$G(r) \sim \frac{1}{r^{d+2-\eta}} \tag{1.25}$$

d jest tu wymiarem przestrzeni, η jest wykładnikiem krytycznym, a jego wartość doświadczalna to około 0.2.

Przedstawione tu zależności wielkości fizycznych, takich jak ciepło właściwe, magnetyzacja, ściśliwość itd., mają ogólnie postać

$$f = \frac{|T - T_c|^\alpha}{T_c}, \qquad \text{czyli} \qquad f = x^\alpha = \left(\frac{x}{x_0}\right)^\alpha \tag{1.26}$$

Dzielimy przez stałą x_0, aby pozbyć się zależności od stałych materiałowych – różne materiały będą miały w tym przypadku różne wartości T_c. Tego typu zależności nazywamy *prawami potęgowymi*. Mają one pewną szczególną i bardzo ważną własność – nie zależą od skali, czyli rozmiarów układu. Weźmy bowiem dwie zależności

$$f_1 = x^\alpha, \qquad f_2 = e^x, \tag{1.27}$$

a następnie przeskalujmy zmienną niezależną x, tzn. weźmy np. $x' = 10x$. dostaniemy

$$f_1' = 10^\alpha x^\alpha = C f_1, \qquad f_2' = (e^x)^{10}, \tag{1.28}$$

czyli f_1' to f_1 pomnożone przez stałą, natomiast f_2' jest już inną funkcją. Z funkcji f_1' do f_1 możemy powrócić zmieniając po prostu skalę wykresu. Możemy więc w takim przypadku badać zachowanie układu w mikroskali, a następnie ekstrapolować je do makroskali. Oczywiście nie wszystkie zjawiska fizyczne i nie w każdych warunkach są niezależne od skali – jest to raczej wyjątek. Stąd też fakt, że układ (pewna jego cecha) zachowuje się zgodnie z prawem potęgowym oznacza, że spełnia on pewne szczególne warunki – być może znajduje się w stanie krytycznym, zachowanie krytyczne jest bowiem opisywane przez prawa potęgowe. Mała zmiana parametru kontrolnego (np. temperatury) prowadzi do zasadniczych zmian zachowania układu jako całości.

Nie są to wszystkie wykładniki krytyczne, lecz te najistotniejsze dla opisu statycznych przejść fazowych.

Okazuje się, że wartości wykładników nie są specyficzne dla jednego układu, lecz przyjmują wartości zbliżone dla czasem nawet bardzo różnych od siebie układów. Mają więc one charakter uniwersalny. Czynniki jakie mogą wpłynąć na ich wartość to symetria (np. 1-osiowa czy układ izotropowy), wymiar przestrzenny próbki oraz wymiar parametru porządku. Dodatkowym czynnikiem jest to czy siły działające mają skończony zasięg działania. Możemy więc istniejące modele (realne układy) podzielić na pewne klasy (*klasy uniwersalności*) w zależności od wartości wykładników krytycznych. Układy należące do tej samej klasy uniwersalności mogą mieć bardzo różne pochodzenie sił działających między cząstkami układu, a więc bardzo różne wartości T_c, H_c itd., lecz te same wykładniki krytyczne. Dlatego też wyznaczenie dla danego układu jego wykładników krytycznych pozwala na zaklasyfikowanie tego układu do określonej klasy uniwersalności. Czynnikami decydującymi o przynależności modelu do danej klasy uniwersalności są: wymiar przestrzenny D oraz wymiar parametru porządku d. Tak więc dwu-wymiarowy model Isinga, dla którego istnieje skończona temperatura krytyczna, należy do innej klasy uniwersalności niż jedno-wymiarowy model Isinga, dla którego $T_c = 0$. Podobnie dwu-wymiarowy model Heisenberga, dla którego wymiar parametru porządku to 3, a $T_c = 0$, należy do innej klasy uniwersalności niż dwu-wymiarowy model Isinga z parametrem porządku o wymiarze 1.

✠

Jakobian funkcji $u, v, \ldots, w$ zmiennych $x, y, \ldots, z$ zapisany jest jako

$$\frac{\partial(u, v, \ldots, w)}{\partial(x, y, \ldots, z)} \tag{1.29}$$

Stąd

$$\left(\frac{\partial u}{\partial x}\right)_{y,\ldots,z} = \frac{\partial(u, y, \ldots, z)}{\partial(x, y, \ldots, z)} \tag{1.30}$$

$$\frac{\partial(u, v, \ldots, w)}{\partial(x, y, \ldots, z)} = -\frac{\partial(v, u, \ldots, w)}{\partial(x, y, \ldots, z)}, \tag{1.31}$$

$$\frac{\partial(v, u, \ldots, w)}{\partial(x, y, \ldots, z)} = \frac{\partial(v, u, \ldots, w)}{\partial(r, s, \ldots, t)} \frac{\partial(r, s, \ldots, t)}{\partial(x, y, \ldots, z)} \tag{1.32}$$

✠

Pokażemy teraz, że związki między wykładnikami można otrzymać z równości termodynamicznych. Jednym z najbardziej znanych takich związków jest **nierówność Rushbrooke'a**:

$$\alpha + 2\beta + \gamma \geq 2, \tag{1.33}$$

gdzie α i γ odnoszą się do podejścia od dołu do T_c.

Mamy

$$\begin{aligned}\left(\frac{\partial S}{\partial T}\right)_M &= \frac{\partial(S, M)}{\partial(T, M)} = \frac{\partial(S, M)}{\partial(T, H)} \frac{\partial(T, H)}{\partial(T, M)} \\ &= \left[\left(\frac{\partial S}{\partial T}\right)_H \left(\frac{\partial M}{\partial H}\right)_T - \left(\frac{\partial S}{\partial H}\right)_T \left(\frac{\partial M}{\partial T}\right)_H\right] \cdot \left(\frac{\partial H}{\partial M}\right)_T \\ &= \left(\frac{\partial S}{\partial T}\right)_H - \left(\frac{\partial S}{\partial M}\right)_T \left(\frac{\partial M}{\partial T}\right)_H\end{aligned}$$

Wykorzystaliśmy tu fakt, że jakobian jest wyznacznikiem i obliczyliśmy jego wartość. Podobnie

$$\left(\frac{\partial S}{\partial M}\right)_T\left(\frac{\partial M}{\partial T}\right)_H \equiv \frac{\partial(S,T)}{\partial(M,T)}\frac{\partial(M,H)}{\partial(T,H)} = -\frac{\partial(S,T)}{\partial(T,M)}\frac{\partial(H,M)}{\partial(H,T)} = -\left(\frac{\partial S}{\partial H}\right)_T\left(\frac{\partial H}{\partial T}\right)_M \quad (1.34)$$

a więc

$$\left(\frac{\partial S}{\partial T}\right)_M = \left(\frac{\partial S}{\partial T}\right)_H + \left(\frac{\partial S}{\partial H}\right)_T\left(\frac{\partial H}{\partial T}\right)_M. \quad (1.35)$$

Samo ciepło właściwe to

$$C_{H,M} = T\left(\frac{\partial S}{\partial T}\right)_{H,M}. \quad (1.36)$$

Z definicji $C_{H,M}$ wynika, że

$$C_M = C_H + T\left(\frac{\partial S}{\partial H}\right)_T\left(\frac{\partial H}{\partial T}\right)_M. \quad (1.37)$$

Drugi wyraz da się przekształcić przy pomocy związków Maxwella, które możemy zapisać tu jako

$$\left(\frac{\partial S}{\partial H}\right)_T = -\left(\frac{\partial^2 S}{\partial H\partial T}\right)_H = \left(\frac{\partial M}{\partial T}\right)_H \quad (1.38)$$

oraz

$$\left(\frac{\partial H}{\partial T}\right)_M\left(\frac{\partial T}{\partial M}\right)_H\left(\frac{\partial M}{\partial H}\right)_T = -1, \quad (1.39)$$

$$T\left(\frac{\partial M}{\partial T}\right)_H\left(\frac{\partial H}{\partial T}\right)_M = -T\left(\frac{\partial M}{\partial T}\right)_H / \left(\frac{\partial T}{\partial M}\right)_H\left(\frac{\partial M}{\partial H}\right)_T = -T\frac{\left(\frac{\partial M}{\partial T}\right)_H^2}{\left(\frac{\partial M}{\partial H}\right)_T} \quad (1.40)$$

Ponieważ wiemy, że $C_M \geq 0$, więc

$$C_H \geq T\frac{\left(\frac{\partial M}{\partial T}\right)_H^2}{\left(\frac{\partial M}{\partial H}\right)_T}. \quad (1.41)$$

W granicy $H \to 0$, $M = M_0$ jest spontaniczną magnetyzacją, natomiast

$$\left(\frac{\partial M}{\partial H}\right)_T = \chi_0. \quad (1.42)$$

podatnością początkową. Mamy więc

$$C_{H=0} \geq T\frac{\left(\frac{\partial M_0}{\partial T}\right)^2}{\chi_0}. \quad (1.43)$$

Weźmy teraz logarytm z obu stron

$$\ln C_{H=0} \geq 2\ln\left(\left|\frac{\partial M_0}{\partial T}\right|\right) - \ln\chi_0 + \ln T \quad (1.44)$$

podzielmy przez $\ln|T_c - T|$ i obliczymy granicę $T \uparrow T_c$. Wówczas pamiętając o definicji wykładników krytycznych,

$$C_H \sim \varepsilon^{-\alpha}, \quad \chi_T \sim \varepsilon^{-\gamma}, \quad \left(\frac{\partial M}{\partial T}\right)_H \sim \epsilon^{\beta-1} \quad (1.45)$$

dostajemy

$$-\alpha \leq 2(\beta - 1) + \gamma. \quad (1.46)$$

Jest to nierówność Rushbrooke'a. Na podstawie różnych danych uważa się, że istnieją tylko dwa niezależne wykładniki, a nierówność Rushbrooke'a jest naprawdę równością. Pozostałe wykładniki można otrzymać poprzez ścisłe związki typu równości Rushbrooke'a.

Związki te, jak widzieliśmy powyżej, wynikają wprost z równowagowej fizyki statystycznej i termodynamiki.

ROZDZIAŁ 2

Przybliźone teorie

1. Klasyczna teoria van der Waalsa

Dotyczy ona przejść typu ciecz-gaz. W swej rozprawie doktorskiej w 1873 r. van der Waals podał następujące równanie stanu dla gazu rzeczywistego, tzn takiego w którym cząstki zajmują pewna objętość (nie są punktowe) i oddziałują pomiędzy sobą

$$\beta P = \frac{\rho}{1 - \rho V_0} - \frac{\beta \alpha \rho^2}{2}, \tag{2.1}$$

gdzie V_0 jest objętością cząstki ($= \frac{4}{3}\pi r_0^3$ w 3D), $\rho = V^{-1}$, $\beta = (kT)^{-1}$ natomiast α wyrażone jest przez potencjał przyciągający między cząstkami cieczy $\phi(r)$

$$\alpha = -\int \phi(r) dr. \tag{2.2}$$

Oprócz potencjału przyciągającego istnieje też w teorii potencjał „twardego rdzenia", oznaczający po prostu, że cząstki nie mogą się nakrywać, lecz pozatem nic o sobie nie wiedzą

$$\varphi(r) = \begin{cases} \infty & : \quad r < r_0 \\ 0 & : \quad r > r_0. \end{cases} \tag{2.3}$$

Wyjściowe równanie często zapisuje się w postaci

$$(V - V_0)\left(P + \frac{\alpha}{2V^2}\right) = kT. \tag{2.4}$$

Dla ustalonych P,T równanie to jest równaniem kubicznym dla V i stąd ma trzy pierwiastki V_1, V_2, V_3. W PK zmienne przyjmują wartości krytyczne p_c, T_c, V_c, a więc

$$(V - V_c)^3 = 0. \tag{2.5}$$

Jednocześnie równanie van der Waalsa ma w tym punkcie postać (jako równanie dla V, przy ustalonych, krytycznych, P_c, T_c)

$$(V - V_0)\left(P_c + \frac{\alpha}{2V^2}\right) = kT_c \tag{2.6}$$

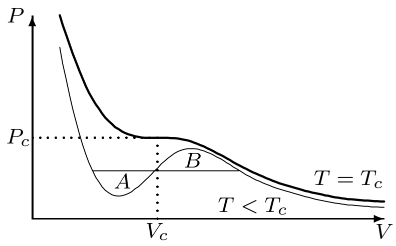
czyli

$$V^3 - \left(V_0 + \frac{kT_c}{P_c}\right)V^2 + \frac{\alpha}{2P_c}V - \frac{\alpha V_0}{2P_c} = 0 \tag{2.7}$$

i musi ono być równoważne równaniu krytycznemu. Porównując współczynniki dostajemy

$$kT_c = \frac{4\alpha}{27V_0}, \qquad P_c = \frac{\alpha}{54V_0^2}, \qquad V_c = 3V_0. \tag{2.8}$$

Wyznaczyliśmy w ten sposób wartości krytyczne poprzez stałe materiałowe α, V_0. Izotermy równania van der Waalsa maja postać jak na rys. 2.1. Dla $T < T_c$ izotermy są niezgodne z warunkiem stabilności $(\partial P/\partial V)_T \leq 0$. Aby pozbyć się tego paradoksu, Maxwell w 1874 r. zaproponował rozwiązanie znane obecnie jako **konstrukcja Maxwella**, polegające na podziale izotermy tak, aby powierzchnia obszaru A była równa powierzchni obszaru B. Izotermy van der Waalsa przeniesione na wykres energii swobodnej ($kTP = \partial F/\partial V$) oznaczają, że F nie jest wypukłą funkcją V. Konstrukcja Maxwella wypełnia to zagłębienie. Jak dotąd brak pełnego uzasadnienia konstrukcji Maxwella, wiadomo tylko, że daje ona dobre wyniki. Teoria van der Waalsa jest poprawna daleko



RYSUNEK 2.1. Izotermy van der Waalsa i konstrukcja Maxwella.

od PK. Aby lepiej zrozumieć przyjęte w tej teorii założenia, spróbujmy wyprowadzić równanie van der Waalsa wychodząc z założonego potencjału (oprzemy się na podejściu Ornsteina z 1908 r.). Suma stanów jest równa

$$Z(V,N,T) = (N!)^{-1} \int \cdots \int_V d^3r_1 \cdots d^3r_N \prod_{1\le i<j\le N} S(|\vec{r_i} - \vec{r_j}|) \exp\Big[-\beta \sum_{1\le i<j\le N} \phi(|\vec{r_i} - \vec{r_j}|)\Big]$$

gdzie

$$S(r) = \begin{cases} 0 & : & r < r_0 \\ 1 & : & r > r_0 \end{cases} \tag{2.9}$$

i związane jest z potencjałem „twardego rdzenia". Załóżmy teraz, że oddziaływanie przyciągające jest bardzo daleko zasięgowe. Wobec tego energia potencjalna dla prawie wszystkich konfiguracji N cząstek będzie równa wartości średniej, tzn.

$$\sum_{1\le i>j\le N} \phi(|\vec{r_i} - \vec{r_j}|) \approx -\frac{\alpha N^2}{2V} = -\frac{\alpha N\rho}{2}. \tag{2.10}$$

Jedna cząstka oddziałuje ze wszystkimi pozostałymi, czyli z gęstością ϱ, z siłą daną przez α. Takich cząstek jest N. To trzeba jeszcze podzielić przez 2, aby nie liczyć dwa razy tych samych oddziaływań. Mamy więc

$$Z(V,N,T) \approx \exp\left(\frac{\alpha N\rho}{2kT}\right)\left[(N!)^{-1}\int \cdots \int_V d^3r_1 \cdots d^3r_N \prod S(|\vec{r_i} - \vec{r_j}|)\right]. \tag{2.11}$$

Wyraz w $[\cdots]$ jest sumą stanów dla układu twardych kul o promieniu r_0. Taką sumę stanów umiemy policzyć tylko w 1D, gdzie

$$Z^{hc}(L,N) = \frac{[L-(N-1)V_0]^N}{N!}. \tag{2.12}$$

L jest tu 1D objętością, a więc długością układu, zaś V_0 jest 1D objętością cząstki. Jeżeli przyjąć, że wyrażenie to jest rozsądnym przybliżeniem w 3D, to energia swobodna na cząstkę w granicy termodynamicznej ma postać

$$\begin{aligned} -\frac{f}{kT} &= \lim_{\substack{L,N\to\infty \\ \rho=const}} N^{-1} \ln Z(L,N,T) \\ &= \frac{\alpha\rho}{2kT} + (N)^{-1} \lim_{\substack{L,N\to\infty \\ \rho=const}} \ln \frac{[L-(N-1)V_0]^N}{N!} \\ &= \frac{\alpha\rho}{2kT} + \lim_{\substack{L,N\to\infty \\ \rho=const}} N^{-1} \ln Z^{hc} = 1 + \frac{\alpha\rho}{2kT} + \ln(V - V_0). \end{aligned}$$

Mamy bowiem

$$
\begin{aligned}
N^{-1}\ln Z^{hc} &= N^{-1}\ln\left\{\frac{N^N}{N!}\left[\frac{L}{N}-V_0+\frac{V_0}{N}\right]^N\right\} \\
&= N^{-1}\ln\frac{N^N}{N!}+\ln\left(\frac{L}{N}-V_0+\frac{V_0}{N}\right).
\end{aligned}
$$

Biorąc granice $L\to\infty$,$N\to\infty$, tak aby $\frac{L}{N}=V$, (czyli granicę termodynamiczną), otrzymamy, korzystając ze wzoru Stirlinga,

$$
\lim_{L,N\to\infty} N^{-1}\ln Z^{hc} =
$$

$$
=\lim_{L,N\to\infty} N^{-1}\ln\frac{N^N}{N!}+\lim_{L,N\to\infty}\ln\left(\frac{L}{N}-V_0+\frac{V_0}{N}\right) \quad = 1+\ln(V-V_0).
$$

Mając już wyrażenie dla energii swobodnej, możemy stąd wyznaczyć ciśnienie

$$
P=-\left(\frac{\partial f}{\partial V}\right)_T = kT\left[(V-V_0)^{-1}-\frac{\alpha\rho^2}{2kT}\right] \tag{2.13}
$$

a stąd otrzymujemy już równanie stanu, czyli równanie van der Waalsa

$$
\beta P=\frac{1}{V-V_0}-\frac{\beta\alpha\rho^2}{2}. \tag{2.14}
$$

Wykorzystując wyrażenie dla Z^{hc} możemy to zapisać

$$
\frac{P}{kT}=\frac{P^{hc}}{kT}-\frac{\alpha\rho^2}{2kT} \tag{2.15}
$$

Uwzględniliśmy to, że P^{hc} (ciśnienie twardych sfer) jest w 1D równe $kT/(V-V_0)$.

Otrzymaliśmy więc równanie van der Waalsa. Jest ono tylko przybliżone, ponieważ założenie, że energia potencjalna jest równa wartości średniej nie jest słuszne. Jest tak kiedy siły oddziaływania mają nieskończony promień działania. Do układu nieskończonego przechodzimy jednak dopiero w granicy termodynamicznej i dopiero **po** tym przejściu możemy przyjąć nieskończony zasięg oddziaływań. Jeden z postulatów mechaniki statystycznej mówi, że potencjał oddziaływania nie może zależeć bezpośrednio od wielkości układu. Lebowitz i Penrose pokazali ściśle, że jeżeli zachowa się poprawną sekwencję granic, wówczas można otrzymać równanie van der Waalsa już z konstrukcją Maxwella. Wynik ten można zinterpretować tak, że teorie klasyczne, których przykładem jest teoria van der Waalsa, są słuszne w granicy nieskończenie daleko-zasięgowego i nieskończenie słabego oddziaływania.

Przedstawione tutaj wyprowadzenie równania van der Waalsa pokazuje również filozofię traktowania problemów w fizyce teoretycznej. Punktem wyjścia jest pewien model (już uproszczony) rzeczywistej sytuacji fizycznej. Kiedy jednak nie udaje się znaleźć rozwiązania analitycznego, dokonujemy przybliżenia, biorąc najbliższą sytuację fizyczną, dla której rozwiązanie takie znamy - w tym przypadku było to równanie stanu dla 1D gazu twardych sfer. Po uzyskaniu ostatecznego wyniku należy sprawdzić jak przyjęte uproszczenie ogranicza stosowalność naszego podejścia.

Zobaczmy teraz jakie wartości wykładników krytycznych daje teoria van der Waalsa. W tym celu znajdziemy najpierw tzw równanie stanów odpowiednich Łącząc wartości krytyczne T_c, P_c, V_c otrzymamy

$$
\frac{P_cV_c}{kT_c}=\frac{3}{8}=0.375. \tag{2.16}
$$

Widać więc, że taka kombinacja wartości krytycznych parametrów nie zależy w ogóle od stałych materiałowych, tzn. wielkości charakteryzujących badany układ (V_0, α). Pomnóżmy równanie van der Waalsa przez $54V_0/\alpha$

$$
(V-V_0)\frac{54V_0}{\alpha}\left(P+\frac{\alpha}{2V^2}\right)=\frac{54V_0kT}{\alpha}=\frac{27V_0}{4\alpha}8kT=8\tilde{T} \tag{2.17}
$$

$$(V - V_0)\frac{54V_0}{\alpha}P + (V - V_0)\frac{27V_0}{V^2} = \left(\frac{V - V_0}{V_0}\right)\left(\tilde{P} + \frac{3}{\tilde{V}^2}\right) = (3\tilde{V} - 1)\left(\tilde{P} + \frac{3}{\tilde{V}^2}\right). \quad (2.18)$$

Mamy więc

$$\left(3\tilde{V} - 1\right)\left(\tilde{P} + \frac{3}{\tilde{V}^2}\right) = 8\tilde{T} \quad (2.19)$$

Wielkości z tildą $(\tilde{P}, \tilde{T}, \tilde{V})$ są odniesione do odpowiednich wartości krytycznych. Widać więc, że jeżeli P, V, T mierzyć w jednostkach „krytycznych", wówczas dla wszystkich substancji otrzymamy to samo równanie stanu,niezależnie od wartości stałych materiałowych. Stąd też jego nazwa – **równanie stanów odpowiednich**.

Dla znalezienia wykładników krytycznych wprowadźmy wielkości

$$p = \tilde{P} - 1 = \frac{P - P_c}{P_c}, \qquad v = \tilde{V} - 1 = \frac{V - V_c}{V_c}, \qquad \epsilon = \tilde{T} - 1 = \frac{T - T_c}{T_c}. \quad (2.20)$$

Podstawiając to do równania stanów odpowiednich dostajemy

$$\left[(1 + p) + 3(1 + v)^{-2}\right][3(1 + v) - 1] = 8(1 + \epsilon) \quad (2.21)$$

Mnożąc przez $(1 + v)^2$ i zbierając podobne wyrazy mamy

$$2p\left(1 + \frac{7}{2}v + 4v^2 + \frac{3}{2}v^3\right) = -3v^3 + 8\epsilon(1 + 2v + v^2). \quad (2.22)$$

Najpierw znajdziemy wykładnik krytyczny δ dla izotermy krytycznej. Na izotermie tej $T = T_c$, a więc $\epsilon = 0$ i stąd

$$p = -\frac{3}{2}v^3\left(1 + \frac{7}{2}v + \cdots\right)^{-1}. \quad (2.23)$$

Ponieważ w pobliżu PK v jest małe, więc

$$p = -\frac{3}{2}v^3\left(1 - \frac{7}{2}v + \cdots\right). \quad (2.24)$$

Wyraz wiodący to v^3, a więc $\delta = 3$.
Wykładnik dla izochory krytycznej (κ_T, gdy $T \downarrow T_c$) otrzymamy kładąc $v = 0$. Ponieważ

$$-(V\kappa_T)^{-1} = \left(\frac{\partial P}{\partial V}\right)_T, \quad (2.25)$$

więc wykorzystując związek P i p oraz V i v, mamy

$$-(V\kappa_T)^{-1} = \frac{P_c}{V_c}\left(\frac{\partial p}{\partial v}\right)_T = \frac{P_c}{V_c}(-6\epsilon) \quad (2.26)$$

czyli

$$\kappa_T = \frac{V_c}{6VP_c}\frac{1}{\epsilon} = const\ \epsilon^{-1}. \quad (2.27)$$

A więc $\gamma = 1$. Dla parametru porządku $(\rho - \rho_c)$ dostajemy $\beta = \frac{1}{2}$, natomiast dla ciepła właściwego $\alpha = 0$. Ostatecznie dla teorii van der Waalsa otrzymujemy następujące wartości wykładników krytycznych

$$\alpha = 0, \qquad \beta = \frac{1}{2}, \qquad \gamma = 1, \qquad \delta = 3. \quad (2.28)$$

2. Metoda pola średniego

Liczba ścisłych wyników dotyczących przejść fazowych nie jest niestety długa – twierdzenie Van Howe'a o tym, że żaden układ 1D o skończonym zasięgu oddziaływań nie może istnieć w stanie uporządkowanym w temperaturze większej niż 0, ścisłe rozwiązanie Onsagera dla 2D modelu Isinga bez zewnętrznego pola oraz twierdzenie Mermina-Wagnera, mówiące, że model Heisenberga i model XY nie mogą w 2D i w $T > 0$ istnieć w stanie uporządkowanym. Dlatego też w badaniach dużą rolą odgrywają metody przybliżone. Dają one wiarygodne wyniki tylko w pewnym zakresie zmienności parametrów, jednak pozwalają na wykonanie szacowań. Kilka z tych metod oparte jest na zasadzie wariacyjnej, która sama ma charakter ścisły.

2.1. Twierdzenie Peierlsa. Niech φ_n będzie dowolnym, zupełnym układem ortonormalnych funkcyj, nie będących funkcjami własnymi hamiltonianu $\mathcal{H}$ rozpatrywanego układu. Wówczas dla sum stanów mamy

$$Z \geq Z_{mod}, \qquad \text{gdzie} \quad Z = \sum_n e^{-\beta E_n}, \quad Z_{mod} = \sum_n e^{-\beta(\varphi_n, \mathcal{H}\varphi_n)}, \tag{2.29}$$

lub dla energii swobodnych

$$F \leq F_{mod}, \quad F = -kT \ln Z, \quad F_{mod} = -kT \ln Z_{mod}. \tag{2.30}$$

E_n to wartości własne $\mathcal{H}$, a więc Z to rzeczywista, nie przybliżona, suma stanów. Wyrażone słowami twierdzenie Peierlsa mówi, że jeżeli uzyskaliśmy przybliżone, czyli modelowe, wyrażenie dla sumy stanów, to leży ono poniżej ścisłej wartości sumy stanów. Odwrotnie jest dla energii swobodnych.

Twierdzenie Peierlsa udowodnimy w postaci

$$Z \geq \sum_n e^{-\beta(\varphi_n, \mathcal{H}\varphi_n)}. \tag{2.31}$$

Oznaczmy

$$Z = \sum_n (\varphi_n, e^{-\beta\mathcal{H}}\varphi_n), \quad Z_{mod} = \sum_n e^{-\beta(\varphi_n, \mathcal{H}\varphi_n)}. \tag{2.32}$$

Dowód oprzemy na lemacie

Lemat

Jeżeli c_n jest zbiorem liczb rzeczywistych spełniających warunki

$$c_n \geq 0, \quad \sum_n c_n = 1 \tag{2.33}$$

a więc mających charakter prawdopodobieństw, oraz

$$\langle f(x) \rangle = \sum_n c_n f(x_n) \quad \text{dla} \quad \forall\, f(x_n), \tag{2.34}$$

wówczas jeżeli

$$f''(x_n) \geq 0, \tag{2.35}$$

to

$$\langle f(x) \rangle \geq f(\langle x \rangle), \quad \text{gdzie} \quad \langle x \rangle = \sum_n c_n x_n. \tag{2.36}$$

Dowód

Skorzystajmy z różniczkowego twierdzenia o wartości średniej

$$f(x_n) = f(\langle x \rangle) + (x_n - \langle x \rangle)\, f'(\langle x \rangle) + \frac{1}{2}\,(x_n - \langle x \rangle)^2\, f''(x_1), \tag{2.37}$$

gdzie x_1 jest pewną ustaloną liczbą rzeczywistą. Wówczas

$$\sum_n c_n f(x_n) = \sum_n c_n f(\langle x \rangle) + 0 + \frac{1}{2}\langle (x_n - \langle x \rangle)^2 \rangle f''(x_1). \tag{2.38}$$

Jeżeli jednak $f''(x_1) \geq 0$, to ostatni wyraz jest nieujemny, $\sum c_n = 1$, a więc

$$\langle f(x)\rangle \geq f(\langle x\rangle). \tag{2.39}$$

Udowodniliśmy w szczególności, że

$$\langle e^x\rangle \quad \geq \quad e^{\langle x\rangle}. \tag{2.40}$$

Powróćmy teraz do dowodu twierdzenia Peierlsa. Niech ψ_n będzie układem funkcyj własnych hamiltonianu rozpatrywanego układu $\mathcal{H}\psi_n = E_n\psi_n$. Istnieje przekształcenie unitarne S_{mn} przeprowadzające funkcje zbioru φ_n w ψ_n

$$\varphi_n = \sum_m S_{mn}\psi_m, \tag{2.41}$$

przy czym

$$\sum_n S^*_{nm}S_{nk} = \sum_n S^*_{mn}S_{kn} = \delta_{mk}, \tag{2.42}$$

a więc

$$\sum_n |S_{kn}|^2 = 1. \tag{2.43}$$

Stąd dostajemy

$$(\varphi_n, \mathcal{H}\varphi_n) = \sum_k\sum_m (S^*_{mn}\psi^*_m, \mathcal{H}S_{nk}\psi_k) = \sum_{km} S^*_{nm}S_{nk}E_k(\psi_m,\psi_k) = \sum_k E_k|S_{nk}|^2. \tag{2.44}$$

A więc

$$Z - Z_{mod} = \sum_n \left(\varphi_n, e^{-\beta\mathcal{H}}\varphi_n\right) - e^{-\beta(\varphi_n,\mathcal{H}\varphi_n)} = \sum_n\left(\sum_k |S_{kn}|^2 e^{-\beta E_k} - e^{\beta\sum|S_{nk}|^2E_k}\right). \tag{2.45}$$

Ale $|S_{nk}|$ gra teraz rolę c_n z lematu, więc

$$\sum_k |S_{nk}|^2 e^{-\beta E_k} = \langle f(E_k)\rangle \tag{2.46}$$

$$\exp\left(-\beta\sum_k |S_{nk}|^2E_k\right) = f(\langle E_k\rangle) \tag{2.47}$$

Mamy więc ostatecznie

$$Z - Z_{mod} = \sum_n \left(\langle f(E_n)\rangle - f(\langle E_n\rangle)\right). \tag{2.48}$$

Każdy wyraz tej sumy jest postaci $\langle f(E_n)\rangle - f(\langle E_n\rangle)$, a to, zgodnie z lematem jest nieujemne, a więc i cała suma jest nieujemna

$$Z - Z_{mod} \geq 0. \tag{2.49}$$

Twierdzenie Peierlsa wykorzystuje się do oszacowań energii swobodnej dla układów oddziałujących. W wielu przypadkach oddziaływanie nie jest silne, a więc w hamiltonianie część swobodna przeważa nad częścią opisującą oddziaływania. Jeżeli teraz konstruujemy hamiltonian przybliżony, to na ogół będzie on zawierał dodatkowy parametr wprowadzony z zewnątrz do teorii. Również i energia swobodna będzie zależała od tego parametru. Z twierdzenia Peierlsa wiadomo, że taka energia jest nie mniejsza niż prawdziwa. Aby więc otrzymać najlepsze, w danej klasie, przybliżenie należy zminimalizować F_{mod} ze względu na wprowadzony parametr.

2.2. MFA dla ferromagnetyka. Jako przykład zastosowania metody wariacyjnej wyprowadzimy równanie stanu dla ferromagnetyka. Dla prostoty przyjmijmy, że jest on opisywany modelem Isinga o spinie 1/2, czyli $S^z = \sigma$.

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - H \sum_i \sigma_i, \tag{2.50}$$

gdzie $\langle ij \rangle$ oznacza, że sumowanie ograniczone jest do najbliższego sąsiedztwa, natomiast H jest zewnętrznym polem magnetycznym przyłożonym wzdłuż kierunku kwantowania spinów, czyli np wzdłuż osi Oz. Ponieważ nie potrafimy rozwiązać ściśle takiego modelu, nawet w 2D, dokonujemy przybliżenia. Idea polega na tym aby zastąpić wyraz opisujący oddziaływania takim, który doprowadzi do tego, że hamiltonian będzie miał postać hamiltonianu jednocząstkowego, tzn. opisującego oddziaływanie jednej cząstki z resztą. Ta reszta będzie to pole efektywne, które pochodzi od wszystkich pozostałych spinów układu. Dla takiego przybliżenia przyjęty jest skrót angielskiej nazwy **MFA – mean field approximation**.

Przyjmijmy, że wychylenia spinów od wartości średniej są niewielkie, tzn., $\langle \sigma \rangle - \sigma$ jest małe. Wobec tego

$$(\langle \sigma \rangle - \sigma_i)(\langle \sigma \rangle - \sigma_j) \simeq 0 = \langle \sigma \rangle^2 - 2\sigma_i \langle \sigma \rangle + \sigma_i \sigma_j \Rightarrow \sigma_i \sigma_j = 2\sigma_i \langle \sigma \rangle - \langle \sigma \rangle^2. \tag{2.51}$$

Przyjęliśmy dla uproszczenia, że układ jest niezmienniczy względem przesunięć, a więc średnia wartość (np magnetyzacji) nie zależy od węzła. Można pokazać, że wprowadzony z zewnątrz parametr $\langle \sigma \rangle = m$, ma sens magnetyzacji na węzeł. $M = \sum_n \langle \sigma_n \rangle = Nm$. Korzystając z (2.51) możemy zamienić hamiltonian dwu-spinowy na jedno-spinowy, zamieniając iloczyn spinów przez funkcję wartości średniej

$$\begin{aligned} \mathcal{H}_{mod} = -J \sum_{\langle i,j \rangle} \left(2\langle \sigma \rangle \sigma_i - \langle \sigma \rangle^2 \right) - H \sum_i \sigma_i \\ = JNz\langle \sigma \rangle^2 - (2Jz\langle \sigma \rangle + H) \sum_i \sigma_i = A - H_{ef} \sum_i \sigma_i, \end{aligned} \tag{2.52}$$

gdzie

$$A \equiv JNzm^2, \qquad H_{ef} = 2zJm + H. \tag{2.53}$$

z jest liczbą najbliższych sąsiadów (liczbą koordynacyjną siatki), natomiast H_{ef} to *pole efektywne* działające na wybrany, ale dowolny, spin i pochodzące od wszystkich pozostałych spinów. Początkowy hamiltonian nie zależał od m, wprowadziliśmy więc zewnętrzny parametr. Należy teraz skorzystać z zasady wariacyjnej (twierdzenia Peierlsa) aby otrzymać możliwie najlepsze przybliżenie prawdziwej sumy stanów.

$$\begin{aligned} Z_{mod} &= \mathrm{Tr}\ e^{-\beta \mathcal{H}_{mod}} \\ &= \mathrm{Tr}\ \exp\left(-\beta A + \beta H_{ef} \sum_i \sigma_i \right) \\ &= e^{-\beta A} Tr \prod_{i=1}^{N} e^{h_{ef} \sigma_i} = e^{-\beta A} \prod_{i=1}^{N} \sum_{\sigma_i = \pm 1} e^{h_{ef} \sigma_i} \\ &= e^{-\beta A} \prod_{1}^{N} 2 \cosh h_{ef} = 2^N e^{-\beta A} \cosh^N (h_{ef}). \end{aligned}$$

gdzie $h_{ef} = \beta H_{ef}$. Dla energii swobodnej mamy

$$F_{mod} = -kT \ln Z_{mod} = -kT \left(-\beta A + N \ln 2 + N \ln \cosh h_{ef} \right). \tag{2.54}$$

Warunek na ekstremum, z którego wyznaczymy m, ma postać

$$\frac{\partial F_{mod}}{\partial m} = 0 \Rightarrow \frac{\partial A}{\partial m} - NkT\ \tanh h_{ef} \frac{\partial h_{ef}}{\partial m} = 2JNzm - 2JNz\ \tanh h_{ef} = 0. \tag{2.55}$$

Stąd już bezpośrednio dostajemy

$$m = \tanh \beta(2Jzm + H), \tag{2.56}$$

co jest równaniem przestępnym na wyznaczenie takiej wartości parametru m, która minimalizuje F_{mod}, a więc daje możliwie najlepsze przybliżenie prawdziwej sumy stanów. Jest to jednocześnie równanie stanu dla ferromagnetycznego modelu Isinga w przybliżeniu MFA. Po wyznaczeniu z niego m, znamy h_{ef}, a więc znamy i sumę stanów oraz energię swobodną. W równaniu (2.56) mamy jedno trywialne rozwiązanie $m = 0$, które odpowiada fazie paramagnetycznej. Jeżeli jednak temperatura jest odpowiednio niska ($T < T_c$), to pojawia się drugie rozwiązanie $m \neq 0$. Rozwiązanie to można znaleźć albo numerycznie rozwiązując równanie (2.56) albo wykreślając w płaszczyźnie (m, h_{ef}) dwie krzywe

$$m = \tanh(h_e), \quad h_{ef} = 2\beta Jzm.$$

Przecięcie tych krzywych wyznacza szukaną wartość m.

Z równania stanu możemy oszacować wartość temperatury krytycznej T_c. Interesuje nas punkt w którym spontaniczne namagnesowanie znika, kładziemy więc $H = 0$. Ponieważ w pobliżu T_c magnetyzacja jest mała, możemy dokonać przybliżenia,

$$m \,=\, \tanh \beta(2Jzm) \simeq (kT_c)^{-1} 2Jzm, \tag{2.57}$$

a stąd

$$kT_c = 2Jz. \tag{2.58}$$

Jak widać, wartość temperatury krytycznej zależy od wartości całki wymiany i liczby najbliższych sąsiadów. Jest to zresztą jedyne miejsce w teorii MFA gdzie wchodzi wymiar modelu. Z tego punktu widzenia nie ma różnicy pomiędzy układem na 2D siatce trójkątnej ($z = 6$) i 3D siatce sześciennej prostej o tej samej liczbie najbliższych sąsiadów. Ponieważ jednocześnie przybliżenie MFA nie odróżnia modelu Isinga i modelu Heisenberga, widać jak bardzo przybliżone są wyniki otrzymane tą metodą. 2D model Isinga ma przejście fazowe, natomiast 2D model Heisenberga – nie ma. Nawet zresztą dla modelu Isinga wyniki otrzymane tą metoda są słuszne tylko jakościowo, a i to nie zawsze. Zaniedbaliśmy bowiem indywidualny charakter oddziaływań między spinami. Będzie to widoczne wszędzie tam, gdzie rola pojedynczych oddziaływań jest duża – w niskich temperaturach, bo tam te efekty są najistotniejsze i w wysokich, gdyż nie pojawi się tam, istniejący w rzeczywistości, bliski porządek.

Otrzymany powyżej wzór dla temperatury krytycznej daje wartości różniące się znacznie od wartości dokładnych, bądź wyznaczonych dokładniejszymi metodami. Np dla 2D modelu Isinga na siatce kwadratowej ($z = 4$) dokładny wynik Onsagera to

$$\frac{kT_c}{J} \,=\, 1.135, \tag{2.59}$$

natomiast MFA daje

$$\frac{kT_c}{J} \,=\, 8. \tag{2.60}$$

Znajdziemy teraz wykładniki krytyczne w teorii średniego pola. Wykorzystując tożsamość trygonometryczną

$$\tanh(x + y) \,=\, \frac{\tanh(x) + \tanh(y)}{1 + \tanh(x)\tanh(y)},$$

dla równania stanu (2.56), możemy je przepisać w postaci

$$m \,=\, \frac{\tanh(h_{ef}) + \tanh(h)}{1 + \tanh(h_{ef})\tanh(h)}, \tag{2.61}$$

gdzie, wykorzystując znaleziony powyżej wzór dla temperatury krytycznej w przybliżeniu MFA

$$h_{ef} \,=\, \frac{2Jz}{kT} m \,=\, \frac{mT_c}{T}, \quad h \,=\, \frac{H}{kT}. \tag{2.62}$$

Wykorzystując (2.61) możemy napisać

$$\tanh(h) = \frac{m - \tanh(mT_c/T)}{1 - m\,\tanh(mT_c/T)}. \tag{2.63}$$

W bezpośredniej okolicy punktu krytycznego, czyli tam gdzie wyznaczamy wykładniki krytyczne, $T \simeq T_c$, m i h są małe, a więc możemy rozłożyć tangens hiperboliczny w szereg i zatrzymać tylko wiodące wyrazy

$$\frac{H}{kT_c} = \left[m\left(1 - \frac{T}{T_c}\right) + \frac{1}{3}m^3\left(\frac{T_c}{T}\right)^3 + ... \right]\left[1 + m^2\left(\frac{T_c}{T} - \frac{1}{3}m^2\left(\frac{T_c}{T}\right)^3 + ..\right)\right]. \tag{2.64}$$

Na izotermie krytycznej $T = T_c$, a więc

$$H = kT_c\frac{1}{3}m^3 \quad \Rightarrow \quad H \sim m^3, \quad \Rightarrow \quad \delta = 3. \tag{2.65}$$

Aby wyznaczyć wykładnik dla zależności parametru porządku (magnetyzacji) połóżmy $H = 0$ w (2.64). Dostaniemy wtedy po podzieleniu przez m

$$\left[\left(1 - \frac{T_c}{T}\right) + \frac{1}{3}m^2\left(\frac{T_c}{T}\right)^3\right] = 0, \tag{2.66}$$

skąd

$$m^2 = 3\left(\frac{T_c}{T} - 1\right)\left(\frac{T}{T_c}\right)^3 = \frac{T_c - T}{T}\left(\frac{T}{T_c}\right)^2. \tag{2.67}$$

Oznacza to, że

$$m = t^\beta, \quad \Rightarrow \quad \beta = \frac{1}{2}. \tag{2.68}$$

Otrzymaliśmy te same wartości co dla teorii Van der Waalsa. Podobnie jest i dla innych wykładników krytycznych. Wartości te nazywamy klasycznymi. Porównując je z danymi doświadczalnymi widać, że różnią się one od nich znacznie. Powodem jest to, że teoria van der Waalsa i przybliżenie pola molekularnego mają charakter przybliżony.

3. Metoda Efektywnego Hamiltonianu

Jest to teoria opracowana przez T. Oguchiego i stanowi wyjście poza ograniczenie pozostałych metod pola średniego, tzn przyjęcie, że pole efektywne jest proporcjonalne do parametru porządku. Inną metodą, gdzie nie stosuje się tego założenia jest *przybliżenie stałego wiązania*, wprowadzone przez Kastelijna i van Kranendonka. Jest to jednak metoda dość skomplikowana.

Istota metody efektywnego hamiltonianu sprowadza się do wyznaczenie macierzy gęstości dla hamiltonianów zawierających odpowiednio jeden i dwa spiny w gronie, a następnie przyrównania, po przeprowadzeniu pewnej operacji, obu macierzy.

Rozważmy spin w węźle R, na który działa oprócz zewnętrznego pola magnetycznego H, także pewne pole efektywne λ, pochodzące od najbliższych sąsiadów. Hamiltonian (jedno cząstkowy) będzie miał w takim przypadku postać

$$\mathcal{H}_1 = -z\lambda S_R^z - (z\delta\lambda + H)S_R^z, \tag{2.69}$$

gdzie drugi wyraz opisuje możliwą zmianę pola efektywnego wywołaną przez włączenie pola zewnętrznego H, tak więc $\delta\lambda \sim H$. Dla prostoty przyjmijmy, że $S = 1/2$, a więc możliwe są dwa stany każdego spinu $|\uparrow> \equiv | \alpha >$ oraz $|\downarrow> \equiv | \beta >$. Interesować nas dalej będzie przede wszystkim wyznaczenie temperatury krytycznej, możemy więc od początku ograniczyć się do niskich pól, gdyż w końcowych wzorach położymy $H = 0$.
Unormowana jedno-cząstkowa macierz gęstości $\hat{\rho}_1$ ma postać

$$\hat{\rho}_1 = \frac{e^{\frac{1}{2}\beta z\lambda}\left[1 + \frac{1}{2}\beta(z\delta\lambda + H)\right] \mid \alpha >< \alpha \mid + e^{-\frac{1}{2}\beta z\lambda}\left[1 - \frac{1}{2}\beta(z\delta\lambda + H)\right] \mid \beta >< \beta \mid}{2\cosh(\frac{1}{2}\beta z\lambda)\left[1 + \frac{1}{2}\beta(z\delta\lambda + H)\tanh(\frac{1}{2}\beta z\lambda\right]}. \tag{2.70}$$

Rozwinęliśmy tu tangens hiperboliczny i zatrzymaliśmy dwa pierwsze wyrazy. Połóżmy teraz $H = 0$, będziemy wtedy mieli także $\delta\lambda = 0$, ponieważ w zerowym polu znika też zmiana pola efektywnego wywołana przez pole H. Macierz gęstości ma postać

$$\hat{\rho}_1 = \frac{e^{\frac{1}{2}\beta z\lambda} \mid \alpha >< \alpha \mid + e^{-\frac{1}{2}\beta z\lambda} \mid \beta >< \beta \mid}{2\cosh(\frac{1}{2}\beta z\lambda)}. \tag{2.71}$$

Rozważmy teraz dwa spiny - jeden w węźle R, drugi w jednym z jego NN - $R + \delta$. Rozpatrzmy model Heisenberga, a więc przyjmijmy, że oddziałują za sobą wektory spinu, a nie tylko ich jedne składowe. Dwu spinowy hamiltonian ma postać

$$\mathcal{H}_2 = -2J\left(\vec{S}_R \cdot \vec{S}_{R+\delta}\right) - (z-1)\lambda\left(S_R^z + S_{R+\delta}^Z\right) - \left[(z-1)\delta\lambda + H\right]\left(S_R^z + S_{R+\delta}^Z\right). \tag{2.72}$$

Wektory własne $\mathcal{H}_2$ mają postać
triplet (stan s)

$$\mid \alpha\alpha >, \quad \frac{1}{\sqrt{2}} \mid \alpha\beta + \beta\alpha >, \quad \mid \beta\beta > \tag{2.73}$$

oraz singlet (stan p)

$$\frac{1}{\sqrt{2}} \mid \alpha\beta - \beta\alpha >, \tag{2.74}$$

gdzie pierwszy wskaźnik odnosi się do spinu w węźle R, natomiast drugi, do spinu węźle $R + \delta$. Dla $H = 0$ unormowana, dwu-spinowa macierz gęstości $\hat{\rho}_2$, ma postać

$$\hat{\rho}_2 = \frac{1}{2}\left[\cosh\beta(z-1)\lambda + e^{\beta J}\cosh\beta J\right]^{-1} * \tag{2.75}$$

$$\Big[e^{\beta(z-1)\lambda} \mid \alpha\alpha >< \alpha\alpha \mid + \frac{1}{2} \mid \alpha\beta + \beta\alpha >< \alpha\beta + \beta\alpha \mid + \tag{2.76}$$

$$e^{-\beta(z-1)\lambda} \mid \beta\beta >< \beta\beta \mid + \frac{1}{2}e^{-2\beta J} \mid \alpha\beta - \beta\alpha >< \alpha\beta - \beta\alpha \mid\Big]. \tag{2.77}$$

Kluczowy dla całej metody warunek samouzgodnienia ma postać

$$Tr_{R+\delta}\hat{\rho}_2 = \hat{\rho}_1. \tag{2.78}$$

Wykonując ten ślad po drugim operatorze i przyrównując współczynniki przy $\mid \alpha\alpha >< \alpha\alpha \mid$ oraz $\mid \beta\beta >< \beta\beta \mid$ w (2.78) dostaniemy

$$\frac{e^{\frac{1}{2}\beta J}}{2\cosh\frac{1}{2}\beta z\lambda} = \frac{1}{2}\frac{e^{\beta(z-1)\lambda} + e^{-\beta J}\cosh\beta J}{\cosh\beta(z-1)\lambda + e^{-\beta J}\cosh\beta J}, \tag{2.79}$$

$$\frac{e^{-\frac{1}{2}\beta J}}{2\cosh\frac{1}{2}\beta z\lambda} = \frac{1}{2}\frac{e^{-\beta(z-1)\lambda} + e^{-\beta J}\cosh\beta J}{\cosh\beta(z-1)\lambda + e^{-\beta J}\cosh\beta J} \tag{2.80}$$

Dodanie do siebie obu wyrażeń daje tożsamość, natomiast odjęcie daje

$$\tanh\frac{1}{2}\beta z\lambda = \frac{\sinh\beta(z-1)\lambda}{\cosh\beta(z-1)\lambda + e^{-\beta J}\cosh\beta J}. \tag{2.81}$$

Równanie to zawiera nieznany parametr – pole efektywne λ. Jak widać, $\lambda = 0$ jest zawsze rozwiązaniem, bez względu na wartość temperatury. Tak samo było też w MFA. Nieco poniżej punktu krytycznego $\beta\lambda$ powinno być niewielkie. Możemy to wykorzystać i rozłożyć obie strony (2.81) w szereg, zatrzymując tylko wyrazy liniowe w $\beta\lambda$. Dostaniemy wtedy

$$\frac{1}{2}\beta_c z\lambda = \frac{\beta_c(z-1)\lambda}{1 + e^{-\beta_c J}\cosh\beta_c J}, \tag{2.82}$$

co da się zapisać prościej jako

$$ze^{-\beta_c J}\cosh\beta_c J = z - 2 \quad \Rightarrow \quad \frac{J}{kT_c} = \frac{1}{2}\ln\frac{z}{z-4}. \tag{2.83}$$

Równanie to wyznacza temperaturę krytyczną dla ferromagnetyka typu Heisenberga o spinie $S = \frac{1}{2}$ w przybliżeniu efektywnego hamiltonianu. Dla modeli dwu-wymiarowych, gdzie $z \leq 4$, $T_c = 0$, zgodnie z bardziej dokładnymi teoriami.

Podobne rozważania można też przeprowadzić dla modelu Isinga, otrzymując następujący wzór dla temperatury krytycznej

$$\frac{J}{kT_c} = \ln \frac{z}{z-2}. \tag{2.84}$$

Tutaj dla układów 1D, gdzie $z \leq 2$, $T_c = 0$, zgodnie z twierdzeniem Van Howe'a. Można też wyznaczyć zależność od temperatury magnetyzacji, a zatrzymując zależność od zewnętrznego pola, znaleźć podatność w obu modelach.

Metoda efektywnego hamiltonianu daje więc o wiele lepsze wyniki niż MFA, aczkolwiek zgodność z danymi doświadczalnymi jest raczej jakościowa. Dla 2D modelu Isinga wartość temperatury krytycznej otrzymana w przybliżeniu efektywnego hamiltonianu to

$$\frac{kT_c}{J} = 1.442, \tag{2.85}$$

natomiast ścisły wynik Onsagera to

$$\frac{kT_c}{J} = 1.135. \tag{2.86}$$

Wartości wykładników krytycznych otrzymanych w tej metodzie są klasyczne, tzn takie same jak dla MFA i teorii Van der Waalsa.

Metodę efektywnego hamiltonianu można też stosować, po odpowiednich zmianach, do wyznaczenia temperatury krytycznej antyferromagnetyków.

4. Teoria Landaua

Jest to teoria fenomenologiczna, której podstawowymi założeniami są:

- Przy ciągłym przejściu fazowym układ przechodzi od stanu o wyższej symetrii (np. paramagnetyk) do stanu o niższej symetrii (ferromagnetyk). Stan o niższej symetrii może być opisany przez jeden, lub więcej, parametrów łamiących symetrię (PP, np. magnetyzację)
- Parametr porządku jest ciągłą funkcją temperatury T. Założenie to nie jest spełnione dla przejść fazowych 1 rodzaju, gdzie PP zmienia się skokowo.
- W otoczeniu PK energia swobodna, ogólnie potencjał termodynamiczny, może być rozłożony w szereg Taylora względem potęg PP. Ukryte tu jest dodatkowe założenie, że energia oprócz normalnych parametrów termodynamicznych, zależy także od PP. Warunek stabilności tej energii ze względu na wprowadzony zewnętrzny parametr (tw. Peierlsa) służy jako samouzgodnienie teorii, zgrywając wprowadzony parametr z wielkością fizyczną, np. magnetyzacją.

Przyjmijmy, że mamy ferromagnetyk w słabym polu magnetycznym H i temperaturę w okolicy T_c. Energia swobodna Gibbsa $G = G(M, T, H)$ może być zapisana poprzez gęstość tej energii

$$G(M, T, H) = \int d^3 r g(\vec{r}, M, T, H). \tag{2.87}$$

Dla uproszczenia notacji nie będziemy dalej zaznaczali jawnie zależności od M, T, H. Skierujmy zewnętrzne pole wzdłuż osi Oz, wówczas dla układu izotropowego, a takim teraz będziemy się zajmowali, magnetyzacja skierowana jest wzdłuż pola, czyli

$$M \equiv M^z. \tag{2.88}$$

W pobliżu T_c, magnetyzacja jest mała, możemy więc szereg Taylora zakończyć na wyrazach 4 rzędu :

$$g(\vec{r}) = g_0(T) - H(\vec{r})M(\vec{r}) + a(T)\left[M(\vec{r})\right]^2 + b(T)\left[M(\vec{r})\right]^4 + c(T)\left[\nabla M(\vec{r}) \cdot \nabla M(\vec{r})\right]. \tag{2.89}$$

Pierwszy, stały, wyraz to gęstość energii swobodnej gdy $M = 0$. Drugi to oddziaływanie spinów z zewnętrznym polem magnetycznym. Wyrazy proporcjonalne do a i b, to oddziaływania spin-spin.

Ponieważ oddziaływania te nie powinny zależeć od znaku spinu (np. zamiany wszędzie $S = 1$ na $S = -1$ przy jednoczesnej zamianie znaku pola magnetycznego $H \to -H$), więc w rozwinięciu nie występują nieparzyste potęgi magnetyzacji. Niezmienniczość ta jest równoważna niezmienniczości przy zamianie znaku strzałki czasu (*time reversal symmetry).* Ostatni wyraz uwzględnia fluktuacje PP (magnetyzacji). Wyraz ten będzie wzmacniał fluktuacje jeżeli zmiana przestrzenna $M(\vec{r})$ będzie prowadzić do zmniejszenia energii swobodnej i będzie tłumił fluktuacje jeżeli energia będzie rosła. Widać więc, że o charakterze zależności decyduje znak współczynnika $c(T)$.

Zgodnie z mechaniką statystyczną, najbardziej prawdopodobna wartość parametru to ta, która minimalizuje energię swobodną. Znajdziemy więc teraz wartość $M(\vec{r})$, która zapewni, że G będzie stacjonarne przy infinitezymalnej zmianie, tzn.

$$M \to M(\vec{r}) + \delta M(\vec{r}), \quad \delta G = G(M + \delta M) - G(M). \tag{2.90}$$

Z dokładnością do wyrazów 1-go rzędu daje to

$$\delta G = \int d^3 r \delta M(\vec{r}) \left(-H(\vec{r}) + 2a(T)M(\vec{r}) + 4b(T)\left[M(\vec{r})\right]^3 - 2c(T)\nabla^2 M(\vec{r}) \right) = 0, \tag{2.91}$$

a ponieważ ma to być spełnione dla każdego δM, więc $(\cdots) = 0$, czyli

$$\left(2a(T) + 4b(T)\left[M(\vec{r})\right]^2 - 2c(T)\nabla^2 \right) M(\vec{r}) = H(\vec{r}). \tag{2.92}$$

W tym miejscu robione jest drastyczne założenie, o zaniedbaniu wszystkich fluktuacji, co odpowiada przyjęciu, że najbardziej prawdopodobna wartość $(M(\vec{r}))$ jest jedyną wartością, a więc i średnią. Rozwiązując, ze względu na $M(\vec{r})$, ostatnią równość i wstawiając uzyskaną wartość do rozwinięcia, uzyskujemy wyrażenie dla energii swobodnej $G(\vec{r})$. Jeżeli zaniedbamy fluktuacje, to gęstość energii swobodnej Gibbsa ma postać

$$g = g_0 - H(\vec{r})M(\vec{r}) + a(T)M^2(\vec{r}) + b(T)M^4(\vec{r}). \tag{2.93}$$

Stąd warunek ekstremum

$$\frac{\partial g}{\partial M} = 0 \quad \Rightarrow \quad 2a(T)M + 4b(T)M^3 = H \tag{2.94}$$

i warunek stabilności

$$\frac{\partial^2 g}{\partial m^2} = 2a(T) + 12b(T)M^2 > 0 \tag{2.95}$$

Wyłączając zewnętrzne pole ($H = 0$), dostajemy z warunku koniecznego dla ekstremum dwa rozwiązania

$$M = 0 \quad \text{oraz} \quad M = \sqrt{-a(T)/2b(T)}. \tag{2.96}$$

Pierwsze rozwiązanie odpowiada fazie paramagnetycznej, drugie ferromagnetycznej. Aby stabilna była faza para musi być

$$a(T) > 0. \tag{2.97}$$

Dla fazy ferro mamy, z żądania stabilności gdy $M \neq 0$, warunek $b(T) > 0$, w przeciwnym przypadku nie jest spełniony warunek konieczny ekstremum. Wobec tego, aby magnetyzacja była nieurojona $a(T) < 0$ w fazie ferro. Stąd wniosek, że współczynnik $a(T)$ zmienia znak przy przejściu przez T_c, dlatego też na ogół przyjmuje się, że współczynnik ten jest proporcjonalny do różnicy $(T - T_c)$, tzn $a(T) = a(T - T_c)$, natomiast b nie zależy od T. Wartości wykładników krytycznych są klasyczne.

4.1. Fluktuacje w ramach teorii Landaua. Ponieważ w prostej teorii Landaua przedstawionej powyżej zaniedbaliśmy korelacje PP, spróbujmy teraz oszacować ich rolę. Rozważać będziemy fluktuacje magnetyzacji $M(\vec{r})$, dane przez $[M(\vec{r}) - \langle M(\vec{r})\rangle]$. Chcemy zobaczyć jak odstępstwo od wartości średniej w jednym punkcie powiązane jest z analogicznym odstępstwem w innym punkcie. Opisuje to funkcja korelacyjna

$$g(\vec{r}, \vec{r}\prime) = \langle [M(\vec{r}) - \langle M(\vec{r})\rangle] \left[M(\vec{r}\prime) - \langle M(\vec{r}\prime)\rangle\right]\rangle. \tag{2.98}$$

W teorii Landaua przyjęliśmy, że korelacje takie są bardzo małe, tymczasem wiadomo, że w okolicy T_c są one bardzo duże. Ogranicza to oczywiście zakres stosowalności teorii.

Postarajmy się teraz znaleźć zakres stosowalności teorii Landaua. Jeżeli hamiltonian zawiera parametr $h(\vec{r})$ i odpowiadający mu parametr ekstenzywny $p(\vec{r})$, (jak np. para magnetyzacja-pole magnetyczne), w kombinacji $\int d\vec{r} h(\vec{r})p(\vec{r})$, wówczas zmiana $h \to h + \delta h$ prowadzi do zmiany $\langle p(\vec{r})\rangle$ wyrażonej przez funkcję korelacyjną. Mamy bowiem

$$\mathcal{H} = \mathcal{H}_0 - \int d\vec{r}\, p(\vec{r})h(\vec{r}) \tag{2.99}$$

oraz

$$\langle p(\vec{r})\rangle = \frac{\mathrm{Tr}\, p(\vec{r})e^{-\beta\mathcal{H}}}{\mathrm{Tr}\, e^{-\beta\mathcal{H}}}. \tag{2.100}$$

Po podstawieniu $h \to h + \delta h$, otrzymamy

$$\begin{aligned}
\langle p(\vec{r})\rangle + \delta\langle p(\vec{r})\rangle &= \frac{\mathrm{Tr}\, p(\vec{r})e^{-\beta\mathcal{H}_0+\beta\int d\vec{r}p(\vec{r})h(\vec{r})+\beta\int d\vec{r}p(\vec{r})\delta h(\vec{r})}}{\mathrm{Tr}\, e^{-\beta\mathcal{H}_0+\beta\int d\vec{r}p(\vec{r})h(\vec{r})+\beta\int d\vec{r}p(\vec{r})\delta h(\vec{r})}} \\
&\simeq \frac{\mathrm{Tr}\, p(\vec{r})e^{-\beta\mathcal{H}}\left(1+\beta\int d\vec{r'}p(\vec{r'})\delta h(\vec{r'})\right)}{\mathrm{Tr}\, e^{-\beta\mathcal{H}}\left(1+\beta\int d\vec{r'}p(\vec{r'})\delta h(\vec{r'})\right)} \\
&= Z^{-1}\left(\mathrm{Tr}\left[p(\vec{r})e^{-\beta\mathcal{H}}\right] + \beta\ \mathrm{Tr}\, p\int d\vec{r'}p(\vec{r'})\delta h(\vec{r'})e^{-\beta\mathcal{H}}\right)\left(1-\beta\int d\vec{r'}\,\langle p(\vec{r'})\rangle\delta h(\vec{r'})\right) \\
&= \left[\langle p(\vec{r})\rangle + \beta\int d\vec{r'}\langle p(\vec{r})p(\vec{r'})\rangle\delta h(\vec{r'})\right]\left[1-\beta\int d\vec{r'}\langle p(\vec{r'})\rangle\delta h(\vec{r'})\right] \\
&= \langle p(\vec{r})\rangle + \beta\int d\vec{r'}\left[\langle p(\vec{r})p(\vec{r'})\rangle - \langle p(\vec{r})\rangle\langle p(\vec{r'})\rangle\right]\delta h(\vec{r'})
\end{aligned}$$

A więc

$$\delta\langle p(\vec{r})\rangle = \beta\int d\vec{r'}\langle[p(\vec{r}) - \langle p(\vec{r})\rangle]\,[p(\vec{r'}) - \langle p(\vec{r'})\rangle]\rangle\delta h(\vec{r'}). \tag{2.101}$$

Jeżeli więc zmienimy pole magnetyczne H dodając mały przyczynek δH, wówczas $\langle p(\vec{r})\rangle$, czyli średnia zmiana magnetyzacji dana jest przez

$$\delta\langle M(\vec{r})\rangle = \beta\int d^3r' g(\vec{r},\vec{r'})\delta H(\vec{r'}). \tag{2.102}$$

Wykorzystajmy teraz wyprowadzony poprzednio warunek stabilności, w którym zamienimy $M \to M + \delta M, \quad H \to H + \delta H$. Otrzymamy wówczas

$$\left[2a + 12b\langle M(\vec{r})\rangle^2 - 2c\nabla^2\right]\delta\langle M(\vec{r})\rangle = \delta H(\vec{r}). \tag{2.103}$$

Podstawiając tu znalezione przed chwilą wyrażenie dla $\delta\langle M\rangle$, dostajemy

$$\int d^3r'\left(\left[2a + 12b\langle M(\vec{r})\rangle^2 - 2c\nabla^2\right]g(\vec{r},\vec{r'}) - kT\delta(\vec{r'}-\vec{r})\right)\delta H(\vec{r'}) = 0. \tag{2.104}$$

Ponieważ δH jest dowolne, więc $(\cdots) = 0$, co daje równanie dla $g(\vec{r},\vec{r'})$

$$\left[2a + 12b\langle M(\vec{r})\rangle^2 - 2c\nabla^2\right]g(\vec{r},\vec{r'}) = kT\delta(\vec{r}-\vec{r'}) \tag{2.105}$$

Jeżeli $H = 0$, to dla $T > T_c$ mamy $\langle M\rangle = 0$, czyli uwzględniając to, że $a = A(T - T_c)$

$$\left[2A(T-T_c) - 2c\nabla^2\right]g(\vec{r},\vec{r'}) = kT\delta(\vec{r}-\vec{r'}) \qquad (T > T_c) \tag{2.106}$$

Natomiast dla $T < T_c$ kwadrat magnetyzacji $M \sim (T_c - T)$, czyli

$$\left[4B(T_c-T) - 2c\nabla^2\right]g(\vec{r},\vec{r'}) = kT\delta(\vec{r}-\vec{r'}) \qquad (T < T_c). \tag{2.107}$$

W trzech wymiarach rozwiązaniem tych równań jest

$$g(\vec{r},\vec{r'}) = |\vec{r}-\vec{r'}|^{-1}\exp\left(-\frac{|\vec{r}-\vec{r'}|}{\xi}\right)\frac{kT}{8\pi c} \tag{2.108}$$

gdzie

$$\xi = \sqrt{c/A}(T - T_c)^{-1/2} \qquad (T > T_c), \tag{2.109}$$

$$\xi = \sqrt{c/B}(T_c - T)^{-1/2} \qquad (T < T_c). \tag{2.110}$$

Otrzymane rozwiązanie dla g jest słuszne jeżeli $|\vec{r} - \vec{r}\prime|$ jest dostatecznie duże. Chodzi tu o to, że stosujemy klasyczną fizykę statystyczną do opisu problemu kwantowo-mechanicznego. Jeżeli jednak obszar jaki badamy ($|\vec{r} - \vec{r}\prime|$) zawiera kilka komórek elementarnych, możemy wówczas uśrednić po tych komórkach i uśrednienia te wytłumią efekt kwantowy. Znalezione wzory dla ξ, czyli promienia korelacji wskazują, że przy $T \to T_c$ promień ten dąży do nieskończoności jak $\varepsilon^{-1/2}$, co daje wartość wykładnika $\nu = 1/2$ w teorii Landaua. Aby teoria Landaua mogła być stosowana fluktuacje nie mogą być zbyt duże – część z nich da się wytłumić poprzez uśrednianie. Nie można jednak uśredniać po zbyt dużym obszarze gdyż wtedy teoria traci swój lokalny charakter, rozwijamy bowiem $M(\vec{r})$. Maksymalną odległością po jakiej możemy uśrednić jest promień korelacji ξ.

Tak więc aby teorię Landaua można było stosować, fluktuacje PP na odległościach rzędu promienia korelacji powinny być małe. W szczególności powinny być dużo mniejsze od samego PP

$$(\delta M)^2_\omega = N_\omega \sum_\omega \left[\langle S_0 S_i\rangle - \langle S\rangle^2\right] \ll M^2_\omega \tag{2.111}$$

ω jest tu obszarem (rzędu promienia korelacji) po którym uśredniamy. Z twierdzenia fluktuacyjno-dyssypacyjnego możemy wielkość w nawiasach kwadratowych utożsamić z podatnością – ściślej z tą jej częścią, która przypada na obszar ω.

$$(\delta M)^2_\omega = C N_\omega \chi_{q=0} \tag{2.112}$$

N_ω jest liczbą spinów w obszarze ω. Z drugiej strony całkowita magnetyzacja w obszarze ω związana jest ze zredukowaną magnetyzacją m poprzez

$$M_\omega = N_\omega m. \tag{2.113}$$

Warunek na to by fluktuacje były małe (w porównaniu z PP), czyli *warunek Ginzburga* ma postać

$$N_\omega \chi \ll N^2_\omega m^2. \tag{2.114}$$

Wiemy, że w pobliżu T_c

$$\chi \sim \varepsilon^{-\gamma}, \quad m \sim \varepsilon^{\beta}. \tag{2.115}$$

Ilość spinów w pewnym obszarze jest proporcjonalna do objętości tego obszaru, który w naszym przypadku jest rzędu promienia korelacji $N \sim V \sim \xi^{d+n}$. n jest liczbą nieujemną, zależną od rozpatrywanego zagadnienia. Sam promień korelacji wokół T_c zachowuje się jak

$$\xi \sim \varepsilon^{-\nu}, \tag{2.116}$$

a więc kryterium Ginzburga stosowalności teorii Landaua można przepisać jako

$$\varepsilon^{-\gamma} \ll \varepsilon^{-(d+n)\nu + 2\beta}. \tag{2.117}$$

Aby związek ten był spełniony dla dowolnie małych ε, musi być

$$\gamma - (d+n)\nu + 2\beta < 0, \tag{2.118}$$

mamy bowiem

$$1 \ll \varepsilon^{\gamma - (d+n)\nu + 2\beta}. \tag{2.119}$$

Ostatecznie otrzymujemy z kryterium Ginzburga tzw. *wymiar krytyczny* d^* powyżej którego teoria Landaua może być stosowana

$$d^* = \frac{\gamma + 2\beta}{\nu} - n. \tag{2.120}$$

Jak widać, dużo zależy od liczby n. Np. dla układu izotropowego o siłach krótkozasięgowych korelacje we wszystkich kierunkach zachowują się tak samo, tzn. są rozbieżne jak ξ, a więc $N_\omega \sim \xi^d \Rightarrow n = 0$. Ponieważ znamy wartości pozostałych wykładników w teorii Landaua, $\beta = \frac{1}{2}, \gamma = 1, \nu = \frac{1}{2}$, możemy więc wyznaczyć d^*, które jest równe 4. Morał jest więc taki, że teorii Landaua (przynajmniej do opisu zjawisk krytycznych) nie można stosować do układów dwu i trój-wymiarowych. Jeżeli mamy magnetyk jednoosiowy, wówczas promień korelacji będzie w tym kierunku rozbieżny silniej $\xi_{\|} \sim \xi^2$, a stąd $N_\omega = \xi\xi\xi_{\|} \sim \xi^4$, czyli $n = 1$. Ponieważ wykładniki pozostają takie same, więc $d^* = 3$. Oznacza to, że teoria Landaua z poprawkami (logarytmicznymi) powinna poprawnie opisywać 3D 1-osiowe ferroelektryki. Doświadczenie potwierdza te wyniki

$$\chi_{q=0} \sim |\varepsilon|^{-1} |\ln \varepsilon|^{1/3} \tag{2.121}$$

$$M \sim \varepsilon^{1/2} |\ln \varepsilon|^{1/3} \tag{2.122}$$

$$C \sim |\ln \varepsilon|^{1/3}. \tag{2.123}$$

ROZDZIAŁ 3

Jakościowa analiza układów dynamicznych

Kinetyka wielu procesów, nie tylko zresztą fizycznych, opisywana jest najczęściej przez układ równań różniczkowych typu

$$\frac{dx_1}{dt} = F_1(x_1, \ldots, x_n), \ldots, \frac{dx_n}{dt} = F_n(x_1, \ldots, x_n), \tag{3.1}$$

gdzie $x = x(t)$, natomiast funkcje F_i mogą być, i najczęściej są, nieliniowe. Przykładem może być układ równań otrzymany z równania Newtona

$$m\frac{d^2x}{d^2t} = F(x), \tag{3.2}$$

które po podstawieniu $y = dx/dt$ ma postać układu równań

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= m^{-1}F(x). \end{aligned}$$

W ogólnym przypadku zmiennymi mogą być współrzędne i pędy uogólnione (q, p). Wówczas płaszczyzną w jakiej opisany jest ruch rozpatrywanego układu dynamicznego jest po prostu przestrzeń fazowa. Jak wiadomo z mechaniki klasycznej, zmienne q, p nie muszą mieć wymiaru położenia i pędu.

Jeżeli funkcje F są nieliniowymi funkcjami swoich argumentów, wówczas rozwiązanie może być na ogół otrzymane tylko na drodze numerycznej, a typ rozwiązania zależy w sposób decydujący od warunków początkowych. Jeżeli więc uda się nam znaleźć numerycznie rozwiązanie dla pewnych wartości początkowych parametrów to nadal nie znamy rozwiązań przy innych warunkach początkowych. Otrzymujemy bowiem w ten sposób tylko rozwiązanie szczególne badanego układu równań różniczkowych, nie znamy jednak rozwiązania ogólnego. Na to aby otrzymać inne rozwiązanie szczególne należy przeprowadzić od nowa wszystkie obliczenia numeryczne, a otrzymanie kilku (czy kilkunastu) takich rozwiązań nie przybliża nas do znajomości rozwiązania ogólnego.

Na szczęście w wielu przypadkach interesuje nas nie tyle dokładna znajomość rozwiązania ogólnego, co poznanie możliwych typów rozwiązań. Jest więc to problem analizy jakościowej, a nie ilościowej. Jednym z częściej stawianych pytań jest problem istnienia stabilnych rozwiązań, czyli stabilności układu przy niewielkiej zmianie wartości parametrów. Odpowiedzi na to pytanie udziela jakościowa teoria układów dynamicznych, która nie podaje jawnej zależności $x_i = x_i(t)$, lecz mówi o tym jaki będzie typ zachowania układu w zależności od funkcji $F_1, ..., F_n$.

1. Portret fazowy

Rozważmy najprostszy układ dynamiczny opisywany przez dwie zmienne x, y i odpowiednie równania różniczkowe

$$\frac{dx}{dt} = P(x, y), \qquad \frac{dy}{dt} = Q(x, y). \tag{3.3}$$

Stan układu w chwili t reprezentowany jest przez punkt w przestrzeni (tu powierzchni) fazowej. Ciąg takich punktów, wyznaczający historię układu, tworzy krzywą w przestrzeni fazowej, tzw. **trajektorię fazową**.

Dzieląc stronami równanie (3.3) rugujemy parametr t i dostajemy równanie

$$\frac{dy}{dx} = \frac{Q(x,y)}{P(x,y)} \tag{3.4}$$

Rozwiązanie ogólne tego równania $y = y(x,C)$ zależy od stałej C, którą należy wyznaczyć z warunku początkowego. Zgodnie z twierdzeniem Cauchy'ego o jednoznaczności rozwiązań, przez każdy punkt przestrzeni (x,y) może przechodzić tylko jedna krzywa całkowa o kącie nachylenia zadanym przez stosunek $\frac{P(x,y)}{Q(x,y)}$. Jeżeli jednak mamy jednocześnie

$$P(x,y) = 0, \qquad Q(x,y) = 0 \tag{3.5}$$

to nachylenie krzywej całkowej w takim punkcie nie jest określone i przez ten punkt może przechodzić więcej niż jedna krzywa całkowa. Takie punkty nazywamy **punktami osobliwymi**. Punkty te mają specjalny charakter, wynikający z ich definicji. Jeżeli bowiem

$$P(x,y) = 0 \Rightarrow \frac{dx}{dt} = 0. \tag{3.6}$$

Podobnie i dla $Q(x,y)$. Jednak równość zeru ostatnich pochodnych oznacza, że w punkcie osobliwym współrzędne nie zmieniają się, a więc punkt osobliwy odpowiada stanowi równowagi układu dynamicznego. (Każdy punkt w przestrzeni fazowej to stan układu !)

Rozwiązanie postawionego zadania, tzn. określenie charakteru rozwiązań w zależności od parametrów czy też warunków początkowych można osiągnąć poprzez wyznaczenie położenia punktów osobliwych, a następnie wykreślenie trajektorii fazowych (krzywych całkowych), co łącznie tworzy **portret fazowy**.

Przykład 1

$$\frac{dx}{dt} = A = const \neq 0 \qquad \frac{dy}{dt} = B = const \neq 0 \tag{3.7}$$

Stąd

$$\frac{dy}{dx} = \frac{B}{A} \tag{3.8}$$

i układ nie ma punktów osobliwych. Rozwiązania to proste o nachyleniu B/A.

Przykład 2

$$\frac{dx}{dt} = -x, \qquad \frac{dy}{dt} = -y \tag{3.9}$$

Tym razem układ posiada punkt osobliwy $x = y = 0$. Wraz z upływem czasu trajektorie fazowe dążą do punktu osobliwego, który nazywamy **stabilnym** lub **trwałym**.

Przykład 3

$$\frac{dx}{dt} = x, \qquad \frac{dy}{dt} = y \tag{3.10}$$

Układ ma ten sam punkt osobliwy, jednak teraz trajektorie fazowe będą się od niego oddalały wraz z upływem czasu. Taki punkt osobliwy nazywamy **niestabilnym** bądź **nietrwałym**.

Na to aby określić nie tylko samą trajektorię fazową, do czego wystarcza równanie nie zawierające już t, ale i kierunek ruchu wzdłuż trajektorii, trzeba skorzystać z układu równań (3.3).

2. Kryteria stabilności punktu osobliwego

Fizyk chcąc sprawdzić czy dany stan układu reprezentuje położenie równowagi trwałej czy nietrwałej, dokonuje małego wychylenia z tego położenia, przykładając niewielką siłę zewnętrzną. Jeżeli układ powróci do poprzedniego stanu, mamy do czynienia z równowagą trwałą, jeśli nie – z nietrwałą. W ten sam sposób będziemy postępowali i teraz. Oznaczmy przez (x^*, y^*) współrzędne punktu osobliwego (punktu równowagi układu), którego stabilność będziemy badali. Niewielkie wychylenie z tego położenia oznaczmy przez (ξ, η), czyli

$$x = x^* + \xi, \qquad y = y^* + \eta \tag{3.11}$$

Podstawmy to do wyjściowego równania (3.3), rozłóżmy prawe strony w szereg Taylora wokół punktu równowagi oraz pomińmy (wobec małości zaburzenia) wyrazy wyższe niż liniowy. Otrzymamy wówczas

$$P(x,y) = P(x^*,y^*) + \left.\frac{\partial P}{\partial x}\right|_{x^*,y^*} \xi + \left.\frac{\partial P}{\partial y}\right|_{x^*,y^*} \eta,$$
$$Q(x,y) = Q(x^*,y^*) + \left.\frac{\partial Q}{\partial x}\right|_{x^*,y^*} \xi + \left.\frac{\partial Q}{\partial y}\right|_{x^*,y^*} \eta.$$

Ale z definicji punktu osobliwego wiemy, że

$$P(x^*,y^*) = Q(x^*,y^*) = 0. \tag{3.12}$$

Tak więc równania te redukują się do

$$\frac{dx}{dt} = \frac{d\xi}{dt} = \xi a_{11} + \eta a_{12},$$
$$\frac{dy}{dt} = \frac{d\eta}{dt} = \xi a_{21} + \eta a_{22}.$$

Otrzymaliśmy w ten sposób układ równań liniowych, dla którego rozwiązań szukamy w postaci

$$\xi = Ae^{\lambda t}, \qquad \eta = Be^{\lambda t}, \tag{3.13}$$

co prowadzi do równania charakterystycznego dla λ o dwu pierwiastkach i daje znane rozwiązania

$$\xi = C_{11}e^{\lambda_1 t} + C_{12}e^{\lambda_2 t},$$
$$\eta = C_{21}e^{\lambda_1 t} + C_{22}e^{\lambda_2 t},$$

przy czym niezależne amplitudy C należy wyznaczyć z warunków początkowych. Jeżeli oba rozwiązania równania charakterystycznego są rzeczywiste, możliwe są następujące przypadki:

(1) $\lambda_1 < 0$ i $\lambda_2 < 0$. Z upływem czasu zarówno ξ jaki i η dążą do zera. Oznacza to, że zaburzenie jest tłumione, ponieważ ξ, η opisują wychylenie ze stanu równowagi, i układ wraca do położenia początkowego. Równowaga jest trwała, czyli punkt osobliwy jest stabilny.

(2) $\lambda_1 > 0$ i $\lambda_2 > 0$, ξ i η rosną z czasem, układ oddala się od położenia początkowego, punkt osobliwy jest niestabilny (równowaga jest nietrwała).

(3) Znaki λ_1 i λ_2 są różne. Dla dostatecznie długich czasów dominuje wyraz zawierający dodatnią potęgę t, a więc układ oddala się od położenia początkowego, które odpowiada równowadze nietrwałej i niestabilnemu punktowi osobliwemu. Wyjątkowy przypadek mamy gdy współczynniki C_{ij} w rozwiązaniu, stojące przy dodatniej potędze t są równe zeru. Wtedy w rozwiązaniu występują tylko wyrazy z tłumieniem.

$$\xi = C_{12}e^{-|\lambda_1|t}, \qquad \eta = C_{22}e^{-|\lambda_2|t} \tag{3.14}$$

Prowadzi to do powstania punktu siodłowego, dzielącego dwa baseny przyciągania, np. dwa działy wodne (prosta SW - NE na rysunku 3.1).

Z następnym przypadkiem mamy do czynienia gdy λ jest zespolone

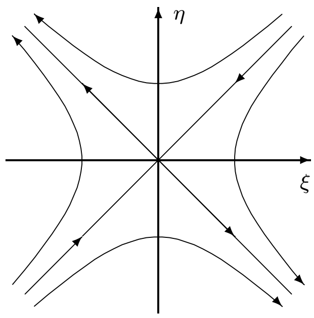
$$i\lambda_{1,2} = -\delta \pm i\omega. \tag{3.15}$$

Rozwiązanie ogólne ma więc postać

$$\xi(t) = e^{-\delta t}\left(C_1 \cos\omega t + C_2 \sin\omega t\right). \tag{3.16}$$

Typ zachowania układu będzie więc w zupełności określony poprzez wartość δ, ponieważ człon w nawiasie oscyluje.

- Jeżeli $\delta = 0$, to rozwiązanie będzie oscylujące - ruch układu odbywa się po elipsie, w której środku znajduje się punkt osobliwy, który nazywamy **centrum** lub **środkiem**.
- Jeżeli $\delta < 0$, to trajektoria w przestrzeni fazowej będzie rozwijającą się spiralą, a punkt osobliwy nazywamy **nietrwałym ogniskiem**.



RYSUNEK 3.1. Siodło, przypadek rzeczywistych pierwiastków o różnych znakach.

- Jeżeli $\delta > 0$, wówczas trajektorie fazowe dążą do centrum po zwijających się spiralach. Punkt osobliwy nazywamy **trwałym ogniskiem**.

Warto zauważyć, że bardzo niewielkie zmiany parametru δ, lecz odpowiadające zmianie znaku δ, prowadzą do drastycznej zmiany zachowania układu i jakościowych zmian (**bifurkacji**) portretu fazowego.

ROZDZIAŁ 4

Metoda Grupy Renormalizacyjnej

1. Hipoteza Skalowania

Hipoteza skalowania, zwana też hipotezą podobieństwa, została wprowadzona w 1966r. niezależnie przez L. Kadanoffa w USA oraz A. Pataszyńskiego i W. Pokrowskiego w Rosji. Pozwala ona na otrzymanie, posługując się ogólnymi argumentami, związków między wykładnikami krytycznymi. Stanowi też punkt wyjścia dla rozważań teorii grupy renormalizacyjnej. Ponieważ do tej pory założenia tego podejścia nie zostały ściśle udowodnione, mówimy o hipotezie, jednak uważa się, że założenia te są w większości fizycznie interesujących przypadków słuszne.

✠

Mówimy, że funkcja $f(x)$ jest **jednorodna w sensie Eulera** rzędu p jeżeli

$$f(\lambda x) = \lambda^p f(x) \quad \text{dla} \quad \forall\lambda. \tag{4.1}$$

Oznacza to, że wartość w dowolnym punkcie możemy otrzymać znając wartość funkcji w wybranym punkcie x_0 ($x = \lambda x_0$) poprzez zmianę skali (mnożenie przez λ). Analogicznie określa się funkcje jednorodne większej liczby zmiennych

$$f(\lambda x, \lambda y \ldots) = \lambda^p f(x, y, \ldots). \tag{4.2}$$

Można też zdefiniować uogólnioną funkcję jednorodną

$$f(\lambda^a x, \lambda^b y,) = \lambda^p f(x, y) \quad \text{dla } \forall\lambda. \tag{4.3}$$

Funkcję jednorodną możemy też zapisać inaczej, wybierając $\lambda = 1/y$. Wtedy

$$f\left(\frac{x}{y}, 1\right) = y^{-p} f(x, y) \qquad\qquad f(x, y) = y^p f\left(\frac{x}{y}, 1\right) := y^p F\left(\frac{x}{y}\right) \tag{4.4}$$

Funkcję jednorodną możemy więc zapisać jako iloczyn funkcji od stosunku zmiennych i y^p. Jest to też warunek na to aby funkcja była jednorodna.

✠

Hipoteza podobieństwa mówi, że energia swobodna (Gibbsa w przypadku magnetyka) $G(T, H)$ jest uogólnioną funkcją jednorodną, tzn. mamy

$$G(\lambda^{a_t} t, \lambda^{a_H} H) = \lambda G(t, H) \quad \text{dla} \quad \forall\lambda \tag{4.5}$$

gdzie

$$t \equiv \frac{T - T_c}{T_c}. \tag{4.6}$$

Hipoteza skalowania zakłada w ten sposób, że układ fizyczny w otoczeniu PK jest niezmienniczy wobec zmiany skali. Z podobnymi spostrzeżeniami, choć dotyczącymi problemów natury geometrycznej, spotykamy się przy badaniu układów samopodobnych (fraktalnych). Mając sformułowaną hipotezę skalowania, spróbujmy znaleźć związki między wykładnikami krytycznymi.

W tym celu zróżniczkujmy względem H obie strony równania wyrażającego hipotezę skalowania

$$\lambda^{a_H} \frac{\partial G(\lambda^{a_t} t, \lambda^{a_H} H)}{\partial(\lambda^{a_H} H)} = \lambda \frac{\partial G(t, H)}{\partial H} \tag{4.7}$$

ale $\partial G/\partial H = -M$, więc

$$\lambda^{a_H} M(\lambda^{a_t} t, \lambda^{a_H} H) = \lambda M(t, H) \tag{4.8}$$

Mając związek dla magnetyzacji, weźmy najpierw $H = 0$. Otrzymamy wtedy β.

$$\lambda^{a_H} M(\lambda^{a_t} t, 0) = \lambda M(t, 0), \tag{4.9}$$

$$M(t, 0) = \lambda^{a_H - 1} M(\lambda^{a_t} t, 0). \tag{4.10}$$

Jako λ weźmy teraz $\lambda = (-1/t)^{1/a_t}$, co daje

$$M(t, 0) = (-t)^{\frac{1-a_H}{a_t}} M(-1, 0). \tag{4.11}$$

Z definicji wykładników krytycznych wiemy, że gdy $t \uparrow 0$ to $M(t, 0) \sim (-t)^{\beta}$, a więc

$$\frac{(1 - a_H)}{a_t} = \beta \tag{4.12}$$

Podobnie znajdujemy związek a_H i a_t z δ. Weźmy $t = 0$ oraz $\lambda = H^{-1/a_H}$, co da

$$\delta = \frac{a_H}{1 - a_H}. \tag{4.13}$$

Rozwiązując oba uzyskane równania dostaniemy

$$a_t = \frac{1}{\beta} \frac{1}{\delta + 1}, \qquad a_H = \delta \frac{1}{\delta + 1}. \tag{4.14}$$

Biorąc pochodne G względem H dwa razy, otrzymujemy (dla $H = 0$) podatność i kładąc $\lambda = (-t)^{-1/a_t}$ dostaniemy $\gamma = (2a_H - 1)/a_t$. Łącząc otrzymane wyrażenia dostajemy równość Widoma

$$\gamma = \beta(\delta - 1). \tag{4.15}$$

Różniczkując względem t otrzymamy równości Griffithsa i Rushbrooke'a:

$$\alpha + \beta(\delta + 1) = 2, \qquad \alpha + 2\beta + \gamma = 2. \tag{4.16}$$

Widać, że hipoteza skalowania zamiast nierówności daje równości, co podejrzewano już przed sformułowaniem hipotezy skalowania, oraz że wykładniki krytyczne dadzą się wyrazić poprzez dwa parametry a_t i a_H. Oznacza to, że niezależne są tylko dwa wykładniki - fakt ten czasami nazywa się *hiperskalowaniem*. Tak więc hipoteza skalowania nie daje wprawdzie wartości wykładników krytycznych, daje jednak wszystkie związki między nimi.

2. Konstrukcja Kadanoffa

Dotyczy ona zjawisk zachodzących w pobliżu PK. Jak wiemy, korelacje są tam bardzo daleko zasięgowe i dlatego to wszystko co dzieje się na mniejszych (w porównaniu z promieniem korelacji) odległościach, nie powinno być istotne. Możemy więc pominąć w rozważaniach fluktuacje krótkozasięgowe, co odpowiada opisowi gruboziarnistemu. Idea Kadanoffa polega na systematycznej eliminacji nieistotnych (w pobliżu PK) stopni swobody (zmiennych termodynamicznych).

Dla ustalenia uwagi rozpatrzmy siatkę o stałej $a = 1$ (dla uproszczenia zapisu), w węzłach której umieszczone są spiny Isingowskie $\{s\}$. Bezwymiarowe wielkości jakimi tu operujemy to oddziaływania $k := \beta J$, pole $h := \beta H$ oraz hamiltonian (energia)

$$\mathcal{H}\{s\} = -k \sum_{<i,j>} s_i s_j - h \sum_{i=1}^{N} s_i. \tag{4.17}$$

W okolicy T_c (zakładamy, że istnieje) wprowadzamy nową zmienną

$$t = \frac{T - T_c}{T_c} \tag{4.18}$$

i suma stanów

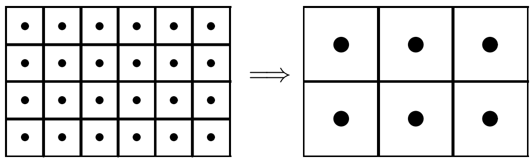
$$Z_N(h, t) = e^{-Ng(t,h)} = \sum_{\{s\}} e^{-\beta\mathcal{H}}. \tag{4.19}$$

Podzielmy teraz nasz układ na komórki o wymiarach L^d, gdzie $L \gg 1$, ale też $L \ll \xi$ (promienia korelacji), d jest wymiarem przestrzennym. Ponieważ w okolicy T_c korelacje są duże, więc w ramach jednej komórki spiny są silnie skorelowane i zachowują się podobnie. Możemy je więc zastąpić przez jeden wypadkowy spin s', również Isingowski. Istnieje wiele sposobów wyboru tego spinu, m.in. tzw *reguła większości* zaproponowana przez Niemeijera i Van Leeuwena. Zgodnie z tą regułą na siatkach w których komórka ma nieparzystą liczbę węzłów, wypadkowy spin całej komórki jest taki jaka jest większość spinów w komórce. Jeżeli komórka ma np trzy węzły (siatka trójkątna) i w pewnej komórce dwa spiny (dla przypadku $S = 1/2$) są skierowane ku górze a jeden w dół, to wypadkowy spin komórki będzie skierowany ku górze.

Po podziale na bloki (komórki) liczba spinów (zmiennych) w układzie zmniejszy się o czynnik L^d

$$N' = NL^{-d} \tag{4.20}$$

wzrosła też L-krotnie stała sieci. Dokonaliśmy w ten sposób skalowania o czynnik L.



Rysunek 4.1. Skalowanie dla $L = 2$

Założenie Kadanoffa polega na tym, że w pobliżu PK opis fizyczny układu poprzez spiny wyjściowe i spiny bloków powinien prowadzić do tych samych wyników. Oznacza to, że powinny być sobie równe odpowiednie sumy stanów – $Z_N(h, T)$ zdefiniowana powyżej dla spinów s, oraz

$$Z_N(h', t') = \exp\left(-N'g(t', h')\right) = \sum_{\{s'\}} \exp\left(-\beta\mathcal{H}\{s'\}\right). \tag{4.21}$$

gdzie

$$\mathcal{H}\{s'\} = -k' \sum_{<i,j>} s'_i s'_j - h' \sum_{i=1}^{N'} s'_i. \tag{4.22}$$

Sumowanie odbywa się po blokach. Z założonej równości sum stanów wynika

$$\exp\left(-N'g(t', h')\right) = \exp\left(-Ng(t, h)\right) \quad \Rightarrow \quad g(t, h) = L^{-d}g(t', h'). \tag{4.23}$$

W opisie posługujemy się dwoma parametrami – t i h (lub t' i h'), które powinny być ze sobą powiązane, ponieważ jeżeli $h \to 0$, to i $h' \to 0$. Podobnie jest dla t. Oznacza to, że możemy przyjąć

$$h \sim h', \qquad t \sim t' \tag{4.24}$$

czyli

$$h' = L^x h \qquad \text{oraz} \qquad t' = L^y t. \tag{4.25}$$

gdzie x, y to wielkości na razie nie znane. Mamy więc

$$g(t, h) = L^{-d}g(L^y t, L^x h). \tag{4.26}$$

Gdyby L było dowolne, oznaczałoby to, że g jest uogólnioną funkcją jednorodną. Ale $1 \ll L \ll \xi$. Jeżeli jednak zignorujemy to ograniczenie i będziemy uważali g za uogólnioną funkcję jednorodną, to

$$Lg(t, h) = g(L^{y/d}t, L^{x/d}h) \tag{4.27}$$

co przy przyjęciu $y/d = a_t$ i $x/d = a_H$, daje hipotezę skalowania.

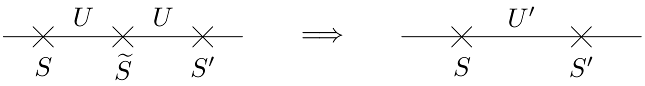
2.1. Przykład – 1D model Isinga. Jest to jedyny model dla którego można przeprowadzić w sposób ścisły konstrukcję blokową Kadanoffa. Niestety w 1D model Isinga nie ma przejścia fazowego. Suma stanów dana jest przez macierz przejścia

$$Z_N = \mathrm{Tr}\ P^N \tag{4.28}$$

gdzie

$$P = \begin{pmatrix} 1/uv & u \\ u & v/u \end{pmatrix}. \tag{4.29}$$

a $u = e^{-\beta J}$, $v = e^{-\beta H}$. Dla ferromagnetyka obie zmienne $u, v \in [0,1]$.



RYSUNEK 4.2. Skalowanie dla 1D modelu Isinga.

Wysumujmy po środkowym spinie $\tilde{s}$, a więc zwiększmy dwukrotnie stałą sieci ($L = 2$).

$$\begin{aligned}\langle s|P'|s'\rangle &= \langle s| \sum_{\tilde{s}=\pm 1} \exp\left(K(s\tilde{s} + \tilde{s}s') + h/2(s + 2\tilde{s} + s')\right)|s'\rangle \\ &= \langle s| \exp\left(K(s+s') + h/2(s+s'+2)\right) + \exp\left(-K(s+s') + h/2(s+s'-2)\right)|s'\rangle\end{aligned}$$

Daje to

$$P' \equiv P^2 = \begin{pmatrix} u^2 + (uv)^{-2} & v + v^{-1} \\ v + v^{-1} & u^2 + v^2u^{-2} \end{pmatrix}. \tag{4.30}$$

Żądamy teraz aby P' miało tę samą postać co P, tzn.

$$P' = C \begin{pmatrix} 1/(u'v') & u' \\ u' & v'/u' \end{pmatrix}. \tag{4.31}$$

Stałą C musimy dodać do rozważań, ponieważ przyrównanie obu wyrażeń dla P daje trzy równania (2 elementy są identyczne) dla u' i v'. Mamy więc następujące warunki

$$u'C = v + v^{-1} \tag{4.32}$$

$$\frac{C}{u'v'} = u^2 + (uv)^{-2} \tag{4.33}$$

$$\frac{Cv'}{u'} = u^2 + \left(\frac{u}{v}\right)^{-2}. \tag{4.34}$$

Rozwiązaniami są

$$u' = \frac{(v + 1/v)^{1/2}}{(u^4 + u^{-4} + v^2 + v^{-2})^{1/4}} \tag{4.35}$$

$$v' = \frac{(u^4 + v^2)^{1/2}}{(u^4 + v^{-2})^{1/2}} \tag{4.36}$$

$$C = \left(v + v^{-1}\right)^{1/2} \left(u^4 + u^{-4} + v^2 + v^{-2}\right)^{1/4}. \tag{4.37}$$

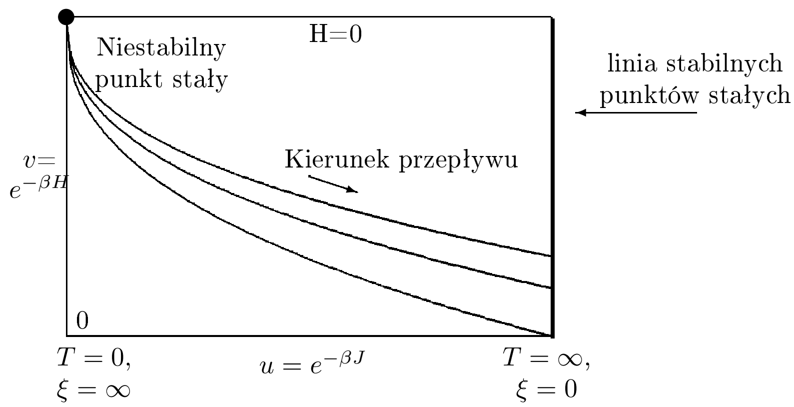
Równania powyższe jest to przepis na odwzorowanie $(u, v) \to (u', v')$, czyli transformacja blokowa Kadanoffa. Powtarzając tę transformację dokonujemy kolejnych przesunięć punktu wyjściowego (u_0, v_0) w przestrzeni parametrów. Otrzymujemy w ten sposób pewną trajektorię. Zbiór takich trajektorii nazywa się zwykle **diagramem przepływu** (**flow diagram**). Pokazuje on co dzieje się z układem, który opisywany jest przez parametry u, v, w trakcie kolejnych transformacji blokowych.

Ponieważ chcemy aby przekształcenia blokowe nie zmieniały fizyki, a więc w szczególności wielkości fluktuacji PP, muszą także zmieniać się pod wpływem tych przekształceń spiny układu:

$$s' = b^{-1}s \tag{4.38}$$

gdzie b zależy od hamiltonianu. Zmianę stałej oddziaływania $u \sim J$, pola zewnętrznego $v \sim H$, czy spinów, nazywamy **skalowaniem** tych wielkości. Powtarzając transformację Kadanoffa (dla rozpatrywanego 1D modelu Isinga) można przekonać się, że po pewnej liczbie przekształceń parametry u, v przestaną się zmieniać. Mówimy, że osiągnęliśmy **punkt stały** transformacji blokowej. Dla naszego przypadku istnieją dwa punkty stałe, tzn dwie pary wartości u, v które pozostają niezmiennicze względem tej transformacji. Są to $u = 0$ i $v = 1$ co odpowiada nieskończenie silnemu oddziaływaniu i zerowemu polu oraz $u = 1$, dowolne v czyli brak oddziaływań i dowolne pole.

Punkt $(0, 1)$ jest niestabilny, tzn. startując z dowolnego punktu w jego otoczeniu, kolejne transformacje będą nas od niego oddalały. Punkty na osi $u = 1$ są stałe. Diagram przepływu ma postać przedstawioną na rysunku 4.3.



RYSUNEK 4.3. Diagram przepływu dla 1D modelu Isinga.

W punktach stałych funkcja korelacyjna jest niezmiennicza względem transformacji blokowej, tzn może być albo $\xi = 0$ albo $\xi = \infty$. Jeżeli interesuje nas PK, to oczywiście ξ powinno być nieskończone. Dla 1D modelu Isinga odpowiada temu punkt $(0, 1)$, który jest jednak punktem niestabilnym, a więc nieosiągalnym, co wyraża fakt, że w modelu tym nie ma przejścia fazowego w skończonych temperaturach. W przypadku ogólnym szukać będziemy takich stabilnych punktów stałych dla których $\xi = \infty$, co będzie oznaczało, że parametry układu osiągają wartości krytyczne.

Stabilny punkt stały na diagramie przepływu ma tę własność, że wszystkie układy znajdujące się w jego sąsiedztwie, będą z biegiem czasu dążyły do tego punktu. Możemy punkt ten nazwać **atraktorem**, natomiast obszar jego 'wpływu'– jego **basenem przyciągania**. W układach deterministycznych łatwo sobie można wyobrazić dwa typy atraktorów – **stałe**, których przykładem może być położenie spoczynkowe wahadła, oraz **atraktor periodyczny**, np wahadło w zegarze napędzanym ciężarkami. W przestrzeni fazowej ruch taki ma postać okręgu. Jak się okazało, możliwy jest jeszcze jeden typ atraktora – **dziwny atraktor**, prowadzący do ruchu chaotycznego.

3. Metoda Grupy Renormalizacyjnej

Metoda grupy renormalizacyjnej (GR) oparta jest na filozofii transformacji blokowych Kadanoffa, eliminacji stopni swobody poprzez sumowanie, bądź całkowanie, po zmiennych które uważamy za nieistotne dla rozpatrywanego zagadnienia. Jak widzieliśmy, w okolicy PK istotną rolę grają tylko korelacje na dużych odległościach, można więc wysumować po małych obszarach.

Rozważmy teraz d-wymiarową siatkę. Położenia węzłów wyznaczane są przez wektor $\vec{x} = (x_i, i = 1, 2, ..., d)$. W węzłach sieci umieszczamy spiny S_x o n składowych $(S_x^k, k = 1, 2, ..., n)$.

$n = 1$ odpowiada modelowi Isinga o spinie $\frac{1}{2}$, $n = 2$ modelowi XY , $n = 3$ modelowi Heisenberga. Możemy jednak rozważać również ciągłe wartości zarówno wymiaru przestrzennego jak i spinu. Jeżeli przejdziemy z n do nieskończoności otrzymamy model sferyczny Berlina i Kaca. Dla $n = -2$ mamy model gaussowski. Jeżeli ograniczyć się do układów izotropowych i krótko-zasięgowych, wówczas wpływ na wykładniki krytyczne mają tylko dwie wielkości – n i d.

Oznaczmy wyjściowy hamiltonian przez $\mathcal{H}$. Dokonując na nim (na jego parametrach) transformacji Kadanoffa otrzymamy nowy hamiltonian $\mathcal{H}'$ o mniejszej liczbie stopni swobody

(1) $\mathcal{H} \quad \Rightarrow \quad \mathcal{H}' = R[\mathcal{H}], \qquad \mathcal{H}'$ jest nadal translacyjnie niezmiennicze.

(2) Operator R nazywamy operatorem GR. Operator ten zmniejsza liczbę stopni swobody układu z N do $N' = N/b^d$, gdzie b jest tzw. czynnikiem skalowania przestrzennego. Działanie operatora R prowadzi więc do zmniejszenia liczby stopni swobody (zmiennych spinowych) o $(N - N')$. Ponieważ krótko-zasięgowe fluktuacje są nieistotne, możemy wycałkować po dużych wektorach falowych, czyli wziąć ślad po transformatach spinów zależących od takich wektorów.

$$e^{\mathcal{H}'} = \mathrm{Tr}\, e^{\mathcal{H}}. \tag{4.39}$$

(3) Ponieważ transformacja R nie może zmieniać fizyki, suma stanów musi pozostać bez zmian, czyli

$$Z_{N'}[\mathcal{H}'] = Z_N[\mathcal{H}]. \tag{4.40}$$

(4) Stała musi też być przestrzenna gęstość spinów, czyli wektory sieci prostej muszą się skalować

$$x \quad \Rightarrow \quad x' = x/b, \tag{4.41}$$

co oznacza, że pędy skalują się odwrotnie

$$p \quad \Rightarrow \quad p' = pb. \tag{4.42}$$

(5) Same spiny też muszą się skalować

$$S_x \quad \Rightarrow \quad S'_x = S_x/c \tag{4.43}$$

przy czym c, zależy od hamiltonianu, $c = c[\mathcal{H}]$.

Z własności (1) i (2) wynika, że gęstość energii swobodnej spełnia równanie

$$g[\mathcal{H}'] = b^d g[\mathcal{H}]. \tag{4.44}$$

Dla dwu-spinowej funkcji korelacyjnej mamy

$$G[x; \mathcal{H}] = c^2 G[x/b; \mathcal{H}']. \tag{4.45}$$

Procedura GR jest następująca: Operacja R stosowana jest wielokrotnie

$$\mathcal{H}' = R[\mathcal{H}], \quad \mathcal{H}'' = R[\mathcal{H}']... \tag{4.46}$$

tak długo aż, dla danych parametrów wyjściowego hamiltonianu znajdziemy punkt stały transformacji, tzn. taką wartość $\mathcal{H}^*$, że

$$R[\mathcal{H}^*] = \mathcal{H}^*. \tag{4.47}$$

Hamiltonian $\mathcal{H}^*$ jest krytyczny, tzn. określające go parametry przyjmują wartość krytyczne – odpowiada mu nieskończenie długo zasięgowa funkcja korelacyjna. Mamy bowiem z równania na punkt stały i przepisu skalowania dla wektorów przestrzennych (a więc i dla promienia korelacji ξ)

$$\xi^* = \xi^*/b \quad \Rightarrow \quad a) \qquad \xi^* = 0, \quad b) \qquad \xi^* = \infty. \tag{4.48}$$

Przypadek $\xi = \xi^* = 0$ jest trywialny – odpowiada mu zupełny brak korelacji, czyli pozostaje możliwość b) świadcząca o zachowaniu krytycznym układu.

3.1. Wykładniki krytyczne w metodzie GR. Aby znaleźć wykładniki krytyczne zlinearyzujmy operator R w działaniu wokół $\mathcal{H}^*$:

$$\mathcal{H}' = R[\mathcal{H}] = R[\mathcal{H}^* + hQ] = \mathcal{H}^* + hLQ + \mathcal{O}(h^2). \tag{4.49}$$

✠

$$f(x) = \mathcal{O}\left(\phi(x)\right) \quad \text{jeżeli} \quad \lim_{x\to\infty} \frac{f(x)}{\phi(x)} = const \tag{4.50}$$

✠

h są to współczynniki, natomiast Q to funkcje własne (operatorowe) zlinearyzowanego operatora L

$$LQ_j = \Lambda_j Q_j. \tag{4.51}$$

Λ_j to wartości własne, które podobnie jak L, zależą od b (czynnika skalowania przestrzennego). Ponieważ R jest operatorem półgrupy, więc Λ_j powinny być postaci

$$\Lambda_j = b^{\lambda_j}, \tag{4.52}$$

przy czym λ_j nie zależą od wyboru czynnika skalującego b. Λ_j dają spektrum wartości krytycznych. Λ_1 zwykle identyfikuje się z gęstością energii, Λ_2 z PP. Związane z tym wykładniki λ_1 i $\lambda_2 > 0$. Przyjmując, że Q_j tworzą układ zupełny (co jest ścisłe tylko asymptotycznie), rozkładamy hamiltonian $\mathcal{H}$ w pobliżu $\mathcal{H}^*$:

$$\mathcal{H} = \mathcal{H}^* + \sum_j h_j Q_j. \tag{4.53}$$

Wielkości h_j nazywamy liniowymi polami skalującymi. Wstawiając taką postać hamiltonianu do równania GR otrzymujemy

$$\begin{aligned}\mathcal{H}' = R[\mathcal{H}] = \mathcal{H}^* + \sum h_j LQ_j + \mathcal{O}(h^2) = \mathcal{H}^* + \sum h_j \Lambda_j Q_j + \mathcal{O}(h^2)\\ = \mathcal{H}^* + \sum h_j' Q_j + \mathcal{O}(h^2),\end{aligned} \tag{4.54}$$

czyli

$$h_j' = \Lambda_j h_j = b^{\lambda_j} h_j. \tag{4.55}$$

Ponieważ dla energii swobodnych mieliśmy

$$g[\mathcal{H}'] = b^d g[\mathcal{H}] \tag{4.56}$$

więc

$$g(h_1', h_2', \ldots) = b^d g(h_1, h_2, \ldots) \tag{4.57}$$

$$g(h_1, h_2, \ldots) = b^{-d} g(h_1 b^{\lambda_1}, h_2 b^{\lambda_2}, \ldots). \tag{4.58}$$

Jest to związek dla uogólnionych funkcyj jednorodnych. Aby więc otrzymać energię swobodną w postaci funkcji jednorodnej, należało dokonać linearyzacji operatora GR. Jest to, jak widać, warunek otrzymania punktu wyjściowego hipotezy skalowania.

Postarajmy się teraz znaleźć wykładniki krytyczne. Ponieważ operator Q_1 związany jest z energią, a Q_2 z magnetyzacją, wybierzmy $h_1 = k_1 t$ i $h_2 = k_2 H$, gdzie t i H to zredukowana temperatura i zewnętrzne pole. Równanie dla energii swobodnej możemy przepisać w postaci

$$bg(h_1, h_2, \ldots) \sim g(h_1 b^{\lambda_1/d}, h_2 b^{\lambda_2/d}, \ldots) = g(k_1 b^{\lambda_1/d} t, k_2 b^{\lambda_2/d} H, \ldots). \tag{4.59}$$

Porównajmy to z analogicznym wzorem otrzymanym przy pomocy hipotezy skalowania

$$Lg(t, H) = g(L^{y/d} t, L^{x/d} H). \tag{4.60}$$

Widać stąd, że ($L \equiv b$)

$$\lambda_1 = y, \quad \lambda_2 = x \tag{4.61}$$

a ponieważ uprzednio pokazaliśmy, że

$$d/y = 2 - \alpha, \quad (d-x)/y = \beta, \tag{4.62}$$

więc

$$\lambda_1 = y = \frac{d}{2-\alpha} = \frac{1}{\nu}, \quad \lambda_2 = x = d\left(\frac{2-\alpha-\beta}{2-\alpha}\right), \tag{4.63}$$

natomiast

$$\lambda_2/\lambda_1 = 2 - \alpha - \beta = 1/2(\gamma + 2 - \alpha) := \Delta. \tag{4.64}$$

Wielkość Δ nosi nazwę wykładnika szczelinowego.
Tak więc przyjmując, że istnieje punkt stały transformacji R dla GR, dany przez $\mathcal{H}^*$, wokół którego możemy zlinearyzować R, otrzymujemy skalowanie dla energii swobodnej. W naturalny sposób, jako wartości własne zlinearyzowanej GR, pojawiają się wykładniki krytyczne. Metoda GR pozwoliła więc na pokazanie, że wykładniki krytyczne nie są sztuczną konstrukcją, lecz należą w sposób integralny do teorii przejść fazowych. Podobnie równanie Schrödingera tłumaczy istnienie linii widmowych.
Również w sposób naturalny otrzymujemy podział na klasy uniwersalności. Widzieliśmy już w trakcie omawiania podstaw GR, że wykładniki krytyczne nie zależały od szczegółów hamiltonianu. Wartości wykładników oraz charakter przejścia fazowego określane są przez punkt stały. Do jednej klasy uniwersalności należeć więc będą wszystkie układy, których hamiltoniany zbiegać się będą do tego samego punktu stałego. Możemy podzielić operatorowe funkcje własne R na trzy rodzaje:

(1) $\Lambda_j > 1$, lub $\lambda_j > 0$. Startujemy z hamiltonianu bliskiego pewnemu punktowi stałemu i dokonujemy kolejnych transformacji GR. Wówczas $h_j Q_j$ rośnie podczas iteracji GR i wyprowadza hamiltonian do innego punktu stałego, charakteryzowanego przez inny zestaw λ_j. Mamy wtedy do czynienia ze zjawiskiem przejścia układu z jednej klasy uniwersalności do innej (*crossover phenomena*). Operatory Q_j różne od $Q_1 \sim E$ i $Q_2 \sim M$, nazywamy *istotnymi*. Dla takich operatorów wprowadza się zwykle *wykładnik przejścia* $\phi_j = \lambda_j/\lambda_1$. Przykładem operatora istotnego może być dodanie do izotropowego hamiltonianu Heisenberga członu anizotropowego.

(2) $\Lambda_j < 1$ lub $\lambda_j < 0$. W takim przypadku $h_j Q_j$ znika przy iteracjach i dodanie do hamiltonianu wyrazu zawierającego Q_j nie zmienia punktu stałego hamiltonianu. Operator taki jest więc *nieistotny*. Może on mieć co najwyżej wpływ na poprawki do czystego prawa skalowania.

(3) $\Lambda_j = 0$ lub $\lambda_j = 0$. Operator taki nazwiemy *marginalnym*. $h_j Q_j$ jest wolno zmienne. Może istnieć linia punktów stałych wzdłuż której zmienia się uniwersalność. Wykładnik przejścia jest równy zeru a pozostałe wykładniki krytyczne zmieniają się w sposób ciągły wraz ze zmianą pola f_i w hamiltonianie zawierającego operator marginalny.

ROZDZIAŁ 5

Układy z domieszkami

Rozpatrywać tu będziemy głównie układy magnetyczne, aczkolwiek niejednorodność może oczywiście dotyczyć także innych zjawisk. Niejednorodność w układach magnetycznych może być spowodowana przez:

(1) nieregularności sieci (magnetyk amorficzny)
(2) nieidentyczność momentów magnetycznych jonów tworzących układ. W tym ostatnim przypadku powodem może być to, że
   - część atomów magnetycznie czynnych wymieniona jest na inne o różnych momentach magnetycznych
   - atomy domieszek pozbawione są w ogóle momentów magnetycznych – mówimy wówczas o **rozrzedzonym** magnetyku
(3) fakt, że wiązania pomiędzy magnetycznie czynnymi atomami nie są jednakowe – np. część z nich odbywa się poprzez inne atomy.

Przypadkiem (1) nie będziemy się dalej zajmowali. Jest on zresztą najtrudniejszy i wymaga innego podejścia.
W przypadku (2) mówimy że jest to **problem węzłów** – **site problem**. Przykładem mogą być stopy ferromagnetyczne, takie jak Fe_xNi_{1-x}. Hamiltonian takiego układu ma postać

$$\mathcal{H}_s = -\sum_{i,j} J_{ij} S_i \xi_i S_j \xi_j, \tag{5.1}$$

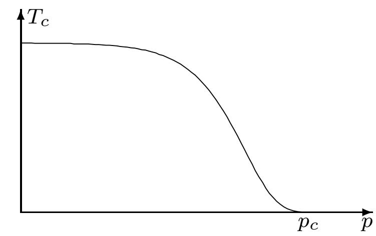
W przypadku stopu podwójnego $\xi_i = 1$ odpowiada składnikowi A (Fe), natomiast $\xi_i = -1$ – składnikowi B (Ni). Dla układu rozrzedzonego $\xi_i = 0, 1$.
W przypadku (3) mówimy o **zagadnieniu wiązań** – **bond problem**. Hamiltonian ma postać

$$\mathcal{H}_b = -\sum_{i,j} J_{ij} \xi_{ij} S_i S_j \tag{5.2}$$

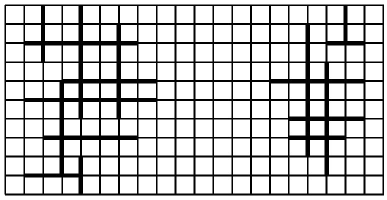
W przypadku nieporządku wiązań atomy w węzłach są takie same, tylko pomiędzy węzły dostają się atomy domieszek, modyfikując w ten sposób oddziaływania. Przykładem może tu być oddziaływanie atomów tlenu w warstwie $Cu - O$ nadprzewodnika $YBa_2Cu_3O_{6+c}$. Bezpośrednie oddziaływanie między tlenami jest odpychające, natomiast poprzez atomy miedzi jest przyciągające. Tak więc wartości współczynników ξ_i są teraz równe ± 1, natomiast wartości „spinów” to $0, 1$, ponieważ dany węzeł może albo być obsadzony przez atom tlenu, albo nie$>$ Dlatego też zmienne S_i mają charakter zmiennych obsadzeniowych. Możemy oczekiwać, że jeżeli będziemy „rozrzedzali” układ magnetyczny poprzez dodawanie niemagnetycznych domieszek, wówczas temperatura krytyczna w funkcji koncentracji zajętych węzłów będzie dążyła do zera. Może istnieć pewna koncentracja krytyczna, p_c, poniżej której $T_c = 0$. Odpowiedni wykres magnetyzacji ma postać jak na rysunku (5.1).

Z zagadnieniem rozrzedzonego magnetyka związany jest problem **perkolacji**.
Niech w krysztale pewna część węzłów (albo, co jest równoważne – plakietek) będzie zajęta, a część pusta. Ograniczmy oddziaływanie między spinami do najbliższego sąsiedztwa. Mówimy, że dwa spiny (węzły) należą do tego samego grona (klastra) jeżeli istnieje łańcuch zajętych węzłów łączących te spiny.

Rozmiar grona zależy oczywiście od koncentracji p zajętych węzłów i rośnie ze wzrostem p.



RYSUNEK 5.1. Zależność temperatury przejścia od koncentracji domieszek.

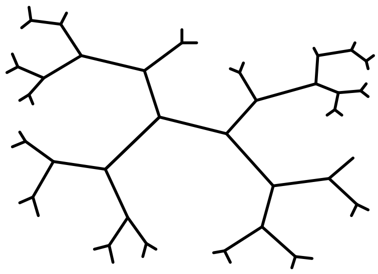


RYSUNEK 5.2. Po lewej układ powyżej, a po prawej poniżej progu perkolacji.

Mówimy, że w układzie wystąpiło zjawisko perkolacji jeżeli dwie przeciwległe krawędzie układu (np. górna i dolna) są połączone nieprzerwaną liną zajętych węzłów lub plakietek. Najniższą wartość koncentracji, p_c, nazywamy *koncentracją* lub **progiem perkolacji**. Dla układu nieskończonego oznacza to, że powyżej progu perkolacji mamy już nieskończone grona połaczonych spinów. Można też rozważać zagadnienie perkolacji, gdy najmniejszy zajęty obszar nie jest elementarną plakietką, lecz jest np. prostokątem złożonym z kilku plakietek. W ferromagnetyźmie poniżej progu perkolacji, a więc dla $p < p_c$ mogą istnieć tylko skończone grona – nie ma więc spontanicznej magnetyzacji i $T_c = 0$. Zagadnienia perkolacji można też przedstawić w postaci problemu wiązań i problemu węzłów. Przejście przewodnik – izolator da się opisać jako problem wiązań. Droga pomiędzy dwoma atomami może być albo otwarta albo zamknięta dla elektronu przewodnictwa. Jeżeli prawdopodobieństwo p otwartej drogi jest mniejsze od p_c, elektrony mogą poruszać się tylko w ograniczonym obszarze i nie ma przewodnictwa. Kryształ staje się przewodnikiem dla $p \geq p_c$. p_c można też np. interpretować jako próg kiedy zaraza staje się epidemią. Istnieje bardzo wiele możliwych interpretacji, również poza fizyką, np. w ekonomii czy naukach społecznych, zjawiska perkolacji. Sam problem pojawił się w 1957r. przy opracowywaniu modelu filtra pochłaniającego pył węglowy. Przykładowe wartości koncentracji krytycznych to (p_c^{bond} i p_c^{site}):

	p_c^{bond}	p_c^{site}
łańcuch	1	1
siatka kwadratowa	0.5	0.59
siatka trójkątna	0.3473	0.5
siatka prosta kubiczna	0.247	0.31
siatka FCC	0.119	0.195

Jak widać, próg perkolacji maleje wraz ze stopniem „komplikacji” siatki, wyrażającym się m.in. rosnącą liczbą najbliższych sąsiadów, stwarza to bowiem możliwości nowych połączeń. Kryterium liczby koordynacyjnej siatki (liczby najbliższych sąsiadów) nie jest jednak jedynym, jak widać z porównania wyników dla siatki trójkątnej i prostej kubicznej, które obie mają liczbę koordynacyjną równą 6. Przy rozwiązywaniu zagadnień perkolacji korzysta się często z metod graficznych, przy czym ścisłe wyniki udaje się otrzymać najczęściej na tzw. **siatkach Bethego**. Są to siatki bez pętli, tzn. wychodząc z pewnego punktu nie możemy już do niego powrócić. Na siatkach Bethego problemy węzłów i wiązań dają te same wyniki. Z problemem perkolacji związany jest też problem **zapchania**, (ang. *jamming*). Tu również na siatce umieszczane są w sposób losowy obiekty, tym



Rysunek 5.3. Przykład siatki Bethego o liczbie koordynacyjnej $\sigma = 3$ (w każdym węźle spotykają się trzy drogi – jedna wchodząca i dwie wychodzące)

razem jednak aż do chwili, kiedy nie da się już umieścić następnego obiektu. Liczba umieszczonych tak obiektów determinuje **próg zapchania**. Dotąd nie uwzględnialiśmy sposobu w jaki rozmieszczone są domieszki. Mogą one brać udział w tworzeniu równowagi termodynamicznej układu jako całości – mówimy wtedy o układzie (opisie) **annealed** – **wyżarzonym**, lub też mogą one być rozmieszczone przypadkowo – mamy wówczas układ (opis) **quenched** – **schłodzony**. Pierwszy układ możemy wyobrazić sobie jako podgrzany do bardzo wysokiej temperatury, a następnie powoli ostudzany, tak aby domieszki mogły znaleźć położenia zgodne z równowagą termodynamiczną dla całości układu. Przy drugim podejściu ochłodzenie jest tak raptowne, że domieszki „zastygają" w swych początkowych (losowych) położeniach. Dla układu A (annealed) energia swobodna będzie więc miała postać

$$F_A = -kT \ln \langle \langle Z \rangle_s \rangle_c \tag{5.3}$$

gdzie $\langle ... \rangle_s$ oznacza uśrednianie po spinach (wzięcie śladu), natomiast $\langle ... \rangle_c$ – uśrednianie po rozkładzie domieszek. Oznacza to, że sumę stanów Z_A liczymy jako średnią po konfiguracjach domieszek, które w procesie tworzenia równowagi termodynamicznej grają taką samą rolą jak spiny

$$Z_A \sim \sum_{\xi} Z(\xi) \tag{5.4}$$

ξ jest tu ustaloną konfiguracją domieszek.Dla układu quenched (Q) znajdujemy energię swobodną licząc energię swobodną stanów dla zadanej konfiguracji, a dopiero później uśredniając po możliwych rozkładach domieszek

$$F_Q = (N^*)^{-1} \sum_{\xi} F(\xi) \; = \; \langle F(\xi) \rangle_c \tag{5.5}$$

gdzie $N^* = \sum_{\xi} 1$ jest całkowitą liczbą konfiguracji wiązań. Mamy więc ostatecznie następujące możliwości – układ wiązań (Q lub A) oraz układ węzłów (Q lub A). Ścisłych twierdzeń o związkach pomiędzy poszczególnymi podejściami (można je traktować jako różne przybliżenia) jest niewiele i dotyczą one głównie albo układów 1D albo na siatce Bethego. Najbardziej widoczna różnica w zachowaniu układów Q i A polega na tym ,że „site A" nie wykazują zjawiska perkolacji ponieważ warunek stabilności (minimum energii Gibbsa) wymaga aby układ miał maksimum liczby wiązań. Prowadzi to do powstawania agregatów – gdy tylko $p \neq 0$ pojawia się nieskończone grono. W przypadku nieporządku wiązań najczęściej stosowanym rozkładem jest (np. gdy mamy dwie możliwości – $+J_1$ i $-J_1$) rozkład gaussowski o zerowej szerokości

$$P(J) = p\delta(J - J_1) + (1 - p)\delta(J + J_1) \tag{5.6}$$

co dla przypadku rozrzedzenia ma postać

$$P(J) = p\delta(J - J_0) + (1 - p)\delta(J). \tag{5.7}$$

Oczywiście możliwe są także bardziej ogólne rozkłady – gaussowski o skończonej szerokości lub prostokątny. Rozpatrzmy teraz 1D łańcuch Isingowski

$$\mathcal{H}_N = -\sum J_{i+1}\sigma_i\sigma_{i+1}, \tag{5.8}$$

gdzie J_{i+1} jest całką wymiany, której wartość (a być może także i znak) zależy od funkcji gęstości prawdopodobieństwa $P(J)$. Suma stanów ma postać

$$Z_N = \prod_{i=1}^{N} 2\cosh\beta J_i \tag{5.9}$$

Przy podejściu Q energia swobodna f na cząstkę jest równa

$$f_Q = \lim_{N\to\infty} N^{-1}F_N^Q = -kT\int \ln(2\cosh\beta J)P(J)dJ \tag{5.10}$$

Wstawiając odpowiedni typ rozkładu $P(J)$ otrzymujemy wielkości termodynamiczne (C_H, U, itd.).

W podejściu A

$$f_A = \lim_{N\to\infty} N^{-1}F_N^A = -kT\ln\int 2\cosh(\beta J)P(J)dJ. \tag{5.11}$$

Jak widać, w przypadku quenched uśrednia się poza logarytmem, natomiast w przypadku annealed - pod logarytmem, co odzwierciedla fakt różnego traktowania zmiennych obsadzeniowych w obu podejściach. Układy typu A są, naogół, łatwiejsze do rozwiązania i w pewnych przypadkach można je przeprowadzić w odpowiednie układy „czyste" i w ten sposób rozwiązać je ściśle – o ile oczywiście znane są rozwiązania odpowiednich układów "czystych".

1. Przykład – rozrzedzony model Isinga

Zastosujemy podejście typu bond-annealed. Hamiltonian ma postać

$$\mathcal{H} = -J\sum \eta_{ij}\sigma_i\sigma_j \tag{5.12}$$

η_{ij} jest zmienną obsadzeniową, tzn. $\eta = 0,1$, natomiast $\sigma = \pm 1$. Ponieważ przyjęliśmy podejście annealed, domieszki biorą udział w tworzeniu równowagi termodynamicznej, podobnie jak spinowe stopnie swobody. Mamy więc dwa typy zmiennych – spinowe i obsadzeniowe η_{ij}. Te ostatnie wprowadzają do danego modelu zmienną liczbę wiązań – powinniśmy więc korzystać z wielkiego rozkładu kanonicznego, gdzie potencjał chemiczny μ związany jest ze zmienną η. Aby znaleźć średnie namagnesowanie dla magnetyka jednorodnego, opisanego hamiltonianem

$$\mathcal{H}_j = J\sum S_iS_j + H\sum_i S_i \tag{5.13}$$

wystarczyło policzyć pochodną

$$\langle m\rangle = \frac{1}{N}\frac{\partial \ln Z_j}{\partial H} \tag{5.14}$$

Z_j oznacza tu sumę stanów dla modelu jednorodnego. Podobnie powinniśmy postąpić i dla zmiennych obsadzeniowych. Odpowiednia suma stanów ma postać (jak dla wielkiego rozkładu kanonicznego)

$$Z = \mathrm{Tr}_\sigma\ \mathrm{Tr}_\eta\ e^{-\beta\mathcal{H}+\beta\mu\sum\eta_{ij}} \tag{5.15}$$

Chcąc znaleźć gęstość wiązań p, policzymy pochodną z sumy stanów

$$p = \frac{1}{N_w}\frac{\partial \ln Z}{\partial \beta\mu} = \frac{1}{N_w}\langle\sum\eta_{ij}\rangle \tag{5.16}$$

gdzie N_w jest liczbą wiązań. Chcemy teraz sprowadzić nasz układ z domieszkami do układu „czystego", opisywanego poprzez parametry K, A, które będziemy wyznaczać poprzez wyjściowe parametery J, μ (lub ewentualnie gęstość wiązań).

Zażądajmy teraz aby wzięcie częściowego śladu (po zmiennych losowych) sprowadzało nasz układ z domieszkami do układu czystego, tzn. dla pojedynczego wiązania

$$\mathrm{Tr}_\eta\ \exp\left(\beta J\eta\sigma\sigma\prime + \beta\mu\eta\right) = \exp\left(\beta K\sigma\sigma\prime + A\right) \tag{5.17}$$

K i A to stałe, które należy wyznaczyć. Wykonując ślad dostajemy

$$\exp\left(\beta J\sigma\sigma\prime + \beta\mu\right) + 1 = \exp\left(\beta K\sigma\sigma\prime + A\right), \tag{5.18}$$

a stąd przypisując poszczególne wartości zmiennym σ

$$\exp(\beta J + \beta\mu) + 1 = \exp(\beta K + A),$$
$$\exp(-\beta J + \beta\mu) + 1 = \exp(-\beta K + A)$$

Daje to

$$e^{2\beta K} = \frac{\exp(\beta J + \beta\mu) + 1}{\exp(-\beta J + \beta\mu) + 1}, \qquad e^{2A} = \left(e^{\beta J + \beta\mu} + 1\right)\left(e^{-\beta J + \beta\mu} + 1\right) \tag{5.19}$$

Koncentrację istniejących wiązań, p, już znaleźliśmy. Ale suma stanów Z_j dla równoważnego modelu czystego (o stałej wiązania K) powinna być równa Z, a więc

$$p = \frac{1}{N_w}\frac{\partial \ln Z_j}{\partial \beta\mu}, \qquad Z_j = \mathrm{Tr}_\sigma \exp\left(\beta K \sum \sigma_i\sigma_j + \frac{Nz}{2}A\right) \tag{5.20}$$

$Nz/2$ jest to całkowita liczba wiązań w układzie, a więc

$$p = \frac{\langle \sum \eta_{ij}\rangle}{Nz/2} = \frac{2}{Nz}\frac{\partial \ln Z}{\partial \beta\mu} = \frac{2}{Nz}\frac{\partial \ln Z_j}{\partial \beta\mu} =$$
$$= \frac{2}{Nz}\left(\langle \sum \sigma\sigma\prime\rangle \frac{\partial \beta K}{\partial \beta\mu} + \frac{Nz}{2}\frac{\partial A}{\partial \beta\mu}\right) = \varepsilon\frac{\partial \beta K}{\partial \beta\mu} + \frac{\partial A}{\partial \beta\mu} \tag{5.21}$$

ε jest funkcją korelacyjną (dwuspinową jeżeli, tak jak tutaj, rozpatrujemy oddziaływanie do najbliższych sasiadów) czystego modelu Isinga. Nasz wzór zawiera nieznany parametr - potencjał chemiczny. Możemy się go pozbyć

$$e^{2\beta K} = \frac{\exp(\beta J + \beta\mu) + 1}{\exp(-\beta J + \beta\mu) + 1} \quad \Rightarrow \quad e^{\beta\mu} = \frac{1 - e^{2\beta K}}{e^{\beta J}(e^{2\beta K - 2\beta J} - 1)} \tag{5.22}$$

Pochodne występujące we wzorze dla p są równe

$$\frac{\partial A}{\partial \beta\mu} = \frac{1}{2}\left(\frac{\exp(\beta J + \beta\mu)}{\exp(\beta J + \beta\mu) + 1} + \frac{\exp(-\beta J + \beta\mu)}{\exp(-\beta J + \beta\mu) + 1}\right) \tag{5.23}$$
$$\frac{\partial \beta K}{\partial \beta\mu} = \frac{1}{2}\left(\frac{\exp(\beta J + \beta\mu)}{\exp(\beta J + \beta\mu) + 1} - \frac{\exp(-\beta J + \beta\mu)}{\exp(-\beta J + \beta\mu) + 1}\right) \tag{5.24}$$

Podstawiając tutaj znalezioną wartość potencjału chemicznego ($e^{\beta\mu}$), dostajemy

$$2p = \frac{\exp 2\beta K - 1}{\exp 2\beta J - 1}\left[\frac{\varepsilon + 1}{\exp(2\beta K - 2J\beta)} + 1 - \varepsilon\right], \tag{5.25}$$

co pozwala napisać

$$e^{2\beta J} = \frac{2p - \left(e^{2\beta K} - 1\right)(1 - \varepsilon)}{2p - (1 - e^{-2\beta K})(1 + \varepsilon)}. \tag{5.26}$$

Aby wyznaczyć zależność T_c wyjściowego modelu (o stałej sprzężenia J) od koncentracji wiązań magnetycznych p, należy podstawić zamiast ε i βK ich wartości krytyczne dla modelu „czystego". Dla siatki kwadratowej wynoszą one

$$e^{2\beta_c K} = \sqrt{2} + 1 \quad \text{i} \quad \varepsilon_c = 1/2. \tag{5.27}$$

Otrzymujemy w ten sposób zależność temperatury krytycznej od koncentracji

$$e^{2\beta_c J} = \frac{2p + \sqrt{2} - 1}{2p - 1}. \tag{5.28}$$

Widać stąd, że koncentracja krytyczna, przy której $T_c(p) = 0$, a więc $\beta_c J \to \infty$, osiągnięta zostanie przy

$$p = p_c = 0.5. \tag{5.29}$$

Wyznaczyliśmy więc w sposób ścisły krytyczną koncentrację dla modelu Isinga na siatce kwadratowej w przybliżeniu bond-annealed. Wartości $\beta_c K$ i ε znane są też dla siatki trójkątnej $e^{2\beta_c K} = \sqrt{3}$ oraz $\varepsilon = 2/3$, co daje

$$p_c^{tr} = 0.3522. \tag{5.30}$$

2. Gaz sieciowy

Pojęcie gazu sieciowego zostało wprowadzone w 1952 roku przez Lee i Yanga jako pewien model rzeczywistego gazu. Zacznijmy od opisu gazu znajdującego się w objętości V i zawierającego $\mathcal{N}$ cząstek. Jak zwykle gdy liczba cząstek jest zadana, posługiwać się będziemy wielkim rozkładem kanonicznym (aby uniknąć stałego uwzględniania warunku $\mathcal{N}$ =const). Wielki rozkład wybieramy tak aby średnia liczba cząstek $< N >$ wyliczona przy jego pomocy była równa zadanej ($\mathcal{N}$). Wielka suma ma postać

$$Z_G(T,V,\lambda) = \sum_{N=0}^{V} \lambda^N \, Z(T,V,N) \tag{5.31}$$

gdzie $\lambda = \exp(\beta\mu)$ to aktywność, μ - potencjał chemiczny, zaś $Z(T,V,N)$ to kanoniczna suma stanów dla układu o N cząstkach. Przyjęliśmy, że maksymalna liczba molekuł gazu równa jest objętości V. Wielki potencjał termodynamiczny Ω wprowadzamy przez

$$Z_G = e^{-\beta\Omega}, \quad \Omega = -kT\ln Z_G \tag{5.32}$$

Znajdziemy teraz związek wielkiej sumy stanów i ciśnienia, co będzie przydatne przy konstrukcji równania stanu. Związek kanonicznej i wielkiej sumy stanów jest taki jak dla potencjałów

$$-\Omega = F + \mu N, \tag{5.33}$$

$$-d\Omega = dF + Nd\mu \tag{5.34}$$

Mamy też

$$dF = pdV + SdT \tag{5.35}$$

gdzie $F = F(V,T)$ to energia swobodna Helmholtza. Potencjał chemiczny to energia swobodna Gibbsa, $G = G(T,p)$, na cząstkę, czyli

$$G = N\mu \;\Rightarrow\; dG = Nd\mu \tag{5.36}$$

tak więc łącząc (5.33), (5.34), (5.35) i (5.36) dostajemy

$$-d\Omega = pdV + SdT + dG \tag{5.37}$$

Z drugiej jednak strony, ponieważ $G = G(T,P)$, więc ze związków Maxwella

$$dG = -SdT + Vdp \tag{5.38}$$

i stąd

$$-d\Omega = pdV + SdT - SdT + Vdp = d(pV) \tag{5.39}$$

Ostatecznie więc, z dokładnością do stałej niezależnej od p, T,

$$-\Omega = pV \tag{5.40}$$

lub

$$p = \frac{kT}{V}\ln Z_G \tag{5.41}$$

Ponieważ

$$N = -\left(\frac{\partial\Omega}{\partial\mu}\right) \tag{5.42}$$

więc gęstość gazu sieciowego

$$\rho = \frac{N}{V} = \frac{kT}{V}\frac{\partial\ln Z_G}{\partial\mu} \tag{5.43}$$

Otrzymaliśmy wyrażenia dla ciśnienia i gęstości gazu poprzez wielką sumę stanów. Do rozwiązania zagadnienia wystarczy więc znalezienie wyrażenia dla Z_G. Dotąd rozważania były całkiem ogólne. Teraz wprowadzimy model gazu sieciowego. Załóżmy, że para molekuł odległych o q oddziałuje potencjałem $\phi(q)$. Stąd suma stanów dla takiego układu dana jest przez

$$Z = (N!)^{-1}\int_V e^{-\beta\phi(q)}dq, \tag{5.44}$$

Podzielmy teraz naszą objętość na V jednakowych komórek o jednostkowej objętości. Przyjmijmy, że $N < V$. W każdej komórce może być conajwyżej jedna molekuła, lub komórka może być pusta. Molekuły mogą w każdej chwili swobodnie przechodzić z jednej komórki do drugiej, o ile ta ostatnia nie jest zajęta. Oddziaływania międzymolekularne przedstawimy jako oddziaływania między komórkami, tak aby możliwie dobrze przybliżały oddziaływania w realnym gazie. Będziemy więc mieli krótko zasięgowe nieskończenie silne odpychanie (molekuły nie mogą znajdować się w tym samym miejscu – komórce), przyciąganie na skończonej odległości i brak oddziaływań jeżeli molekuły są dostatecznie daleko. Rozpatrujemy oddziaływanie par, tzn.

$$\phi = \sum \phi_2, \tag{5.45}$$

i zgodnie z tym co przyjęliśmy powyżej, ϕ_2 ma postać:

$$\phi_2 = \begin{cases} \infty & \text{gdy 2 molekuły są w tej samej komórce (twarde sfery)} \\ -\epsilon & \text{gdy komórki są swymi najbliższymi sąsiadami} \\ 0 & \text{gdy są dalszymi sąsiadami} \end{cases} \tag{5.46}$$

Przy takim podziale objętości na komórki, całka we wzorze na sumę stanów zamienia się na sumę. Każdej komórce przypisujemy zmienną $t_j = 1$ jeżeli komórka jest zajęta i $t_j = 0$, jeżeli jest pusta. t_j jest nazywana zmienną obsadzeniową. Dla sumy stanów możemy napisać

$$Z = \sum_{\{t_j\}} e^{\beta\epsilon \sum_{\langle i,j \rangle}^{V} t_i t_j} \tag{5.47}$$

z dodatkowym warunkiem aby suma wszystkich molekuł była równa $\mathcal{N}$, tzn

$$\sum_{i=1}^{V} t_i = \mathcal{N}. \tag{5.48}$$

Ograniczenia wynikającego ze stałości liczby cząstek możemy się pozbyć wprowadzając wielką sumę stanów:

$$Z_G = \sum_{N=1}^{V} \left(e^{\beta\mu}\right)^N \sum_{\{t_i\}} e^{\beta\epsilon \sum t_i t_j} = \sum_{\{t_i\}} e^{\beta\epsilon \sum t_i t_j + \beta\mu \sum t_i}. \tag{5.49}$$

Wyrażenie to możemy podstawić do wzorów (5.41,5.43). Warto jednak zauważyć, że jest ono, przynajmniej formalnie, analogiczne do wzoru na kanoniczną postać sumy stanów dla modelu Isinga o spinie 1/2:

$$Z_{Is} = \sum_{\sigma} e^{\beta J \sum \sigma_i \sigma_j + \beta h \sum \sigma_i}, \tag{5.50}$$

gdzie $\sigma = \pm 1$. Aby przeprowadzić wielką sumę stanów w sumę stanów dla modelu Isinga dokonajmy podstawienia $\sigma_i = 1 - 2t_i$, a wtedy

$$\begin{aligned} Z_{Is} &= \sum_{\{t\}} \exp \left\{ \beta J \sum_{\{t_i\}} (1-2t_i)(1-2t_j) + \beta h \sum (1-2t_i) \right\} \\ &= \exp\left\{V(1/2\beta J z + \beta h)\right\} \sum_{\{t\}} \exp\left\{4\beta J \sum t_i t_j - 2\beta(Jz+h)\sum t_i\right\} \end{aligned} \tag{5.51}$$

Aby więc mieć $Z_{Is} = Z_G$, należy dokonać przyporządkowania

$$4J = \epsilon, \tag{5.52}$$

$$2(Jz+h) = -\mu \quad \Rightarrow \beta h = -1/2 \ln\lambda - \beta J z, \tag{5.53}$$

$$\ln Z_{Is} = V\beta(1/2Jz + h) + \ln Z_G. \tag{5.54}$$

Wracając do wyprowadzonych na początku tego paragrafu związków miedzy wielką sumą stanów i ciśnieniem oraz gęstością dostajemy dla ciśnienia (na cząstkę)

$$p = kT \ln Z_G = -1/2J\,z - h + kT\,\ln Z_{Is} = -1/2J\,z - h - f_{Is}, \tag{5.55}$$

gdzie f_{Is} jest gęstością energii swobodnej dla modelu Isinga. Gęstość gazu to

$$\rho = N/V = \frac{\sum t_i}{V}. \tag{5.56}$$

Można to wyrazić poprzez magnetyzację modelu Isinga m

$$m = \frac{1}{2V}\sum_{i=1}^{V}[(1+\sigma_i)-(1-\sigma_i)] =$$

$$= 1/V\sum_i \sigma_i = 1/V\sum_i(1-2t_i) = 1-2\frac{\sum t_i}{V} = 1-2\rho, \tag{5.57}$$

czyli

$$m = 1-2\rho. \tag{5.58}$$

Liczba spinów „w górę” $n_\uparrow = 1/2\sum_i(1+\sigma_i) = V-\rho$,
Liczba spinów „w dół” $n_\downarrow = 1/2\sum(1-\sigma_i) = \rho$.

Widać więc, że model gazu sieciowego i model Isinga są całkowicie równoważne. Własności termodynamiczne gazu sieciowego można otrzymywać wprost ze znanych wyników dla modelu Isinga, posługując się następującą tabelką równoważności:

Model Isinga	Gaz sieciowy
4 J	ϵ
β h $+\beta$ J z	$-\frac{1}{2}\ln\lambda$
m	1 - 2ρ
$n_\downarrow$	ρ
$n_\uparrow + n_\downarrow$	V
$-\frac{1}{2}$ J z + h + f_{Is}	-p

Dla modelu Isinga mamy izotermy (m,h), natomiast dla gazu sieciowego izotermy (p,V). Jeżeli wiemy, że izoterma modelu Isinga ma nieciągłość dla pewnej wartości m, to odpowiadać temu będzie nieciągłość izotermy w funkcji V, co oznacza istnienie przejścia fazowego dla gazu sieciowego. Jest to kondensacja wywołana przez krótkozasięgowe oddziaływania przyciągające. Tak więc w gazie sieciowym, a przynajmniej w jego najprostszej wersji rozpatrywanej powyżej, istnieją dwie fazy – ciekła i gazowa.

Widać też, że gaz sieciowy jest równoważny modelowi Isinga w zewnętrznym polu magnetycznym, a więc ścisłe rozwiązanie dla gazu sieciowego istnieje tylko w 1D. W ostatnich latach model gazu sieciowego był bardzo często stosowany przy użyciu techniki automatów komórkowych, co pozwoliło na rozpatrywanie układów dynamicznych, dla których nie można było otrzymać rozwiązania poprzez równania różniczkowe.

Innym przykładem modelu równoważnego modelowi Isinga jest *stop podwójny*, w którym węzeł i może być zajęty albo przez atom typu A ($\sigma_i = +1$), albo przez atom typu B ($\sigma - i =$ -1). W zależności od typów oddziaływań pomiędzy cząstkami, będą teraz trzy typy tych oddziaływań, możemy mieć np. separację faz, czyli oddzielnie część układu bogatą w cząstki A i inną bogatą w cząstki B.

3. Magnetyczny gaz sieciowy

Jest to gaz sieciowy, w którym cząstki maja dodatkowy stopień swobody, posiadają bowiem wypadkowy moment magnetyczny. Jak zobaczymy, będzie to prowadziło do bogatszego diagramu fazowego. Rozpatrzmy układ cząstek gazu posiadających momenty magnetyczne. Cząstki te mogą oddziaływać zarówno siłami magnetycznymi (J_1), jak i niemagnetycznymi (J_2). Na układ ten możemy też spojrzeć jak na rozrzedzony magnetyk. Zastosujmy przybliżenie annealed i nieporządek typu węzłów, ponieważ niektóre węzły będą osadzone, a niektóre nie. Hamiltonian takiego systemu ma postać

$$\mathcal{H} = -J_1 \sum t_i \sigma_i t_j \sigma_j - J_2 \sum t_i t_j, \tag{5.59}$$

gdzie $t = 0, 1$, $\sigma = \pm 1$. Do rozwiązania problemu zastosujmy przybliżenie pola molekularnego, zarówno do zmiennych obsadzeniowych, czy też gazowych (t), jak i magnetycznych (σ). Równe traktowanie obu typów zmiennych wynika z przyjętego podejścia annealed.

$$\mathcal{H} \equiv \mathcal{H}_{MFA} = Nz\left(J_1 \langle t\sigma\rangle^2 - J_2 \langle t\rangle^2\right) - 2J_1 \langle t\sigma\rangle \sum_i t_i \sigma_i - 2J_2 \langle t\rangle \sum_i t_i. \tag{5.60}$$

Oznaczając $\langle t\sigma\rangle = m$, $\langle t\rangle = \rho$, mamy

$$\mathcal{H} = Nz\left(J_1 m^2 + J_2 \rho^2\right) - 2J_1 m \sum_i t_i \sigma_i - 2J_2 \rho \sum_j t_j. \tag{5.61}$$

Odpowiednio dla sumy stanów Z mamy

$$Z = A \, \mathrm{Tr}_t \; \mathrm{Tr}_\sigma \; e^{-\beta\left(\mathcal{H} - \mu \sum_i t_i\right)}, \tag{5.62}$$

gdzie

$$A = \exp\left(-\beta N z (J_1 m^2 + J_2 \rho^2)\right). \tag{5.63}$$

Dla sumy stanów mamy

$$\begin{aligned} Z = A \; \mathrm{Tr}_t \; \mathrm{Tr}_\sigma \; e^{\sum t_j (2J_1 m \sigma_j - 2J_2 \rho)} &= \\ = A \; \mathrm{Tr}_t \, e^{\beta(2J_2\rho + \mu)\sum t_j} \sum_{\sigma=\pm 1} e^{2\beta J_1 m t_i \sigma_i} & \\ = e^{-\beta N\left(J_1 m^2 + J_2 \rho^2\right)} \left(1 + e^{2\beta J_2 \rho + \mu} \cosh 2\beta J_1 m\right)^N 2^N & \end{aligned} \tag{5.64}$$

Stąd energia swobodna na węzeł

$$f = \left(J_1 m^2 + J_2 \rho^2\right) - kT \ln 2 - kT \ln\left(1 + e^{2\beta J_2 \rho + \mu} \cosh 2\beta J_1 m\right). \tag{5.65}$$

Warunki stabilności, czyli optymalny wybór parametrów m i ρ, otrzymujemy z żądania (twierdzenie Peierlsa)

$$\frac{\partial f}{\partial \rho} = 0, \qquad \frac{\partial f}{\partial \mu} = 0, \tag{5.66}$$

a stąd już wyrażenia dla m i ρ

$$m = \frac{e^{2\beta J_2 \rho + \mu} \sinh 2\beta J_1 m}{1 + e^{2\beta J_2 \rho + \mu} \cosh 2\beta J_1 m}, \tag{5.67}$$

$$\rho = \frac{e^{2\beta J_2 \rho + \mu} \cosh 2\beta J_1 m}{1 + e^{2\beta J_2 \rho + \mu} \cosh 2\beta J_1 m}. \tag{5.68}$$

Jeżeli przyjmiemy, że gęstość ρ jest zadanym zewnętrznym parametrem, wówczas z ostatniego równania możemy wyrugować μ,

$$e^{2\beta J_2 \rho + \mu} \cosh 2\beta J_1 m (1 - \rho) = \rho. \tag{5.69}$$

Pozwala to zapisać magnetyzację jako

$$m = \rho \tanh 2\beta J_1 m. \tag{5.70}$$

Drugie równanie stanu otrzymamy pamiętając, że dla gazu sieciowego energia swobodna jest równa ciśnieniu (ze znakiem minus)

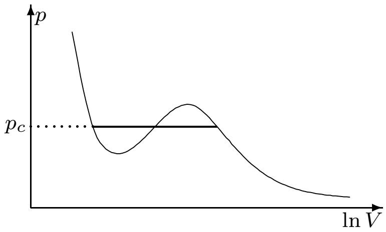
$$p = -f. \tag{5.71}$$

Stąd dwa równania stanu

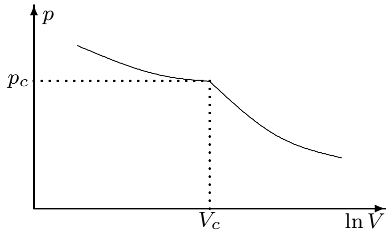
$$p = J_1 m^2 + J_2 \rho^2 - kT \ln 2 + kT \ln(1 - \rho), \tag{5.72}$$

$$m = \rho \tanh 2\beta J_1 m. \tag{5.73}$$

W zależności od temperatury, gęstości cząstek oraz stosunku stałych oddziaływania, układ może znaleźć się w jednej z trzech faz: o niskiej gęstości i zerowej magnetyzacji (niemagnetyczny gaz – NG), wysokiej gęstości i zerowej magnetyzacji (niemagnetyczna ciecz – NC) oraz wysokiej gęstości i niezerowej magnetyzacji (magnetyczna ciecz – MC). Diagram fazowy w płaszczyźnie (p, T) otrzymujemy w każdym przypadku wykreślając izotermy (p, V) z wykorzystaniem równania dla magnetyzacji. Ze względu na stosowane przybliżenie MFA pojawią się niefizyczne części izoterm, w stosunku do których trzeba zastosować konstrukcję Maxwella (rysunek 5.4). Przejścia fazowe ciągłe otrzymamy jako zagięcia na izotermach (rysunek 5.5). Diagram fazowy ma postać jak na

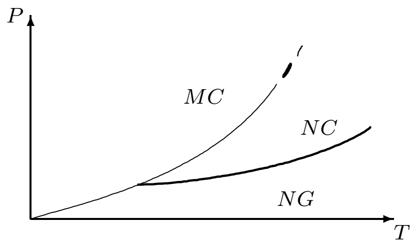


Rysunek 5.4. Przejścia fazowe nieciągłe dla magnetycznego gazu sieciowego.



Rysunek 5.5. Przejścia fazowe ciągłe dla magnetycznego gazu sieciowego.

rysunku (5.6). Na to aby istniała faza magnetycznej cieczy, oddziaływania magnetyczne muszą



Rysunek 5.6. Diagram fazowy dla magnetycznego gazu sieciowego.

być dostatecznie silne.

Jest rzeczą jasną, że rozpatrywany układ może reprezentować bardzo wiele rzeczywistych sytuacji – układów charakteryzowanych przez dwa (powiązane ze sobą) parametry porządku. Jeden reprezentuje gęstość cząstek – drugi ich dodatkowy stopień swobody.

ROZDZIAŁ 6

Metoda szeregów wysokotemperaturowych

Badanie zjawisk zachodzących w pobliżu przejścia fazowego wymaga zastosowania metod przybliżonych, gdyż jedyny ścisły wynik dotyczy dwuwymiarowego modelu Isinga. Techniką, która jest powszechnie uznawana za dającą najbardziej wiarygodne wyniki, jest metoda szeregów wysokotemperaturowych. Została ona opracowana po raz pierwszy jeszcze w latach 30-tych przez W. Opęchowskiego. Pozwala ona na dokładne wyznaczenie wartości temperatury krytycznej oraz, już z mniejszą dokładnością, wykładników krytycznych. Przedstawimy ją tutaj na szczególnie prostym przykładzie modelu Isinga o spinie 1/2 na siatce kwadratowej z oddziaływaniem ograniczonym do najbliższego sąsiedztwa (NN).

Suma stanów ma postać

$$Z_N = \mathrm{Tr}\, e^{\sum\limits_{\langle i,j\rangle} K\sigma_i\sigma_j} = \sum_{\sigma} \prod_{\langle i,j\rangle} e^{K\sigma_i\sigma_j}, \tag{6.1}$$

gdzie $K \equiv \beta J$. Ponieważ $\sigma^2 = 1$, więc

$$\exp(K\sigma_i\sigma_j) = \cosh K + \sigma_i\sigma_j \sinh K = \cosh K\,(1 + \sigma_i\sigma_j\kappa) \tag{6.2}$$

$\kappa = \tanh K$, a więc

$$Z_N = \sum_{\sigma} \prod_{\langle i,j\rangle} \cosh K\,(1 + \sigma_i\sigma_j\kappa) = (\cosh K)^{\frac{Nqz}{2}} \sum_{\sigma} \prod_{\langle i,j\rangle} (1 + \sigma_i\sigma_j\kappa)\,, \tag{6.3}$$

gdzie z to liczba najbliższych sąsiadów. Tak więc suma stanów jest szeregiem nieskończonym, ponieważ w granicy termodynamicznej wielomian przechodzi w szereg. Szansa na wyznaczenie wszystkich współczynników tego szeregu jest oczywiście niewielka. Co więc stało by się gdyby szereg „oberwać" na pewnym wyrazie ? Oznacza to zamianę szeregu na wielomian, a o ile nieskończony szereg może mieć osobliwości, o tyle jedyną osobliwością wielomianu jest jego zerowanie się. Jeżeli więc przybliżymy szereg przez wielomian, wyznaczymy stąd sumę stanów i dalej różniczkując - magnetyzację i podatność magnetyczną, to ta ostatnia, nawet w granicy termodynamicznej, nie będzie rozbieżna.

Idea metody szeregów wysokotemperaturowych (z angielskiego *high temperature series expansion – HTSE*) polega na tym aby obliczając kolejne współczynniki szeregu, który w T_c będzie rozbieżny (np. podatność), otrzymać ciąg przybliżeń „prawdziwej" podatności i ekstrapolując uzyskane wyniki otrzymać położenie punktu osobliwego. Sprowadza się to do wyznaczenia promienia zbieżności szeregu dla podatności. Zmienną jest tu $\kappa = \tanh K$, a więc temperatura, a raczej jej odwrotność. Znane kryteria z algebry wymagają jednak znajomości wszystkich wyrazów szeregu. Niestety odpowiedź na pytanie czy na podstawie znajomości pewnej, skończonej liczby N wyrazów możemy określić promień zbieżności szeregu, brzmi **NIE**. Mimo tego fizycy wierzą, że rozpatrywane przez nich wielkości fizyczne mają na tyle porządne zachowanie, że taka ekstrapolacja jest możliwa. Wiara ta jak narazie, w ciągu ponad 40 lat stosowania tej techniki, nas nie zawiodła. Dla dwu-wymiarowego (2D) modelu Isinga, gdzie znany jest wynik dokładny, różnica między nim a wynikiem otrzymanym metodą HTSE wynosi około 0.01 %. Dlatego też wyniki uzyskane tą metodą uważa się za bardzo dokładne. Na podstawie metody HTSE Domb uzyskał oszacowanie dla zależności magnetyzacji od temperatury 2D modelu Isinga jako $M(T) \sim (T_c - T)^{1/8}$, a więc wynik otrzymany później ściśle przez Onsagera.

Nim zajmiemy się konstrukcją szeregu, napiszmy formalnie szereg dla podatności początkowej $(H=0)$ ferromagnetyka

$$\chi_0(T)=\sum_{n=0}^{\infty} a_n \kappa^n. \tag{6.4}$$

Wszystkie znane dotąd wyrazy są, dla ferromagnetyka, dodatnie, co oznacza, że najbliższa początku układu (w płaszczyźnie zespolonej κ) osobliwość leży na dodatniej rzeczywistej osi κ. Ma ona więc charakter fizyczny, gdyż odpowiada jej rzeczywista, dodatnia temperatura – identyfikujemy ją z T_c. Jeżeli jest to dostatecznie silna osobliwość, a tak na ogół jest, określa ona asymptotyczne zachowanie współczynników. Aby wyznaczyć T_c należy więc znaleźć promień zbieżności $\chi_0(T)$. Można posłużyć się np, wzorem Cauchy'ego

$$r=\lim_{n\to\infty} |a_n|^{1/n}, \tag{6.5}$$

albo też przyjąć, że w okolicy osobliwości szereg ma postać

$$\chi_0 \sim A\left(1-\frac{\kappa}{\kappa_c}\right)^{-\gamma}, \tag{6.6}$$

co oczywiście odpowiada postaci χ_0 wynikającej z granicy termodynamicznej. Gdybyśmy jeszcze wiedzieli, ale nie wiemy, że wszystkie współczynniki a_n są dodatnie, to na pewno najbliższą osobliwością byłoby κ_c. Napiszmy teraz wzór dwumienny Newtona dla (6.6) i współczynniki tego rozwinięcia porównajmy z a_n ze wzoru (6.4). Przy $n\to\infty$ mamy

$$a_n \sim \frac{A}{\kappa_c^n}\frac{(n+\gamma-1)!}{n!(\gamma-1)!}. \tag{6.7}$$

Biorąc stosunek dwu kolejnych współczynników dostajemy

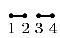
$$r_n=\frac{a_n}{a_{n-1}}=\frac{1}{\kappa_c}\left(1+\frac{\gamma-1}{n}\right). \tag{6.8}$$

Tak więc dla dużych n $(n\to\infty)$ wykres zależności $r_n(n^{-1})$ powinien stawać się linią prostą, przecinającą oś rzędnych (odwrotnej temperatury) w punkcie κ_c^{-1} i z nachyleniem $\kappa_c^{-1}(\gamma-1)$. Możemy stąd wyznaczyć zarówno temperaturę krytyczną jak i wykładnik krytyczny γ.

Zobaczymy teraz jak należy konkretnie budować szereg (6.4). Zacznijmy od Z_N. Nie jest trudno zauważyć, wyrazy szeregu można znaleźć albo na drodze algebraicznej albo geometrycznej. Dla prostoty rozważmy siatkę kwadratową, aczkolwiek rozumowanie jest takie samo i dla innych typów siatek.

σ_i to dwu-wartościowa zmienna (spin $=\pm1$) ulokowana w węźle i. $\kappa_{ij}\equiv\tanh K_{ij}$ to oddziaływanie między węzłami i oraz j. σ reprezentować będziemy przez węzły, natomiast κ to linie łączące te węzły. Jak widać ze wzoru (6.3) mamy

$$\begin{aligned} Z_N &= \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.9)\\ &=(\cosh K)^{\frac{Nq}{2}}\sum_{\sigma_i}(1+\sigma_1\sigma_2\kappa)(1+\sigma_1\sigma_3\kappa)(1+\sigma_1\sigma_4\kappa)(1+\sigma_1\sigma_5\kappa)(1+\sigma_2\sigma_1\kappa)(1+\sigma_2\sigma_6\kappa)\ldots\\ &=(\cosh K)^{\frac{Nq}{2}}\sum_{\sigma_i}\left(1+\kappa\sum_{\langle i,j\rangle}\sigma_i\sigma_j+\kappa^2\sum_{\langle i,j\rangle}\sum_{\langle k,l\rangle}\sigma_i\sigma_j\sigma_k\sigma_l\ldots\right). \end{aligned}$$

Sumowanie przebiega po $i,j\neq k,l$. Możemy teraz każdej parze zmiennych $\sigma_i\sigma_j$ przyporządkować wiązanie (linię) łączącą węzły i,j, które są swymi najbliższymi sąsiadami. Linia taka będzie miała „wagę” κ. Wyrazy szukanego szeregu staja się grafami na rozważanej siatce. Dla pierwszego wyrazu będzie to lub , dla drugiego a)  lub b)  Odpowiednie wyrażenia będą miały postać:

w przypadku a) $\kappa^2(\sigma_1\sigma_2\sigma_2\sigma_3)=\kappa^2(\sigma_1\sigma_3)$,

i w przypadku b) $\kappa^2(\sigma_1\sigma_2\sigma_3\sigma_4)$.

Jeżeli więc zmienna pojawia się parzystą liczbę razy, to ponieważ przyjmuje ona wartości ±1, więc możemy ją zastąpić przez 1. Ponieważ sumowanie w (6.9) przebiega po różnych parach węzłów, więc też i grafy nie mogą się pokrywać. Co najwyżej mogą mieć wspólne węzły, jednak nie linie.

Aby wyznaczyć Z_N należy wziąć ślad, czyli wysumować po $\sigma = \pm 1$. Jeżeli w którymś z wyrazów rozwiniętego iloczynu zmienna σ_i pojawia się w nieparzystej potędze, to wyraz ten zniknie, ponieważ $\sum_{\sigma_i=\pm 1} = 0$. Ostatecznie dla sumy stanów mamy

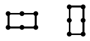
$$Z_N = (\cosh K)^{\frac{Nz}{2}} \sum_{r=0}^{\infty} n(r)\kappa^r, \tag{6.10}$$

gdzie $n(0) = 1$, zaś $n(r)$ jest liczbą nieznikających grafów o r wiązaniach, czyli liczbą sposobów na jakie na danej siatce można umieścić graf składający się z r linii i spełniający dwa warunki

(1) każde połączenie może w danym grafie pojawić się tylko raz,
(2) w każdym węźle siatki, który bierze udział w sumowaniu, spotyka się parzysta liczba linii.

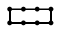
Tak więc cały problem sprowadzony został do wyliczenia na ile sposobów poszczególne grafy mogą być rozmieszczone na sieci.
Na siatce kwadratowej dla $r = 1, 2, 3$ nie ma grafów spełniających oba warunki. Dla $r = 4$ jest jeden – kwadrat. Ponieważ może on być umieszczony w N punktach (zakładamy periodyczne warunki brzegowe), więc

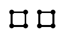
$$n(4) = N. \tag{6.11}$$

Dla $r = 5$ żaden graf nie spełnia nałożonych warunków. Dla $r = 6$ są dwa 
przy czym każdy może być umieszczony w N punktach, a więc

$$n(6) = 2N. \tag{6.12}$$

$r = 7$ też nie daje wkładu. Dla $r = 8$ mamy

,  (* 2),  (* 4)

czyli razem 7N. Tutaj jednak pojawia się coś nowego – dwa grafy rozłączne  Pierwszy może być znów umieszczony w N węzłach, ale drugi już tylko w $N - 5$, ponieważ kwadraty nie mogą mieć ani wspólnych boków, ani też nie mogą się pokrywać. Ponieważ oba kwadraty są identyczne, należy to podzielić przez 2. Ostatecznie więc wkład od r =8 to

$$n(8) = 7N + 1/2N(N-5) = 1/2\,N^2 + 9/2\,N. \tag{6.13}$$

Dalsza procedura jest analogiczna, choć oczywiście coraz bardziej skomplikowana. Zatrzymajmy się na poziomie κ^8. Mamy

$$Z_N = (2\cosh^2 K)^N \left(1 + N\kappa^4 + 2N\kappa^6 + (\frac{1}{2}N^2 + \frac{9}{2}N)\kappa^8 + \ldots\right). \tag{6.14}$$

W wyrażeniu tym występują wyrazy proporcjonalne do N^2, natomiast $\ln Z_N$ powinien, jako wielkość addytywna, być proporcjonalnym do N. Jeżeli jednak zlogarytmujemy obie strony (6.14), a następnie rozłożymy w szereg według wzoru, co możemy zrobić gdyż interesuje nas otoczenie punktu krytycznego gdzie argument logarytmu jest mały

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \ldots \tag{6.15}$$

to, zachowując wyrazy do κ^8 włącznie, otrzymamy

$$\begin{aligned} \ln Z_N &= N\ln(2\cosh^2 K) + N\kappa^4 + 2N\kappa^6 + \frac{1}{2}(N^2 + \frac{9}{2}N)\kappa^8 - \frac{1}{2}(N\kappa^4)^2 = \\ &= N\left(\ln(2\cosh^2 K) + \kappa^4 + 2\kappa^6 + \frac{9}{2}\kappa^8\right). \end{aligned} \tag{6.16}$$

Tak więc wyrazy proporcjonalne do N^2 uprościły się, i tak jest we wszystkich rzędach rozwinięcia. Otrzymany rozkład energii swobodnej ($\ln Z_N$) poprzez rozkład $\ln(1+x)$ nosi nazwę **rozkładu**

według kumulantów lub półniezmienników. Jest on często stosowany także w innych zagadnieniach i dlatego warto jest go przedstawić w ogólniejszej postaci. Weźmy

$$\exp(-\beta\mathcal{H}) = 1 - \beta\mathcal{H} + \frac{1}{2}(\beta\mathcal{H})^2 + \dots \tag{6.17}$$

Dla sumy stanów

$$\begin{aligned} Z_N &= Tr(1) - \beta\, Tr(\mathcal{H}) + \frac{1}{2}(\beta)^2 Tr\,(\mathcal{H})^2 + \dots \\ &= Tr(1)\left(1 - \beta\, tr(\mathcal{H}) + \frac{1}{2}\, tr\,(\mathcal{H})^2 + \dots\right) \end{aligned} \tag{6.18}$$

gdzie $\mathrm{Tr}(1) = 2^N$ jest liczbą stanów układu, natomiast

$$tr\,(\mathcal{H})^a = Tr(1)^{-1} Tr\,(\mathcal{H})^a . \tag{6.19}$$

Wielkości

$$b_n = tr\,(\mathcal{H})^n \tag{6.20}$$

noszą nazwę *momentów hamiltonianu.*
Rozkładając $\ln(1+x)$ w szereg mamy

$$\ln Z_N = \ln(Tr(1)) + \ln\left(1 - \beta tr\mathcal{H} + \frac{1}{2} tr(\mathcal{H})^2 + \dots\right). \tag{6.21}$$

Dokonując tu rozkładu logarytmu dostajemy

$$\ln Z_N = \sum_{r=0}^{\infty}\left(\frac{c_r}{r!}\right)\beta^r, \tag{6.22}$$

gdzie c_r to kumulanty (półniezmienniki). Ich postać można otrzymać porównując (6.21) i (6.22). Pierwsze trzy mają postać

$$\begin{aligned} c_0 &= \ln\left[tr(1)\right], \\ c_1 &= -b_1 = -tr(\mathcal{H}), \\ c_2 &= b_2 - b_1^2 = tr(\mathcal{H})^2 - (tr\mathcal{H})^2 . \end{aligned} \tag{6.23}$$

Mamy więc, z dokładnością do κ^8, wyrażenie dla energii swobodnej (6.16). Jak się jednak okazuje, szeregi dla ciepła właściwego, które można dostać z wyrażenia na energię swobodną przez różniczkowanie względem T, są raczej wolno zbieżne. Dlatego też do wyznaczenia wartości temperatury krytycznej korzysta się na ogół z szeregów dla początkowej podatności χ_0, tzn. reakcji układu na niewielkie zewnętrzne pole magnetyczne, które zostało następnie szybko wyłączone.

Pokażemy teraz jak należy postępować dla znalezienia szeregu dla podatności. Rozpatrywać znów będziemy ten sam najprostszy przypadek, czyli model Isinga o spinie $1/2$ na siatce kwadratowej z oddziaływaniami ograniczonymi do NN. Hamiltonian takiego układu ma postać

$$\mathcal{H} = -K\sum_{<i,j>}\sigma_i\sigma_j - h\sum_i \sigma_i. \tag{6.24}$$

Stąd dla sumy stanów

$$Z_N = Tr_\sigma \exp\left\{K\sum_{i,j}\sigma_i\sigma_j + L\sum_i\sigma_i\right\}, \qquad \text{gdzie} \quad K = \beta J,\ L = \beta H. \tag{6.25}$$

Postępując tak jak poprzednio, dostajemy

$$Z_N = (\cosh K)^{2N}(\cosh L)^N\, Tr_\sigma \prod_{<i,j>}(1 + \sigma_i\sigma_j\kappa)\prod_i(1 + \sigma_i h),$$

gdzie $\kappa = \tanh K, h = \tanh L$. Wprowadźmy unormowane ślady, $\mathrm{tr}(1) = 1$, a wtedy

$$Z_N = 2^N(\cosh K)^{2N}(\cosh L)^N\, \mathrm{tr}_\sigma \prod_{<i,j>}(1 + \sigma_i\sigma_j\kappa)\prod_i(1 + \sigma_i h) \tag{6.26}$$

W dalszych obliczeniach rozpatrywać będziemy tylko wyrazy proporcjonalne do h^2, co nie jest żadnym uproszczeniem, ponieważ podatność magnetyczna jest drugą pochodną energii swobodnej względem pola, a po wykonaniu różniczkowania zewnętrzne pole będzie wyłączone. Tak więc wkład do podatności dadzą tylko te wyrazy. Ograniczeniem jest natomiast rozpatrywanie tylko wyrazów zawierających κ w potędze nie wyższej niż 8. Wzór (6.26) można przepisać jako

$$\begin{aligned} Z_N = 2^N\left(\cosh K\right)^{2N}\left(\cosh L\right)^N tr\left[1+\sum\kappa\sigma_i\sigma_j+\kappa^2\sum\sigma_i\sigma_j\sigma_k\sigma_l+\ldots\right]*\\ \left[1+h\sum\sigma_i+h^2\sum\sigma_i\sigma_j+\ldots\right]=\\ =\mathcal{A}\,\mathrm{tr}\left[1+\kappa h^2\sum\sigma_1^2\sigma_2^2+\kappa^2h^2\sum\sigma_1^2\sigma_2^2\sigma_3^2+\ldots\right] \end{aligned}\tag{6.27}$$

gdzie ze względu na operację śladu zostały tylko te wyrazy, które operację tę przetrwają. Jeżeli teraz przejdziemy do „reprezentacji" graficznej, to pole zewnętrzne działające na spin w węźle, będziemy przedstawiać jako krzyżyk. Znak $\sum$ oznacza sumowanie po wszystkich węzłach sieci, jak również sumę wszystkich grafów danego rzędu, tzn. zawierających tę samą liczbę linii (wiązań). Mamy więc

$$\begin{aligned} \frac{1}{N}\ln Z_N = \ln\mathcal{A}+\frac{1}{N}\ln\Big(1+h^2\Big[\kappa\,\mathrm{tr}\sum\times\!\!-\!\!\times+\kappa^2\,\mathrm{tr}\sum\ +\ldots+\kappa^8\,\mathrm{tr}\sum\ \Big]+\\ +\kappa^4\,\mathrm{tr}\sum\square+\kappa^8\,\mathrm{tr}\sum\square\,\square+\ldots\Big). \end{aligned}\tag{6.28}$$

Zgodnie z tym co przyjęliśmy powyżej, rozłóżmy teraz ln(1+x) w szereg. Ponieważ rozpatrujemy przypadek kiedy pola są słabe, a temperatury wysokie, więc x jest małe.

$$\begin{aligned} \frac{1}{N}\ln Z_N = \ln\mathcal{A}+h^2\Big[\kappa\,\mathrm{tr}(\sum\times\!\!-\!\!\times)+\ldots+\kappa^8\,\mathrm{tr}(\sum\ )+\ldots\Big]+\\ +\kappa^4\,\mathrm{tr}(\sum\square)+\ldots+\kappa^8\,\mathrm{tr}(\sum\square\,\square)+\ldots-\frac{1}{2}\kappa^4\,\mathrm{tr}(\sum\square)\cdot\kappa^4\,\mathrm{tr}(\sum\square)+\ldots\\ -h^2\Big[\kappa^4\,\mathrm{tr}(\sum\square)\,\mathrm{tr}(\sum\times\!\!-\!\!\times)\kappa+\ldots+\kappa^4\,\mathrm{tr}(\sum\square)\cdot\kappa^4\,\mathrm{tr}(\sum\ )+\kappa^6\,\mathrm{tr}(\sum\ )\cdot\kappa\,\mathrm{tr}(\sum\times\!\!-\!\!\times)\Big]\ldots \end{aligned}$$

Zacznijmy teraz liczyć wkłady od poszczególnych wyrazów

$$\mathrm{tr}(\sum\times\!\!-\!\!\times)=2N,$$

bo wiązanie może się zacząć w każdym z N punktów siatki, oraz może być poziome i pionowe.

$$\mathrm{tr}(\sum\times\!\!-\!\!\bullet\!\!-\!\!\times)=6N$$

$$\mathrm{tr}(\sum\times\!\!-\!\!\bullet\!\!-\!\!\bullet\!\!-\!\!\times)=18N,$$

$$\mathrm{tr}(\sum\ )=50N,$$

$$\mathrm{tr}(\sum\ )=150N+2N(N-6),$$

$$\mathrm{tr}(\sum\ )=418N+6N(N-8),$$

$$\mathrm{tr}(\sum\ )=1196N+14N(N-10)+6N(N-9)+2N(N-8),$$

$$\mathrm{tr}(\sum\ )=3302N+36N(N-12)+24N(N-11)+2N(N-10),$$

Począwszy od piatego rzędu pojawiają się grafy niespójne, typu tr($\square\times\!\!-\!\!\times$). Pojedyńczy graf z polem może być w $2N$ punktach, ale ponieważ grafy $\square$ i $\times\!\!-\!\!\times$ są niezależne, więc kwadrat nie może dotykac odcinka, czyli odcinek (przy ustalonym położeniu kwadratu) może byc umieszczony tylko w N - 6 węzłach. Stąd łączny wkład od takiego grafu niezwiązanego jest $2N(N-6)$.

Policzmy teraz kumulanty. W najniższym rzędzie mamy

$$\mathrm{tr}(\sum\ )-\mathrm{tr}(\sum\square)\,\mathrm{tr}(\sum\times\!\!-\!\!\times).$$

Teraz jednak sytuacja jest inna niż poprzednio, ponieważ na iloczyn śladów nie są nakładane żadne ograniczenia, tzn zakaz nakrywania się krawędzi. Możliwy jest więc graf . Grafów takich może być $4N$. Oprócz tego mogą w iloczynie pojawić się grafy $\mathrm{tr}(\sum\ )$, więc ostatecznie

$$\mathrm{tr}(\sum\ ) - \mathrm{tr}(\sum\ )(\mathrm{tr}(\sum\ ) =$$

$$= \mathrm{tr}(\widetilde{\sum}\ ) + \mathrm{tr}(\sum\ ) - \mathrm{tr}(\sum\ ) - \mathrm{tr}(\sum\ ) =$$

$$= \mathrm{tr}(\widetilde{\sum}\ ) - \mathrm{tr}(\sum\ ) - \mathrm{tr}(\sum\ ) = 150N - 12N = 138N.$$

$\widetilde{\sum}$ oznacza sumowanie wyłącznie po grafach związanych. Jak więc widać, w rozwinięciu kumulacyjnym uprościły się wszystkie wyrazy $\sim N^2$, a pozostały tylko proporcjonalne do N. Tak jest we wszystkich rzędach rozwinięcia. Ostatecznie mamy dla energii swobodnej na węzeł

$$\begin{aligned} -\beta f &= 2\ln\cosh K + \ln 2 + \ln\cosh L + \\ &+ h^2\left(2\kappa + 6\kappa^2 + 18\kappa^3 + 50\kappa^4 + 138\kappa^5 + 370\kappa^6 + 978\kappa^7 + 2586\kappa^8 + ..\right) + \\ &+ \left(\kappa^4 + 2\kappa^6 + \frac{9}{2}\kappa^8 + ..\right). \end{aligned} \tag{6.29}$$

Początkowa podatność χ_0, to

$$\chi_0 = \frac{(g\mu)^2}{2kT}\left(\frac{\partial^2 f}{\partial L^2}\right)_{L=0}, \tag{6.30}$$

i stąd

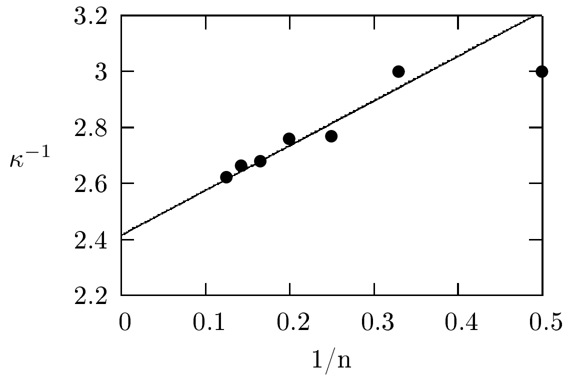
$$\chi_0 = \frac{(g\mu)^2}{4kT}\left(1 + 4\kappa + 12\kappa^2 + 36\kappa^3 + 100\kappa^4 + 276\kappa^5 + 740\kappa^6 + 1956\kappa^7 + 5172\kappa^8 + ...\right) \tag{6.31}$$

Jak widać, współczynniki przy kolejnych potęgach κ rosną bardzo szybko, co odzwierciedla fakt, że wyznaczanie ich staje się coraz bardziej kłopotliwe. Zgodnie z powszechnie przyjętą opinią, wyznaczenie kolejnego współczynnika jest trudniejsze niż wyznaczenie wszystkich poprzednich razem wzięte.

Mamy więc wyznaczonych przez wzór (6.31) pierwszych osiem współczynników szeregu wysokotemperaturowego

$$\chi_0 = \sum_{n=0}^{\infty} a_n \kappa^n.$$

Najprostsza metoda znalezienia promienia zbieżności, a więc zlokalizowania osobliwości, czyli temperatury krytycznej, to metoda stosunków.



Rysunek 6.1. Wyznaczenie krytycznej temperatury metodą stosunków.

W naszym przypadku współczynniki mają następujące wartości

$$a_0 = 1, \quad a_1 = 4, \quad a_2 = 12, \quad ... \quad a_8 = 5172.$$

Biorąc stosunki następnego do poprzedniego współczynnika dostajemy

$$r_1 = \frac{a_2}{a_1} = 4, \quad r_2 = \frac{a_3}{a_2} = 3, \quad r_3 = 3, \quad r_4 = 2.77,$$
$$r_5 = 2.76, \quad r_6 = 2.681, \quad r_7 = 2.664, \quad r_8 = 2.623.$$

Wartości r_n odkładamy na osi pionowej , natomiast na osi poziomej wartości $1/n$. Punkty te możemy dopasować prostą. Przecięcie się tej prostej z osią κ^{-1} pozwala wyznaczyć krytyczną wartość κ_c^{-1}

$$\kappa_c^{-1} = 2.415 \Rightarrow \frac{2kT_c}{J} = 2.2727, \tag{6.32}$$

podczas gdy dokładna wartość temperatury krytycznej znana z pracy Onsagera to

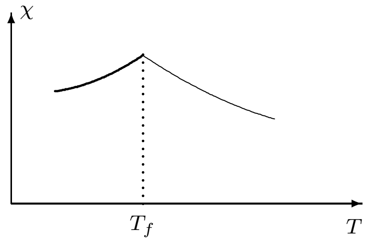
$$\frac{2kT_c}{J} = 2.270.$$

Zgodność jest więc bardzo dobra, mimo iż wykorzystaliśmy tylko osiem wyrazów rozwinięcia.

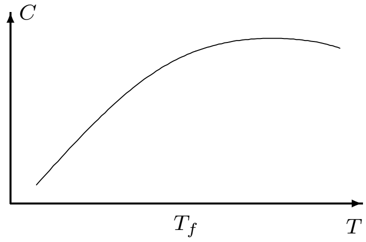
ROZDZIAŁ 7

Szkła spinowe

Szkło spinowe jest to silnie rozrzedzony stop magnetyczny, np. $AuFe$ lub $CuMn$. Typowa koncentracja atomów magnetycznych (Fe lub Mn) to $1-2$ %. Badania doświadczalne takich materiałów wykazały, że przy pewnej temperaturze, T_f, występuje ostre załamanie krzywej podatności. Temperatura ta jest nazywana **temperaturą zamarzania**. Temperatura zamarzania



RYSUNEK 7.1. Schematyczna zależność podatności od temperatury dla szkła spinowego.



RYSUNEK 7.2. Schematyczna zależność ciepła właściwego od temperatury dla szkła spinowego.

jest dość niska ($10-50$ K) a więc szybkość migracji atomów jest mała – mamy wobec tego do czynienia z problemem *quenched*. Szkło spinowe zawiera niewielkie ilości losowo rozłożonych atomów magnetycznych, które oddziałują między sobą siłami daleko-zasięgowymi. Oddziaływanie pomiędzy spinami atomów magnetycznych odbywa się poprzez elektrony atomów niemagnetycznego metalu, wpływając na ich ruch. Jest to tzw. *efekt Kondo*. Samo oddziaływanie dwu spinów można przedstawić w postaci

$$J(r) = A\left[\frac{\cos 2p_F r}{(2p_F r)^3} - \frac{\sin 2p_F r}{(2p_F r)^4}\right].$$

p_F to pęd Fermiego, zaś r to odległość spinów. Tego typu oddziaływanie nosi nazwę oddziaływania $RKKY$, on nazwisk fizyków którzy je wprowadzili (Ruderman, Kittel, Kasuya i Yosida). Rozkład atomów magnetycznych jest losowy, tak więc i oddziaływania pomiędzy momentami magnetycznymi będą losowe i to zarówno jeśli chodzi o ich wartość jak i znak. Siły te mają charakter oscylacyjny, tzn zmieniają znak wraz z odległością. Krzywa zależności ciepła właściwego wykazuje początkowo zależność liniową, osiągając maksimum dla temperatur powyżej T_f. Zarówno powyżej jak i poniżej T_f nie zaobserwowano dalekiego porządku. Istnieją natomiast, poniżej T_f, lokalne pola wywołane przez skorelowane, na niewielkich odległościach, spiny. Wypadkowe momenty „zamrożonych" grup spinów mają przypadkową orientację, z czego wynika brak dalekiego uporządkowania. Przyjmuje się więc, że w T_f nie zachodzi przejście fazowe (przynajmniej w tradycyjnym jego rozumieniu). Losowemu rozkładowi zamrożonych momentów grup spinów odpowiada układ w którym stan podstawowy jest silnie zdegenerowany – istnieje wiele lokalnych minimów energii. Tego typu sytuacja odpowiada stanowi metastabilnemu, gdyż równie prawdopodobne może być znalezienie układu w dowolnym z minimów energii. Stąd też nazwa tych układów w analogii do zwykłych szkieł, które również są układami metastabilnymi, tyle tylko że nie magnetycznymi. Pojęcie szkła spinowego jest stosunkowo nowe – zostało wprowadzone przez Mydosha i Canelliego w 1971r. Przedstawimy dalej pewne podstawowe elementy teorii szkieł spinowych ponieważ z jednej strony pełna teoria jest bardzo skomplikowana, z drugiej jednak strony pojęcia stosowane w tej teorii są wykorzystywane również w innych zastosowaniach fizyki statystycznej.

1. Metoda replik

W przypadku szkła spinowego (*spin glass – SG*) mamy układ quenched, tzn. musimy wyliczyć energię swobodną ze wzoru

$$f = \langle f\{x\}\rangle_x = -\frac{\langle \ln Z\{x\}\rangle_x}{\beta N}, \tag{7.1}$$

gdzie $\{x\}$ oznacza układ zmiennych opisujących nieporządek (tutaj atomów magnetycznych), natomiast $\langle\cdots\rangle_x$ uśrednienie po rozkładzie tych zmiennych. Zadanie to jest bardzo trudne, ponieważ wymaga obliczeń dla hamiltonianu zależnego od tych zmiennych, który ponadto nie jest translacyjnie niezmienniczy. Kłopot polega na tym, że wielkość którą chcemy uśrednić występuje pod logarytmem i tę operację należy najpierw przeprowadzić.
Pomocą w rozwiązaniu tego problemu jest tzw **metoda replik**. Oparta jest ona na tożsamości

$$\ln Z = \lim_{n\to 0} n^{-1}\left(Z^n - 1\right) \tag{7.2}$$

ponieważ dla $n \to 0$

$$Z^n = e^{n\ln Z} \approx 1 + n\ln Z \tag{7.3}$$

Stąd

$$\langle \ln Z\{x\}\rangle_x = \lim_{n\to 0} n^{-1}\left(\langle Z^n\{x\}\rangle_x - 1\right) \tag{7.4}$$

Dla dodatnich, całkowitych n możemy logarytm sumy stanów wyjściowego modelu przedstawić jako iloczyn n identycznych kopii (replik) naszego układu

$$Z^n\{x\} = \prod_{\alpha=1}^{n} Z_\alpha\{x\} = \prod_{\alpha=1}^{n} \exp\left(-\beta\mathcal{H}(x, S_i^\alpha)\right) = \exp\left(-\beta\sum_{\alpha=1}^{n}\mathcal{H}(x, S_i^\alpha)\right) \tag{7.5}$$

Z_α to funkcja rozkładu dla repliki α. Możemy teraz wprowadzić efektywny, translacyjnie niezmienniczy hamiltonian, zdefiniowany w przestrzeni $\{S_i^\alpha\}$ wszystkich replik modelu

$$Z_n \equiv \langle Z^n\{x\}\rangle_x \equiv \mathrm{Tr}_{\{S^\alpha\}}\exp\left(-\beta\mathcal{H}_{ef}(n)\right). \tag{7.6}$$

Dokładniej, jak zostanie to dalej pokazane,

$$Z_n = \mathrm{Tr}_{S^\alpha}\exp\left(\sum_{i\neq j}\left[\frac{1}{2}\beta\bar{J}\sum_{\alpha=1}^{n} S_i^\alpha S_j^\alpha + \frac{1}{4}\left(\beta\Delta\bar{J}\right)^2\sum_{\alpha\beta} S_i^\alpha S_j^\alpha S_i^\beta S_j^\beta\right]\right). \tag{7.7}$$

Wielkości $\bar{J}$ oraz $\Delta\bar{J}$ to wartość średnia i standardowe odchylenie wynikające z (założonego) gaussowskiego rozkładu wiązań

$$P(J_{ij}) = \left(2\pi\Delta\bar{J}^2\right)^{-1/2}\exp\left(\frac{-\left(J_{ij}-\bar{J}\right)^2}{2\Delta\bar{J}^2}\right). \tag{7.8}$$

Inaczej mówiąc, jeżeli model opisany jest Isingowskim hamiltonianem

$$\mathcal{H} = \frac{1}{2}\sum_{ij} J_{ij}S_iS_j, \tag{7.9}$$

to

$$Z_n = \mathrm{Tr}_{S_i^\alpha}\prod_{ij}\int dJ_{ij}P(J_{ij})\exp\left(\beta J_{ij}\sum_{\alpha=1}^{n} S_i^\alpha S_j^\alpha\right). \tag{7.10}$$

Mamy do wykonania dwie operacje uśredniania - po zmiennych spinowych (Tr_{S^α}) oraz po nieporządku wiązań (całka z rozkładem $P(J_{ij})$). Po rozłożeniu eksponenty w szereg otrzymujemy

$$Z_n = \mathrm{Tr}_{S_i^\alpha} \prod_{ij} \left[1 + \langle J_{ij} \rangle_x \beta \sum_\alpha S_i^\alpha S_j^\alpha + \frac{1}{2} \beta^2 \langle J_{ij}^2 \rangle_x \left(\sum_\alpha S_i^\alpha S_j^\alpha \right) \left(\sum_\gamma S_i^\gamma S_j^\gamma \right) + \ldots \right] \tag{7.11}$$

co, jak łatwo sprawdzić, da się zapisać w postaci

$$\begin{aligned} Z_n &= \mathrm{Tr}_{S_i^\alpha} \prod_{ij} \left\{ 1 + \beta/2 J_{c1} \mathcal{S}_{ij} + \beta^2/4 \mathcal{S}_{ij} \mathcal{S}_{ij} \left(J_{c2} + J_{c1}^2 \right) + \cdots \right\} \\ &= \mathrm{Tr}_{S_i^\alpha} \prod_{ij} \left\{ 1 + \beta/2 J_{c1} \mathcal{S}_{ij} + \beta^2/4 \mathcal{S}_{ij} \mathcal{S}_{ij} J_{c2} + \cdots + \beta^2/4 J_{c1}^2 \mathcal{S}_{ij}^2 + \cdots \right\} \\ &= \mathrm{Tr} \prod \exp \left\{ \beta/2 J_{c1} \mathcal{S}_{ij} + \beta^2/4 J_{c2} \mathcal{S}_{ij} \mathcal{S}_{ij} + \cdots \right\} \\ &= \mathrm{Tr} \exp \left(\frac{1}{2} \sum_{ij} \sum_{k=1}^{\infty} \frac{\beta^k}{k!} J_{ck} \mathcal{S}_{ij}^k \right) \end{aligned} \tag{7.12}$$

gdzie

$$J_{ck} \equiv J_{ij}^{cum}(k) \tag{7.13}$$

jest k-tym kumulantem J_{ij} :

$$J_{c1} = \langle J_{ij} \rangle_x \equiv \bar{J}, \tag{7.14}$$

$$J_{c2} = \langle J_{ij}^2 \rangle_x - \langle J_{ij} \rangle_x^2 \equiv \left(\Delta J_{ij} \right)^2, \text{ itd..} \tag{7.15}$$

Dwa pierwsze kumulanty to wielkości dobrze znane - wartość średnia i odchylenie kwadratowe, natomiast

$$\mathcal{S}_{ij} \equiv \sum_{\alpha=1}^{n} S_i^\alpha S_j^\alpha. \tag{7.16}$$

Dwa pierwsze wyrazy szeregu dają zdefiniowaną wyżej sumę stanów dla efektywnego hamiltonianu. W hamiltonianie efektywnym występują oddziaływania między replikami (dla $k > 1$). Wartość tego oddziaływania dana jest przez $1/4 \left(\Delta \bar{J} \right)^2$. Oddziaływań tych nie było w wyjściowym, nieuśrednionym hamiltonianie. Nie był on jednak translacyjnie niezmienniczy. Wyrugowanie jednej trudności spowodowało pojawienie się innej.
Ponieważ $\mathcal{H}_{ef}$ jest translacyjnie niezmienniczy i nie zawiera nieporządku, możemy stosować „klasyczne” metody fizyki statystycznej, w szczególności pole molekularne. Z grubsza rzecz biorąc sprowadza się to do wprowadzenia rozszczepień w wyrażenia wielospinowe

$$S_i^\alpha S_j^\alpha \to \langle S_i^\alpha \rangle S_j^\alpha, \tag{7.17}$$

$$S_i^\alpha S_j^\alpha S_i^\beta S_j^\beta \to \langle S_i^\alpha S_i^\beta \rangle S_j^\alpha S_j^\beta, \quad \text{itd.} \tag{7.18}$$

Odpowiada to wprowadzeniu parametrów porządku

$$m_\alpha \equiv \langle S_j^\alpha \rangle, \quad q_{\alpha\beta} \equiv \langle S_j^\alpha S_j^\beta \rangle \, \alpha \neq \beta, \quad q_{\alpha\beta\gamma} \equiv \langle S_j^\alpha S_j^\beta S_j^\gamma \rangle \, \alpha \neq \beta \neq \gamma \quad \ldots, \tag{7.19}$$

które następnie należy wyznaczyć.

Wszystko to jest jasne tak długo jak liczba replik (n) jest całkowitą liczbą dodatnią. Jednak aby stosować wyjściowy wzór zamieniający $\ln Z$ na granicę, należy przedłużyć analitycznie $\mathcal{H}_{ef}(n)$ dla rzeczywistych, dodatnich n. Prowadzi to do dwu problemów:

(1) Aby zastosować wzór wyjściowy należy zamienić kolejność granic, tzn najpierw wziąć granicę termodynamiczną $N \to \infty$, a potem $n \to 0$. Przez jakiś czas myślano, że pewne patologie w teorii szkieł spinowych (np. ujemna entropia modelu Sherringtona-Kirkpatricka) pochodzą właśnie stąd. Teraz wydaje się, że źródło kłopotów tkwi gdzie indziej.

(2) Przedłużenie analityczne $\mathcal{H}_{ef}(n)$ dla rzeczywistych n nie jest jednoznaczne dla małych n. Tak długo jak n jest całkowite, $\mathcal{H}_{ef}$ jest nieczułe na zmianę kolejności indeksów replik, nie jest to jednak oczywiste dla rzeczywistych n w granicy $n \to 0$. Prowadzi to do pojęcia złamania symetrii replik. Mówiąc najprościej, jeżeli obliczamy parametr porządku

$$q = \langle \langle S_i \rangle_T^2 \rangle_x, \tag{7.20}$$

to stosując metodę replik otrzymamy

$$q_{\alpha\beta} = \langle S_i^\alpha S_i^\beta \rangle_T, \quad (\alpha \neq \beta) \tag{7.21}$$

gdzie α i β to dwie dowolne repliki. Podobnie byłoby dla funkcji wielospinowych. Jak się okazuje, wyliczony w polu średnim parametr $q_{\alpha\beta} = \langle S_i^\alpha S_i^\beta \rangle_{\alpha \neq \beta}$ opisuje układ, który jest niestabilny w obszarze, gdzie powinien być stabilny. Aby otrzymać rozwiązanie stabilne trzeba złamać symetrię replik, czyli ich dokładną równoważność. Z oczywistych względów nie ma powodu aby wyróżniać jakąkolwiek replikę. Repliki są sobie równoważne w tym sensie, że wszystkie wielkości zależące od jednej tylko repliki, są niezależne od tego jaką replikę wybraliśmy, tzn

$$m_\alpha = m_\beta, \qquad \sum_{\alpha=1}^{n} (q_{\alpha\beta})^k = \sum_{\alpha=1}^{n} (q_{\alpha\gamma})^k \tag{7.22}$$

dla dowolnych k oraz α, β. Nie są sobie natomiast równe wszystkie $q_{\alpha\beta}$. Problem jak złamać symetrię replik nie jest jeszcze do końca wyjaśniony.

2. Modele Szkła Spinowego

Pierwszy model SG zaproponowali w 1975 roku Edwards i Anderson (EA). W najprostszej wersji zbudowany jest on ze spinów Isingowskich ($S = 1/2$) w węzłach regularnej sieci. Wiązania między spinami mają charakter losowy i ich rozkład zadany jest przez $P(J_{ij})$. W modelu tym, bez względu na zasięg oddziaływań, wszystkie kumulanty $J_{ck} \equiv 0$ dla $k \geq 3$, a więc efektywny hamiltonian ma szczególnie prostą postać

$$\beta\mathcal{H}_{ef}(n) = -\frac{1}{2}\sum_{ij} \beta J_{ij} \sum_{\alpha=1}^{n} S_i^\alpha S_j^\alpha - \frac{1}{4}\sum_{ij} (\beta\Delta J_{ij})^2 \sum_{\alpha\beta} S_i^\alpha S_j^\alpha S_i^\beta S_j^\beta. \tag{7.23}$$

Model EA rozwiązywany był w przybliżeniu pola średniego. Dla nieskończenie daleko-zasięgowych oddziaływań przybliżenie to jest ścisłe. Najprostszy wariant modelu EA to przyjęcie jako $P(J_{ij})$ rozkładu binarnego

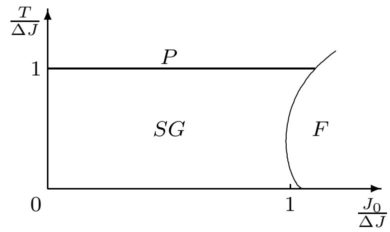
$$P(J_{ij}) = p_1\delta(J_{ij} - J) + (1 - p_1)\delta(J_{ij} + J). \tag{7.24}$$

Model taki jest bardzo uproszczony, stanowi jednak dobry punkt startowy do bardziej skomplikowanych rozważań.
W 1975 roku Sherrington i Kirkpatrick (SK) wprowadzili model w którym rozkład $P(J_{ij})$ opisuje oddziaływanie dowolnej pary spinów niezależnie od ich wzajemnej odległości. Jest więc to od początku model „średnio-polowy". Model ten rozwiązano dla równych koncentracji wiązań ferromagnetycznych i antyferromagnetycznych, oraz dla gaussowskiego rozkładu wiązań o $< J_{ij} >= J_0/N$ i szerokości $\Delta J^2/N$. Aby zachować właściwe wartości w granicy termodynamicznej należy zmienić wartości średnie na

$$\bar{J}_{ij} \propto N^{-1}, \qquad \Delta J_{ij} \propto N^{-1/2} \tag{7.25}$$

Otrzymany wykres fazowy ma postać (SG to faza szkła spinowego, P to faza paramagnetyczna, natomiast F to ferromagnetyczna – wykres jest pokazany dla połowy wartości stałej, tzn. $J > 0$)

Modele EA i SK są typu pola średniego. Ich rozwiązania, których nie będziemy tu przytaczać, sugerują, że poniżej pewnej temperatury T_f różny jest od zera parametr porządku q, który wyróżnia fazę SG od fazy paramagnetycznej. W fazie szkła spinowego znika namagnesowanie, czyli parametr dalekiego porządku, natomiast układ może znajdować się w jednym z wielu stanów, które są energetycznie prawie równoważne i różnią się tylko rozkładem tych samych wiązań. Prowadzi to do różnej orientacji przestrzennej lokalnie silnie skorelowanych grup spinów. Stany te odpowiadają



RYSUNEK 7.3. Diagram fazowy szkła spinowego w modelu SK

„dolinom" (minimom) energetycznym w przestrzeni fazowej. Ze względu na podobne wartości energii w tych dolinach, stan podstawowy układu szkła spinowego wykazuje znaczną degenerację. Nie da się jej usunąć, tak jak w ferromagnetyku, przez wprowadzenie zewnętrznego pola (w ferromagnetyku mamy dwukrotną degenerację – spiny „w górę" i „w dół"). Należy uśredniać po wszystkich tych stanach.

2.1. Model Mattisa. W 1976 roku Mattis wprowadził model, w którym losowość zawarta jest w rozkładzie węzłów

$$\mathcal{H}_M = -\frac{1}{2}\sum_{ij} J_{ij}\varepsilon_i\varepsilon_j\vec{S}_i\vec{S}_j \tag{7.26}$$

$\varepsilon_i = \pm 1$ z rozkładem losowym. $\vec{S}$ jest m-wymiarowym wektorem spinu. Jednak jak łatwo jest zobaczyć, model taki nie jest naprawdę losowy. Wprowadźmy bowiem nowe zmienne

$$\vec{\tau}_i = \varepsilon_i\vec{S}_i. \tag{7.27}$$

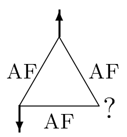
Wówczas

$$\mathcal{H}_M = -\frac{1}{2}\sum_{ij} J_{ij}\vec{\tau}_i\vec{\tau}_j. \tag{7.28}$$

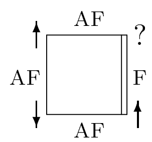
Otrzymaliśmy teraz hamiltonian tzw. *modelu m-wektorowego*, gdzie m to wymiar zmiennej spinowej. Dla $m = 1$ mamy model Isinga, dla $m = 2$ – model XY, dla $m = 3$ – model Heisenberga. Nie zawierają one żadnej losowości, nie jest to więc model SG.

3. Frustracja

Dla SG ważne jest jaką postać ma stan podstawowy, a więc jaka jest tam konfiguracja spinów. Widzieliśmy bowiem, że cechą wyróżniającą fazę SG jest silna degeneracja stanu podstawowego. Przydatna więc byłaby miara tej degeneracji. Pomocne jest tu pojęcie **frustracji** wprowadzone w 1977 roku przez G. Toulouse'a. Weźmy antyferromagnetyk typu Isinga o spinie $\frac{1}{2}$ na siatce trój-



RYSUNEK 7.4. Frustracja na sieci trójkątnej.



RYSUNEK 7.5. Frustracja na sieci kwadratowej.

kątnej. Okazuje się, że nie da się tak ułożyć spinów aby wszystkie leżały zgodnie z wymaganiami

oddziaływań. Taki elementarny trójkąt nazwiemy **sfrustrowanym**. Podobnie jest dla siatki kwadratowej o nieparzystej liczbie wiązań ferro i antyferro.
Rozpatrzmy teraz hamiltonian dla SG bez pola

$$\mathcal{H} = -\sum J_{ij} S_i S_j. \tag{7.29}$$

Nie zmieni się on po transformacji

$$S_i \to -S_i, \qquad J_{ij} \to -J_{ij}, \tag{7.30}$$

która oczywiście jest lokalna. Przy pomocy tego typu transformacji możemy np zlikwidować losowość w modelu Mattisa. Rozpatrzmy siatkę kwadratową i na niej jeden (dowolny) kwadrat o węzłach numerowanych 1, 2, 3, 4.
Odpowiedni hamiltonian możemy zapisać jako

$$\mathcal{H} = -J_{12}\varepsilon_1\varepsilon_2 S_1 S_2 - J_{23}\varepsilon_2\varepsilon_3 S_2 S_3 - J_{34}\varepsilon_3\varepsilon_4 S_3 S_4 - J_{41}\varepsilon_4\varepsilon_1 S_4 S_1. \tag{7.31}$$

Niech teraz dla wybranego kwadratu będzie np. $\varepsilon_1 = -1$, a pozostałe $\varepsilon_i = 1$. Oznacza to, że

$$\mathcal{H} = J_{12} S_1 S_2 - J_{23} S_2 S_3 - J_{34} S_3 S_4 + J_{14} S_1 S_4. \tag{7.32}$$

Dokonajmy lokalnej transformacji

$$J_{12} \to -J_{12}, \quad S_2 \to -S_2, \quad J_{14} \to -J_{14}, \quad S_4 \to -S_4. \tag{7.33}$$

Da to w efekcie

$$\mathcal{H} = J_{12} S_1 S_2 + J_{23} S_2 S_3 + J_{34} S_3 S_4 + J_{41} S_4 S_1. \tag{7.34}$$

Tak więc nieporządek został całkowicie zlikwidowany przez odpowiednią lokalną transformację, którą nazywamy *transformacją cechowania.*

Jeżeli mamy zachowaną pewną wielkość lokalną, to zgodnie z twierdzeniem Noether, odpowiada temu lokalne prawo zachowania odpowiedniego prądu i otrzymujemy równanie ciągłości. Niech teraz wielkość opisująca układ (np. funkcja Lagrange'a) będzie niezmiennicza wobec przekształcenia

$$\varphi(x) \to \varphi(x) e^{i\alpha} \tag{7.35}$$

i jeżeli przekształcenie to jest lokalne, tzn. α zależy od x, wówczas w równaniach pola (równaniu ciągłości) pojawią się dodatkowe wyrazy pochodzące od pochodnych $\alpha(x)$. Aby odzyskać niezmienniczość równań wprowadza się (na ogół wektorowe) pole kompensujące ten efekt. Jest ono generowane przez transformację cechowania i stąd jego nazwa *pole cechowania* lub *pole kompensujące.*
W elektrodynamice polem fizycznym (tzw. *polem materii*) jest pole elektrodynamiczne φ, natomiast przekształcenie (znane jako cechowanie Lorentza)

$$\varphi \to \varphi - \frac{1}{c}\frac{\partial \Lambda}{\partial t} \qquad \text{oraz} \qquad \vec{A} \to \vec{A} + \nabla\Lambda \tag{7.36}$$

pozostawia niezmiennicze równania ruchu, czyli równania Maxwella. Wektorowe pole $\vec{A}$ jest w tym przypadku polem cechowania. Λ jest dowolnym polem skalarnym.

Podobnie możemy postąpić dla SG. Lokalne przekształcenie

$$S_i \to -S_i, \qquad J_{ij} \to -J_{ij}, \tag{7.37}$$

gdzie ij to najbliżsi sąsiedzi, nie zmienia hamiltonianu. Można więc przekształcenie to traktować jako analogon transformacji cechowania (*gauge*) na sieci. Rolę pola cechowania gra teraz J_{ij}, natomiast s_i to pole materii. W elektrodynamice transformacja cechowania związana była z ciągłą grupą $U(1)$, teraz zaś mamy dyskretną grupę Z_2. Wprowadźmy nową wielkość – *funkcję frustracji*

$$\phi = \prod_c J_{ij}/|J|, \qquad (J_{ij} = \pm J) \tag{7.38}$$

będącą iloczynem wiązań wzdłuż konturu c. Jeżeli $\phi = +1$, oznacza to, że na konturze nie ma frustracji, pojawia się ona gdy $\phi = -1$. Na ogół jako kontur c bierze się elementarną komórkę, którą nazywa się *plakietką*. Pojęcie frustracji plakietki można powiązać z jej własnościami geometrycznymi. Konkretnie – jeżeli plakietka nie jest 'sfrustrowana' (czyli $\phi = +1$), wówczas ma ona podobne własności co plakietka płaska – w przeciwnym przypadku mamy do czynienia z plakietką zakrzywioną. Prostą ilustracją jest przesunięcie, zgodne z oddziaływaniami, wektora spinu wzdłuż plakietki. Jeżeli plakietka nie zawiera frustracji, wówczas kierunki wektora przed i po tej operacji będą takie same. Jeżeli natomiast plakietka posiada frustrację, kierunki te będą różne, a w przypadku spinu $\frac{1}{2}$, nawet przeciwne.

Jeżeli model zawiera frustrację, wówczas $\langle\phi\rangle_x \rightarrow 0$, przy czym może to być dążenie *słabe* $\langle\phi\rangle_x \sim e^{-L}$, lub *silne* $\langle\phi\rangle_x \sim e^{-S}$. L to obwód a S to powierzchnia otoczona konturem c. Przy wzroście frustracji mamy, dla pewnej koncentracji x_c, przejście od jednego do drugiego typu zaniku. Ponieważ „złamane" wiązanie występuje zawsze (w 2D) na granicy dwu plakietek, sfrustrowane plakietki pojawiają się parami. Każda taka para podnosi oczywiście energię (wiązania nie mogą być ułożone tak „jakby chciały"). Rozpatrzmy dla przykładu siatkę kwadratową z wiązaniami dodatnimi i ujemnymi, przy czym domieszkami są te ostatnie (jest ich mniej). Rozkład domieszek jest ustalony i nie mają one możliwości migracji, ponieważ rozpatrujemy model quenched. Pojawienie się ujemnego wiązania powoduje, że w plakietce jedno z wiązań dodatnich należy „złamać". Problemem jest które, ponieważ dla pojedynczej plakietki każdy wybór daje tę samą energię. Łącząc środki zakrzywionych plakietek poprzez „złamane" wiązania otrzymujemy „strunę", której długość jest proporcjonalna do wzrostu energii układu wywołanego frustracją. Stąd dla stabilności układu żądać należy minimum długości struny. Przy zadanej gęstości frustracji (liczbie zakrzywionych plakietek) realizowany będzie stan podstawowy odpowiadający minimalnej długości „struny". Oczywiście warunek ten nie wyznacza jednoznacznie konfiguracji i stąd otrzymujemy znaczną degenerację stanu podstawowego. Istnienie w układzie frustracji jest warunkiem koniecznym dla możliwości pojawienia się szkła spinowego. Układ z frustracją posiada bowiem wysoki stopień degeneracji stanu podstawowego. Istnieje więc wiele blisko leżących poziomów energetycznych.

ROZDZIAŁ 8

Sieci neuronowe

1. Podstawowe wiadomości

Celem zastosowania metod i pojęć fizyki statystycznej do opisu układu nerwowego jest próba uzyskania informacji o sposobie działania inteligencji, zarówno naturalnej jak i sztucznej. Oczywiście można mieć zupełnie uzasadnione wątpliwości czy tego typu podejście jest w stanie powiedzieć cokolwiek sensownego o zachowaniu się tak niesłychanie skomplikowanych tworów jakimi są układy biologiczne. Można jednak mieć nadzieję, że uda się w ten sposób znaleźć pewne ogólne prawidłowości, które mogą mieć znaczenie przede wszystkim dla opisu sztucznej inteligencji. Okazało się też, że również metody fizyki statystycznej wzbogaciły się poprzez kontakt z biofizyką.

Centralny układ nerwowy składa się z bardzo dużej liczby *neuronów* (komórek nerwowych). Dla ssaków liczba ta jest rzędu 10^{12}. Pojedynczy neuron składa się z trzech części – *ciała neuronu*, otaczających go wypustek (*dendrytów*) oraz *aksonu*. Najkrócej mówiąc, dendryty stanowią część odbiorczą neuronu – zbierają one informacje pochodzące od innych komórek. Informacje te 'spływają' do ciała neuronu, gdzie jako wynik sumy takich bodźców, może powstać impuls elektryczny, tzw. *potencjał czynnościowy*. Potencjał ten przenoszony jest, nieraz do bardzo odległych miejsc, poprzez długą wypustkę komórki nerwowej, czyli akson. Długość aksonu przekracza nieraz 1 m. Normalnie jednak sygnał z mózgu np do mięśnia nie jest przekazywany przez pojedynczy akson, lecz bierze w tym udział kilka neuronów. Przekazywanie potencjału czynnościowego odbywa się pomiędzy zakończeniem pierwszego (1) neuronu a dendrytami i ciałem neuronu 2. Strefa zmiany nosi nazwę *synapsy*. Znajduje się w niej tzw szczelina synaptyczna, poprzez którą odbywa się przekazanie sygnału. W mózgu najczęstsza jest synapsa chemiczna, gdzie przekazanie sygnału (aktywacja neuronu 2) odbywa się na drodze chemicznej poprzez wyzwolenie w obszarze presynaptycznym (1 neuron) substancji chemicznych (transmiterów), które po dotarciu do obszaru postsynaptycznego (2 neuron) wywołują powstanie potencjału czynnościowego. Potencjał czynnościowy jest to gwałtowna zmiana potencjału wnętrza błony komórkowej w stosunku do otoczenia. Prowadzi to do chwilowego przepolaryzowania obu tych obszarów, po czym, gdy ostre maksimum przemieści się do następnej komórki, różne mechanizmy, np pompy sodowe, doprowadzają napięcie w komórce do stanu pierwotnego.

Okazuje się, że komórka nerwowa może, na skutek bodźców które nadeszły do niej z innych komórek (może ich być do 10^4), albo wysłać bodziec elektryczny dalej, albo nie. Tę własność neuronu nazywa się zachowaniem „wszystko, albo nic". Wybór jednego z nich determinowany jest przez synapsy, ściślej przez ich potencjały, które są wynikiem działalności owych transmiterów chemicznych. Te zaś z kolei określane są przez potencjały komórek presynaptycznych. Dla pojedynczego neuronu jego działalność wyznaczona jest przez sumę ważoną działalności innych neuronów. Wagą są tutaj stany synapsy - mogą one być silne lub słabe, wzbudzające lub hamujące. W przypadku tych ostatnich wzbudzenie komórki presynaptycznej powoduje zanik aktywności w komórce postsynaptycznej. Cały ten proces całkowania odbywa się najprawdopodobniej w aksonie w części poprzedzającej ciało komórkowe.

Jest to oczywiście opis niesłychanie uproszczony. Pewne mechanizmy, czy też zależności, są ciągle jeszcze niezbyt jasne. Nie wiadomo też dobrze jakie mechanizmy należy uważać za najważniejsze. Dlatego też modele fizyczne, jakie się tu stosuje, mogą być bardzo odległe od rzeczywistości. Nie jest bowiem pewne, czy zawierają one wszystkie najistotniejsze składniki.

Zastosowanie fizyki statystycznej do opisu sieci neuronowych tłumaczy się zwykle tym, że w obu przypadkach rozpatrywane zagadnienia dotyczą przekazywania informacji. Ważniejszym chyba powodem jest to, że fizyka statystyczna osiągnęła niewątpliwy sukces opisując w sposób poprawny, choćby jakościowo, złożone zachowanie układu wielu ciał prowadzące do przejścia fazowego. Stąd też nadzieja, że te same metody dadzą rozsądne wyniki przy opisie bardzo wielu neuronów tworzących tak skomplikowany układ jak mózg.

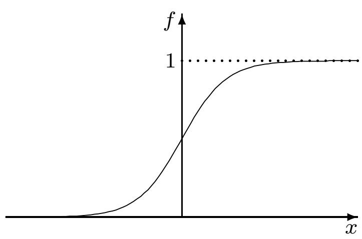
Związki neurofizjologii i fizyki statystycznej można zacząć od pracy McCullocha i Pittsa w 1943 r., którzy pierwsi zaproponowali opis neuronu jako dwustanowego elementu o pewnym progu wzbudzenia. Hebb w 1949 r. zasugerował opis procesu uczenia się jako zmianę stanu synaps. Little w 1974 r. wskazał na związek szumu synaptycznego i temperatury. Hopfield w 1982 r. wprowadził pojęcie „energii", co umożliwiło ściślejsze związki z fizyką statystyczną.

Przedstawione dalej modele oparte są o metody fizyki statystycznej. Oprócz tego istnieją i są obecnie bardziej intensywnie rozwijane, podejścia oparte o metody ściśle informatyczne i matematyczne.

2. Model fizyczny

Najprostszy model neuronu to sprowadzenie go do tzw *neuronu binarnego*, gdzie może on występować w dwu stanach – wzbudzonym i spoczynkowym. Skoro mamy funkcję dwu-wartościową, to możemy wprowadzić zmienne typu Isingowskiego, mówiąc np. że $S_i(t) = +1$ odpowiada neuronowi zlokalizowanemu w pozycji i w chwili t i znajdującemu się w stanie wzbudzonym (tzn. przekazującemu impuls). Stan synapsy między neuronami w pozycjach i i j opiszemy przez J_{ij} , przy czym $J_{ij} > 0$ odpowiada synapsie typu pobudzeniowego, natomiast $J_{ij} < 0$ – synapsie typu hamującego. Zrobiliśmy tutaj niesłychanie upraszczające założenie, że synapsy są symetryczne, tzn $J_{ij} = J_{ji}$. Wprawdzie prawie wszystkie modele fizyczne warunek ten (III prawo Newtona) spełniają, jednak synapsy – nie ! Przyjęcie symetrycznych oddziaływań umożliwia stosowanie formalizmu hamiltonowskiego, dla którego znanych jest wiele metod.

Pozostaje nam teraz zdefiniować mechanizm, który będzie decydował o tym czy neuron będzie w stanie wzbudzonym, czy nie. Opiszemy to wprowadzając funkcję $f(V_i - \Theta_i)$, która zależy od różnicy wartości potencjału V_i zebranego z innych neuronów i dochodzącego do komórki i, oraz wartości progowej Θ_i. Funkcja f ma przebieg sigmoidalny (patrz rysunek (8.1)). Szerokość ob-



RYSUNEK 8.1.

szaru, w którym f zmienia się od 0 do 1 jest miarą stochastyczności procesu i tłumaczone jest jako „szum" synapsy rozmywający, podobnie jak temperatura dla magnetyków, uporządkowanie w układzie. W przypadku granicznym ($T = 0$) mamy funkcję schodkową, co odpowiada procesowi deterministycznemu. Ponieważ jest to prostszy przypadek, dalej przyjmiemy $T = 0$.

Aby opisać pracę takiego „mózgu" trzeba zbadać jak zmienia się on w czasie, a więc określić dynamikę naszego modelu. Wiemy, że prawdopodobieństwo wzbudzenia jest równe

$$Pr(S_i = +1) = f(V_i - \Theta_i). \tag{8.1}$$

W zasadzie istnieje wiele sposobów opisu zależności potencjału V_i od aktywności pozostałych komórek. Na ogół przyjmuje się, że jest to liniowa suma ważona

$$V_i(t+\delta t) = \sum_j J_{ij} S_i(t). \tag{8.2}$$

Czas jest tu dyskretny. Przyjęliśmy, że $T = 0$, a więc f jest funkcją schodkową. Podstawiając ostatnie równanie do poprzedniego dostajemy

$$S_i(t+\delta t) = sign\left[\sum_j J_{ij} S_j(t) - \Theta_i\right]. \tag{8.3}$$

Równanie to opisuje ewolucję w czasie stanu neuronu i.

3. Model pamięci

Naszym celem jest zbudowanie maszyny, lub co na jedno wychodzi, algorytmu, który mógłby działać tak jak pamięć i który byłby:

(1) adresowalny według zawartości, odporny na zakłócenia sygnału,
(2) odporny na zniszczenie,
(3) szybki,

Aby wytłumaczyć te pojęcia posłużmy się następującym przykładem. Niech nasza pamięć będzie czarną skrzynką zawierającą „wejście" i „wyjście" oraz przełącznik umożliwiający pracę pamięci w trybie uczenia się bądź operacyjnym. Zacznijmy od trybu uczenia się. Pokazujemy skrzynce zestaw M wzorców, które ona w jakiś (nieznany) sposób przekształca w M, lub mniej, wzorców (wyjściowych). Możemy je oznaczyć odpowiednio ξ_μ^{in} i ξ_μ^{out}. Może to oznaczać, że jeżeli pokażemy na wejściu obraz pana Wojtka, to na wyjściu otrzymamy jego zdjęcie. W ten sposób skrzynka została nauczona rozpoznawania obrazów. Oczywiście możliwa jest inna organizacja pracy, możemy nauczyć skrzynkę aby do imion na wejściu dopasowywała twarze. Jeżeli teraz skrzynka znajdzie się w trybie operacyjnym, to po pierwsze, powinna ona rozpoznawać wszystkie wzorce jakich ją nauczyliśmy, oraz po drugie, jeżeli wzorzec wejściowy jest nieznacznie (dalej będzie wyjaśnione co to znaczy) zdeformowany, to skrzynka powinna mimo to właściwie rozpoznać wzorzec. Własności te noszą nazwę pamięci odpornej na zakłócenia oraz adresowalnej według zawartości, a nie tak jak w komputerze, według adresu komórki.
Odporność na zniszczenie oznacza, że nawet jeżeli uszkodzimy część skrzynki, rozpoznawanie powinno być poprawne. Możliwe jest to tylko wtedy gdy pamięć jest rozłożona, a więc każdy wzorzec znajduje się w całej sieci.

Możemy teraz zauważyć, że problem pamięci jest związany z zagadnieniem szkła spinowego. Przyjmijmy bowiem, że w danej chwili może się zmienić tylko jeden spin, oraz że $J_{ij} = J_{ji}$, co upraszcza znacznie rachunki lecz, jak wiemy, nie jest słuszne. Przy takiej interpretacji, ostatni wyprowadzony wzór opisuje relaksację spinów Isingowskich. Energia układu SG gdzie spiny oddziałują między sobą z siłą $J_{i,j}$ i z zewnętrznym polem $Theta_i$, jest równa

$$E = -\frac{1}{2}\sum_{i,j} J_{i,j} S_i S_j + \sum_i \Theta_i S_i = -\frac{1}{2}\sum_i \left(\sum_j J_{i,j} S_j - \Theta_i\right) S_i. \tag{8.4}$$

W miarę upływu czasu, zgodnie z naszym równaniem ewolucji, energia będzie malała, spiny będą 'przeskakiwać', aż osiągnięte zostanie lokalne minimum energii, tzn każdy spin będzie ustawiony równolegle w stosunku do lokalnego pola (jest nim Θ). Ważnym założeniem jakie się w tym miejscu robi jest przyjęcie, że *lokalne minima, tzn stany lokalnie stabilne, odpowiadają zapamiętanym wzorcom.* W teorii SG widzieliśmy, że stan podstawowy jest wielokrotnie zdegenerowany - posiada wiele lokalnych minimów. Odpowiada to wielu zapamiętanym wzorcom. Stąd przydatność SG, a nie np ferromagnetyka o dwu tylko stabilnych stanach. Ściśle rzecz biorąc, zagadnienie pamięci jest to odwrotny problem SG. W tym ostatnim zadane są wiązania J_{ij} (ich rozkład), a znaleźć chcemy rozkład minimów energii. W problemie pamięci chcemy ze znanego rozkładu wzorców (minimów E) znaleźć odpowiadający mu rozkład wiązań J_{ij}. Procedura dynamiczna znajdująca te wiązania opisuje proces uczenia się.

4. Model Hopfielda

Przyjmijmy, tak jak poprzednio, że

- w danej chwili zmienia znak tylko jeden spin,
- rozkład J_{ij} jest symetryczny ($J_{ij} = J_{ji}$),
- stany stabilne układu to zapamiętane wzorce.

Założenia te, ich związek ze SG oraz zastosowanie do opisu sieci neuronowej zostały podane w 1982 r przez Hopfielda. Aby model był do końca określony należy jeszcze zdefiniować Θ_i oraz J_{ij}. Hopfield przyjął $\Theta = 0$, natomiast J_{ij} wg przepisu Hebba, gdzie, jak poprzednio, $\xi_{i\mu}$ to wektory opisujące zapamiętane wzorce

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{M} \xi_{i\mu}\xi_{j\mu}, \qquad J_{ii} = 0. \tag{8.5}$$

Wstawiając to do równania na energię, dostajemy

$$E = -\frac{N}{2}\sum_{\mu=1}^{M}(m_\mu)^2 + \frac{M}{2}, \tag{8.6}$$

gdzie

$$m_\mu = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\xi_{i\mu}S_i. \tag{8.7}$$

Wielkość ta określa stopień nakrywania się konfiguracji określonej przez zmienne S_i ze wzorcem μ, zadanym przez zmienne $\xi_{i\mu}$. N jest tu liczbą neuronów, natomiast M - liczbą zapamiętanych wzorców.

Pokażemy teraz, że wzorce odpowiadają minimum energii E. Jeżeli dla ustalonego wzorca konfiguracja 'spinów' jest dowolna, wówczas

$$m_\mu \sim \frac{1}{\sqrt{N}} \quad \Rightarrow \quad E \sim 0. \tag{8.8}$$

Jeżeli natomiast $S_i \sim \xi_{i\mu}$, a więc oba zbiory prawie pokrywają się, to

$$m_\mu \sim 1 \quad \Rightarrow \quad E \sim -\frac{N}{2}, \tag{8.9}$$

co jest mniejszą wartością, czyli minimum energii odpowiada zadanym wzorcom.

W ten sposób określiliśmy model Hopfielda. Mówimy, że pewien wzorzec został zapamiętany przez sieć, jeżeli stan początkowy $S_i(0)$, który jest niezbyt odległy od wzorca ξ_ν, przejdzie, pod wpływem przyjętej przez nas poprzednio reguły dynamicznej, do stanu S_i^*, którego pokrycie z ξ_ν, jest duże, tzn $m_\nu^* \simeq 1$. Okazuje się, że sprawność, czyli zdolność do rozpoznawania wzorców, zależy od ich zapamiętanej liczby

$$\alpha = M/N. \tag{8.10}$$

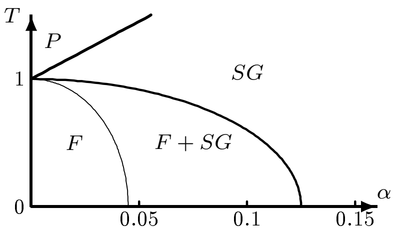
Jeżeli liczba ta jest zbyt duża, rozpoznawanie jest złe. Symulacje przeprowadzone przez Kinzela na układzie 400 spinów pokazały, że przy zapamiętanych 30 wzorcach i sytuacji wejściowej w której zniekształcenie wynosiło 30 procent (tzn. 30 procent „spinów” było „odwróconych” w stosunku do wzorca, potrzeba było 4 kroków aby układ powrócił do wzorca (litery A).

Możliwa jednak jest i taka sytuacja kiedy poziom szumu jest wysoki, tzn. małe jest pokrycie wyjściowej konfiguracji spinów i wzorca. Układ może wówczas dążyć do nieistniejącego (w pamięci) wzorca. Stąd pytania:

- jaka jest graniczna wartość α , poniżej której wzorce są stabilne, tzn basen przyciągania jest duży,

- jaki będzie efekt włączenia stochastyczności (tzn. $T > 0$),
- jaka jest natura tych nowych 'urojonych' wzorców, do których dąży układ w przypadku szumu,
- jakie są obszary przyciągania dla poszczególnych wzorców,
- jakie są granice (na ile ostre) pomiędzy tymi obszarami.

Ponieważ model Hopfielda da się zapisać w postaci hamiltonowskiej, można było próbować rozwiązać go ściśle. Zrobili to w latach 1985-87 Amit, Gutfreund i Sompolinsky, wyznaczając diagram fazowy dla modelu Hopfielda. W przybliżeniu bez złamanej symetrii replik ma on postać Faza „ferromagnetyczna" odpowiada niskim temperaturom i małej liczbie zapamiętanych wzorców



RYSUNEK 8.2. Diagram fazowy dla modelu Hopfielda.

– jest to duże przekrycie z jednym wzorcem. W fazie F + SG stabilna jest większa liczba stanów, ale też i ich przekrycie ze wzorcami jest mniejsze. W fazach SG i P nie udaje się odzyskać wzorców jeżeli układ ma szum. Są to wszystko wyniki dla modelu statycznego, a naprawdę interesuje nas model dynamiczny, tzn. poznanie tego jak wyjściowe zaburzenie wzorca będzie „przetwarzane" przez układ. Tu jednak narazie nie ma ścisłych wyników.

5. Warstwowa sieć neuronowa

Model ten, jako opis pamięci, jest alternatywą dla modelu Hopfielda. W szczególności pozwala on sprawdzić jaka jest rola założeń przyjętych w modelu Hopfielda. Warstwowa sieć neuronowa została zaproponowana w 1986 r. przez Domany'ego, Meira i Kinzela. Istnieje wiele formalnych podobieństw między obu modelami - korzystamy z funkcji dwu-wartościowych (spinów Isingowskich dla $S = \frac{1}{2}$), rozwój układu odbywa się w dyskretnym czasie zgodnie z zasadą dynamiki dla modelu Hopfielda, $\Theta = 0$, wiązania J spełniają związek Hebba. Jednak konstrukcja obu modeli jest inna. W modelu warstwowym mamy L warstw, a w każdej $i = 1, ..., N$ spinów S_i^l. Każdy spin oddziałuje ze wszystkimi spinami sąsiednich warstw, jednakże w sposób ukierunkowany, tzn stan spinu w warstwie $l + 1$ i w chwili $t + 1$ zależy od stanu wszystkich spinów w warstwie l w chwili t oraz wiązań J_{ij}. Jest więc to informacja przekazywana „w dół" czasu od danego spinu do wszystkich w „niższej czasowo" warstwie. Reguła rozwoju (zasada dynamiki) ma postać

$$S_i^{l+1} = sign\left[\sum_{j=1}^{N} J_{ij}^l S_i^l\right]. \tag{8.11}$$

Odpowiada to oddziaływaniu spinu w warstwie $l + 1$ ze wszystkimi „ przodkami" w warstwie l. Zmiana spinów w warstwie następuje jednocześnie. Reguła Hebba ma postać

$$J_{ij}^l = \frac{1}{N}\sum_{\nu=1}^{\alpha N} \xi_{j\nu}^{l+1}\xi_{i\nu}^l. \tag{8.12}$$

$\nu = 1, \ldots \alpha N$ to zapamiętane wzorce, które jak widać, zależą teraz od numeru warstwy. Jest to główna cecha różniąca ten model od modelu Hopfielda. Ma to znaczenie zarówno ideowe jak i techniczne. Ideowe, ponieważ tylko pierwsza warstwa ma tutaj kontakt ze „światem zewnętrznym"

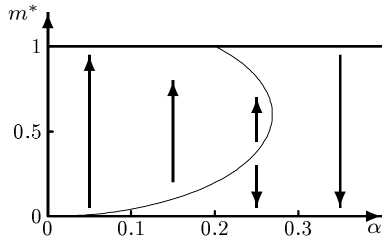
i wzorce są zadane zewnętrznie. Na głębszych poziomach układ dokonuje swobodnie wyboru reprezentacji wewnętrznej $\xi^l_{i\nu}$ wzorca. W najprostszym podejściu przyjmuje się, że wszystkie $\xi^l_{i\nu} = \pm 1$ mają to samo prawdopodobieństwo wyboru. Taki układ nie ma możliwości uczenia się.
Rozwój dynamiczny układu polega na tym, że zadany zewnętrznie stan początkowy determinuje stan w warstwie 2 w chwili $t = 1$, ten determinuje stan w warstwie 3 w chwili $t = 2$ itd. Rozwiązanie problemu polega na podaniu wzoru rekurencyjnego określającego przekrycie w chwili/warstwie, jeżeli znane jest przekrycie w poprzedniej chwili. Wzór ten ma postać

$$m^{l+1} = \frac{1}{\sqrt{2\pi}} \int_0^{a^l} e^{-\frac{y^2}{2}} dy = erf\left(\frac{m^l}{\sqrt{2\alpha g^l}}\right), \tag{8.13}$$

gdzie

$$q^{l+1} = 1 + \frac{2}{\alpha\pi} e^{-(a^l)^2} \quad , \quad a^l = \frac{m^l}{\sqrt{\alpha q^l}}. \tag{8.14}$$

W chwili początkowej m^1 jest zadane, $q^1 = 1$. W granicy dużych czasów i L osiągnięty zostanie punkt stały $m^l = m^*$, którego zależność od α i wartości początkowej m^1, ma postać jak na rysunku 8.3. Dla $\alpha \geq \alpha_c = 0.269$, istnieje tylko jedno stabilne rozwiązanie – $m^* = 0$. Dla $\alpha < \alpha_c$ pojawiają



RYSUNEK 8.3. Zależność stopnia pokrycia (m) od względnej liczby zapamiętanych wzorców (α).

się dwa dodatkowe punkty stałe, z których gałąź dla wyższych m^* jest stabilna, dla niższych m^* – niestabilna. Jak widać z rysunku 8.3, po przekroczeniu granicy α_c, wartość m^* zmienia się skokowo – mamy więc przejście fazowe I rodzaju.

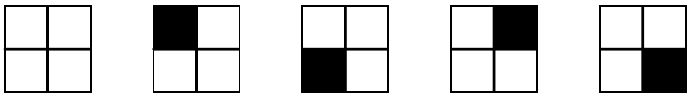
Przy symulacjach rozpatruje się układy skończone, które zamiast skokowej zmiany m^* dają zmianę ciągłą. Wraz ze wzrostem rozmiarów układu krzywe przypominają coraz bardziej wyniki analityczne. Uśrednianie przeprowadzane jest po początkowym rozkładzie wiązań. Efekty rozmiarowe są istotne w pobliżu granicy fazy, natomiast z dala od nich nawet układ o $N = 100$ daje znakomitą zgodność z krzywymi analitycznymi. W pobliżu granicy fazy jest większa możliwość, że układ będzie dążył do „złego" m^* – granica faz oznacza przecież, że stykają się ze sobą dwa baseny przyciągania dwu wzorców. Dlatego możliwość wybrania „złego" jest większa.
Podobnie jak robiliśmy to dla modelu Hopfielda wypiszemy teraz własności modelu warstwowego (*feed-forward*)

(1) wszystkie wiązania są jedno-kierunkowe,
(2) tylko komórki (spiny) w sąsiednich warstwach są ze sobą powiązane. Powiązanie to jest pełne,
(3) asymetria funkcjonalna - pierwsza warstwa to wejście, ostatnia to wyjście z sieci,
(4) stan wejściowy (1 warstwy) jest ustalany zewnętrznie,
(5) losowe wzorce są używane jako reprezentacja wejściowa, wewnętrzna i wyjściowa,
(6) z ostatniej warstwy odczytywane są wzorce wyjściowe. Może to prowadzić, podobnie jak w modelu Hopfielda, do pojawiania się niewłaściwych wzorców,
(7) Dynamika modelu da się rozwiązać ściśle.

Model ten bywa też nazywany wielowarstwowym perceptronem.

6. Perceptron

Perceptron jest to algorytm ('maszyna'), który jest w stanie sklasyfikować pokazywane wzorce zgodnie z zasadą „tak lub nie”, w stosunku do pewnego wcześniej wybranego wzorca. Maszyna taka jest w stanie nauczyć się takiego postępowania, oraz istnieje twierdzenie mówiące, że procedura uczenia się jest zbieżna.
Perceptron został wymyślony przez F. Rosenblatta w latach 1950-60. Jego idea polega na tym, że obraz jest rzutowany na ekran podzielony na kwadratowe komórki, które opisane są dwuwartościową funkcją ($\sigma_j = 0, 1$). Czynnikiem decydującym o wyborze wartości może być np. ilość światła padającego na komórkę. Następnym składnikiem perceptronu jest warstwa elementów decyzyjnych, też o wartościach 0,1, przy czym każdemu elementowi przyporządkowany jest, w sposób jednoznaczny, pewien obszar ekranu. Może to być np. kwadrat o wymiarach 2 komórki na 2. Stan każdego elementu decyzyjnego oznaczmy przez $S_i = 0, 1$. $S_i = 1$ wtedy i tylko wtedy gdy 4 komórki, za które odpowiada element i znajdują się w określonych stanach. Zwykle mówi się, że komórki te mają należeć do określonej *klasy prawdy*. Klasa taka może mieć np postać jak na rysunku 8.4



RYSUNEK 8.4. Przykładowe klasy prawdy.

Oczywiście każdy element decyzyjny może mieć swoją własną klasę prawdy i pozostaje ona stała w procesie.
Jeżeli zrzutujemy obraz na ekran to operacja taka wywoła powstanie w warstwie decyzyjnej zbioru S, złożonego z 1 i 0. Jeżeli wypadkowa (ważona) wartość S^{out} równa jest pewnej, założonej wcześniej, wartości progowej, wówczas na wyjściu otrzymamy sygnał, że pokazany obraz zgodny jest ze wzorcem. Jest on reprezentowany przez zbiór klas prawdy dla poszczególnych elementów decyzyjnych oraz wartość progową P, dopuszczalnego zniekształcenia. Możemy więc napisać

$$S^{out} = \Theta\left(\sum_i J_i S_i - P\right), \tag{8.15}$$

gdzie Θ to funkcja Heaviside'a. Wagi J_i należy tak dobrać aby perceptron optymalnie wykonał zadanie. Osiąga się to poprzez proces uczenia, który przebiega następujaco
Na początku wszystkie wiązania otrzymują wartości losowe. Przygotowujemy zbiór obrazów zawierających zarówno wzorzec, którego rozpoznawania chcemy nauczyć perceptron, jak i nie-wzorce. W przypadku wzorca na wyjściu powinniśmy otrzymać $S^{out} = 1$, natomiast dla nie-wzorca $S^{out} = 0$. Losowo wybrany obraz jest rzutowany na ekran. Na wyjściu otrzymujemy pewną wartość S^{out}. Zależnie od wyniku stosujemy jedno z trzech prawideł

(1) wynik jest prawidłowy, tzn $S^{out} = 1$ dla wzorca i $S^{out} = 0$ dla nie-wzorca. Wiązania pozostawiamy bez zmian,
(2) Perceptron „niedowidzi”, tzn. wzorzec wywołał wynik $S^{out} = 0$. Wówczas
$$J_i \to J_i + \eta S_i, \tag{8.16}$$
(3) „fałszywy alarm”, tzn. $S^{out} = 1$ gdy pokazany obraz nie był wzorcem. Wówczas
$$J_i \to J_i - \eta S_i, \qquad 0 < \eta < 1. \tag{8.17}$$

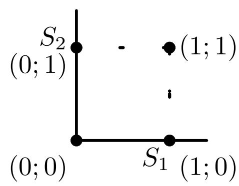
Reguła ta jest prosta do zrozumienia, jeżeli bowiem mieliśmy przypadek 1, to oznacza to iż $\sum J_i S_i$ była zbyt mała. Sumę tę należy więc zwiększyć, co najprościej można osiągnąć zwiększając wagę tych elementów gdzie $S_i = 1$. Odwrotnie postępujemy jeśli $\sum J_i S_i$ jest zbyt duża.

Można udowodnić następujące twierdzenie o zbieżności:
Jeżeli istnieje rozwiązanie J_i^**, wówczas stosując regułę uczenia się perceptronu, rozwiązanie to zostanie osiągnięte w skończonej liczbie kroków, niezależnie od wyjściowego wyboru wag.*

Jest to bardzo mocne twierdzenie, jednakże osłabia je pierwsze słowo – *jeżeli*. Okazuje się, że istnieje szereg problemów, dla których nie ma rozwiązania.
Każdy obraz jest przetwarzany przez warstwę decyzyjną na zbiór S_i, możemy więc zbiór obrazów (składający się zarówno ze wzorca jak i nie-wzorców) przedstawić jako punkty w N-wymiarowej przestrzeni. Zadaniem perceptronu jest znalezienie hiperpowierzchni, która dzieli punkty odpowiadające wzorcom i nie-wzorcom. Można to zilustrować na bardzo prostym przykładzie warunku XOR. Jeżeli mamy dwa elementy wejściowe (binarne), to wynik powinien być równy 1 wtedy i tylko wtedy gdy jeden z elementów = 1. Oznacza to, w naszym zapisie, że należy tak dobrać J_1, J_2, P aby było

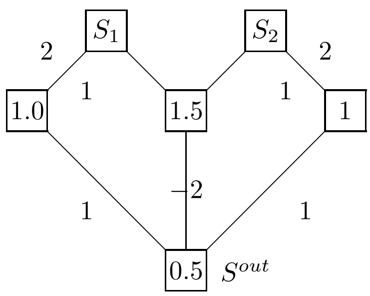
$$J_1S_1 + J_2S_2 > P. \tag{8.18}$$

Że tak być nie może, widać z diagramu w płaszczyźnie (S_1, S_2), nie można bowiem przy pomocy linii prostej oddzielić punktów $(0,0)$ i $(1,1)$ od $(0,1)$ i $(1,0)$. Można to jednak obejść, wprowadzając



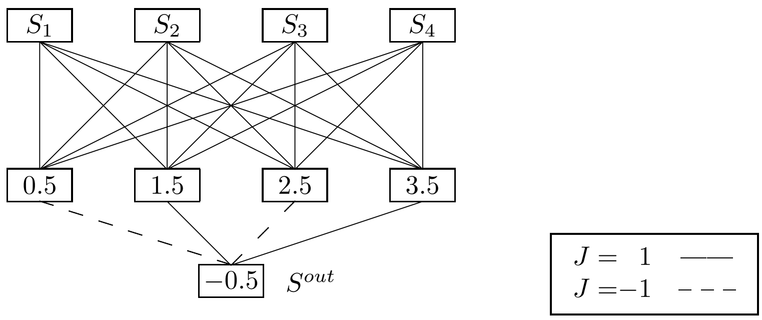
Rysunek 8.5

„ukryte jednostki", tzn. tworząc *perceptron warstwowy*.
Dla warunku XOR będzie on miał postać



Rysunek 8.6. Perceptron o dwu warstwach rozwiązujący problem XOR.

Wartości obok linii to wagi J_i. W kwadratach są wartości progów P_i.
W efekcie przestrzeń nie jest już dwu-wymiarowa i można podzielić punkty na dwie grupy. Ceną jest jednak prostota algorytmu uczenia się. Jedną z możliwości jest *algorytm powracający - back propagation.* Pokażemy go na przykładzie zagadnienia parzystości, gdy chodzi o to aby odróżnić wejście o parzystej od wejścia o nieparzystej liczbie jedynek. Problem XOR to problem parzystości dla dwu jednostek wejściowych. Teraz rozpatrzmy cztery. W ten sposób w drugiej warstwie pudełko po lewej stronie ma $S = 1$ tylko, gdy wszystkie 4 spiny są równe 1, następne – gdy trzy są równe 1 itd. Naprzemienne wartości wag gwarantują poprawny wybór. Dla wielowarstwowych perceptronów brak jest jednak ogólnego twierdzenia o zbieżności.



Rysunek 8.7. Perceptron rozwiązujący problem parzystości dla 4 jednostek wejściowych.