

Numerische Methoden in der Festkörperphysik

Prof. Dr. Nils Blümer

19. Oktober 2007

Inhaltsverzeichnis

1	Monte-Carlo-Simulationen des Ising-Modells	1
1.1	Statistische Eigenschaften von Zeitreihen	2
1.1.1	Motivation	2
1.1.2	Mathematical excursion	3
1.2	Metropolis Monte-Carlo Methode	7
1.3	Statistische Physik im kanonischen Ensemble	11
1.4	Das Ising-Modell	13
1.4.1	Mean-field solution of the Ising model	14
1.4.2	Ising model: solution in 1 dimension	15
1.4.3	Ising model on the 2d square lattice	16
1.4.4	Ising model: simulation	16
1.4.5	Metropolis importance sampling Monte Carlo scheme	18
1.4.6	Ising-Modell (NN-Wechselwirkung) - Kritische Temperaturen .	19
1.4.7	Kritische Exponenten des Ising-Modells in $d=3$	19
1.5	Finite-size scaling	19
1.6	Cluster-MC-Algorithmen	23
1.6.1	Swendsen-Wang Cluster-Update-Algorithmus	24
2	Exakte Diagonalisierung	27
2.1	Matrixdarstellung des Heisenberg-Modells	28
2.2	Mathematischer Exkurs: Eigenwertprobleme	30
2.2.1	Allgemeine Aussagen	31
2.3	Reduktion des Hilbertraums (vor ED)	32
2.3.1	Vergleich mit allgemeinem vollständig verbundenem Modell . .	39
2.4	Numerische Verfahren zur vollständigen Bestimmung aller Eigenwerte und (optional) Eigenvektoren	39
2.4.1	Naive Strategie:	39
2.4.2	Von-Mises-Verfahren (wiederholte Multiplikation mit \mathcal{H}), ggf. mit Orthogonalisierung	40
2.4.3	Iterative Diagonalisierung mit Jacobi-Rotationen	40
2.4.4	Vollständige Tridiagonalisierung symmetrischer Matrizen nach Givens	43
2.4.5	Vollständige Tridiagonalisierung symmetrischer Matrizen mit Householder-Transformationen	44

2.5	Bestimmung aller Eigenwerte und (optional) Eigenvektoren von sym-	
	meterischen Tridiagonalmatrizen	46
2.5.1	Bestimmung der EVs aus charakteristischem Polynom	46
2.5.2	Iterative Diagonalisierung einer Tridiagonalmatrix mittels QR-	
	bzw. QL-Zerlegung	47
2.6	Partielle Tridiagonalisierung mit dem Lanczos-Algorithmus	49
2.6.1	Lanczos-Tridiagonalisierung	49
3	Quantum Monte Carlo (QMC) simulations	53
3.1	Classification of QMC methods	54
3.2	Path Integral Monte Carlo (PIMC)	54
3.2.1	Discrete path integrals	56
3.2.2	Fractal nature	58
3.2.3	Role of temperature	58
3.2.4	Connection to Feynman-Kac path integral	59
3.2.5	Monte Carlo sampling of the partition function	59
3.2.6	Observables	61
3.2.7	Applicaion to vacancies in solid ^4He	62
3.3	Weltlinien-Quanten-Monte-Carlo-Methode (world line QMC)	62
3.3.1	Berechnung der Energie	66
3.4	WL-QMC mit Loop-Updates	67
3.4.1	Abbildung auf 6-Vertex-Modell	68

Kapitel 1

Monte-Carlo-Simulationen des Ising-Modells

1.1 Statistische Eigenschaften von Zeitreihen

1.1.1 Motivation

Experimentelle Messungen und Simulationen liefern im Allgemeinen nicht ein exaktes Ergebnis, sondern eine *Zeitreihe* (d.h. eine geordnete Liste) von Ergebnissen (Skalaren, Vektoren, etc.)

Problem:

Wie lassen sich Zeitreihen (hier: Skalare) charakterisieren und interpretieren? Was sagen sie über das exakte Ergebnis aus?

Beispiel 1: Eine Simulations/Messreihe liefert die Zahlen 47.5, 47.7, 47.2, 46.8 - Charakterisierung?

- alle Werte ungefähr 47
- *Mittelwert:* 47.3
- *Varianz der Einzelmessung:* 0.153
- *Mittelwert mit Standardabweichung:* 47.30 ± 0.20
- *Trend/Transient:* negativ (spätere Werte eher kleiner)
- *Histogramm:* nicht sinnvoll (zu wenige Werte)
- *Autokorrelation:* nicht erfassbar

Frage: how to get reliable results from simulations?

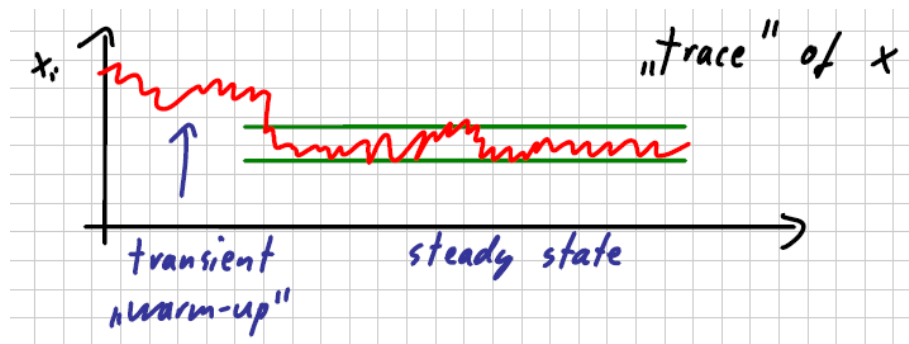
Result: estimate + error bars

- decide whether results are correct or incorrect
- essential for extrapolation (e.g. finite size)

Typical case in computer simulations:

simulations start with non-representative configurations and gradually approach steady state (e.g. thermal equilibrium)

→ both mean and variance of initial measurements can differ strongly from later ones and should be thrown away



Let us now assume that transient parts have been discarded. How to compute best estimate + error bars?

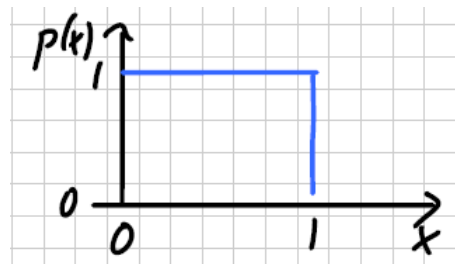
1.1.2 Mathematical excursion

let X be a random variable with *probability distribution* p :

$$P(x \leq X \leq x + dx) = p(x)dx \quad (p(x) \equiv P_X(x)) \quad (1.1)$$

example: uniform distribution

$$p(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{else} \end{cases} \quad (1.2)$$



probability distribution is normalized: $\int_{-\infty}^{\infty} p(X)dX = 1$

$$\text{mean (1st moment): } \langle X \rangle = \int_{-\infty}^{\infty} Xp(X)dX$$

$$2^{\text{nd}} \text{ moment: } \langle X^2 \rangle = \int_{-\infty}^{\infty} X^2 p(X)dX$$

$$\text{variance: } \sigma_X^2 = \langle (X - \langle X \rangle)^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2$$

The mean is always additive; the variance only for uncorrelated random variables:

$$\begin{aligned}
 \langle X + Y \rangle &= \int dX \int dY p(X, Y) (X + Y) \\
 &= \int dX X \int dY p(X, Y) + \int dY Y \int dX p(X, Y) \\
 &= \langle X \rangle + \langle Y \rangle \\
 \sigma_{X+Y}^2 &= \langle [(X + Y) - \langle X + Y \rangle]^2 \rangle \\
 &= \langle [(X - \langle X \rangle) + (Y - \langle Y \rangle)]^2 \rangle \\
 &= \sigma_X^2 + \sigma_Y^2 + \underbrace{2\langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle}_{=0 \text{ for uncorrelated variables}}
 \end{aligned} \tag{1.3}$$

Specifically for $\sigma_X = \sigma_Y \equiv \sigma$ in uncorrelated case:

$$\sigma \left[\frac{X + Y}{2} \right] = \frac{\sigma}{\sqrt{2}} \tag{1.4}$$

and in fully correlated case $X = Y$:

$$\sigma \left[\frac{X + Y}{2} \right] = \sigma \tag{1.5}$$

and for arithmetic average of N uncorrelated random numbers with equal variance σ :

$$\sigma_{\bar{X}} = \sigma \left[\frac{1}{N} \sum_{i=1}^N X_i \right] = \frac{\sigma}{\sqrt{N}} \tag{1.6}$$

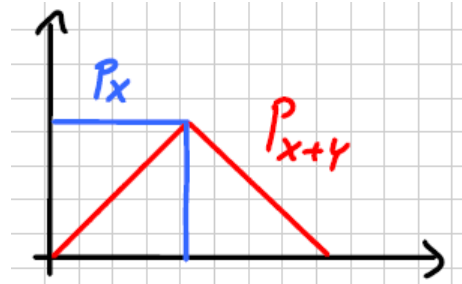
This is one formulation of the *central limit theorem*.

Full probability distribution of sum/average?

$$\begin{aligned}
 P_{X+Y}(z) &= \int dX P_X(X) \int dY P_Y(Y) \delta(z - (X + Y)) \\
 &= \int dX P_X(X) P_Y(z - X)
 \end{aligned} \tag{1.7}$$

specifically for uniformly distributed X, Y :

$$P_{X+Y}(z) = \begin{cases} z & \text{for } 0 \leq z \leq 1 \\ 2 - z & \text{" } 1 \leq z \leq 2 \\ 0 & \text{else} \end{cases} \tag{1.8}$$



Look at higher order cumulants \rightarrow normal distribution

$$P_{\bar{X}}(x) \xrightarrow{N \rightarrow \infty} \frac{1}{\sigma_{\bar{X}} \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\bar{X} - \langle X \rangle}{\sigma_{\bar{X}}} \right)^2 \right]; \quad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}} \quad (1.9)$$

central limit theorem (uncorrelated case!)

Now: application to simulation/measurements

again: random variables X_i with same distribution, non necessarily uncorrelated

Problem:

Probability distribution, mean, variance unknown. Best unbiased estimates?
easy:

$$\langle X \rangle \approx \bar{X} = \frac{1}{N} \sum_{i=1}^N X_i \quad (1.10)$$

\bar{X} is (best) unbiased estimator for $\langle X \rangle$ since

$$\langle \bar{X} \rangle = \frac{1}{N} \sum_{i=1}^N \underbrace{\langle X_i \rangle}_{\langle X \rangle} = \langle X \rangle \quad (1.11)$$

$$\left(\begin{array}{l} \text{also unbiased: } \tilde{X} = \sum_{i=1}^N a_i X_i \text{ with } \sum_{i=1}^N a_i = 1 \\ \text{also estimator: } \tilde{X} = \frac{1}{N-1} \sum_{i=1}^N X_i \text{ (since } \langle \tilde{X} \rangle = \frac{N}{N-1} \langle X \rangle \xrightarrow{N \rightarrow \infty} \langle X \rangle) \end{array} \right) \quad (1.12)$$

difficult: best estimator for error of \bar{X} ? Step (i) unbiased estimator for σ_X ?

$$\begin{aligned} \left\langle \sum_{i=1}^N (X_i - \bar{X})^2 \right\rangle &= N \langle (X_i - \bar{X})^2 \rangle \\ &= N \left\langle \left(X_i - \frac{1}{N} \sum_{j=1}^N X_j \right)^2 \right\rangle \\ &= N \left\langle \left[(X_i - \langle X \rangle) - \frac{1}{N} \sum_{j=1}^N (X_j - \langle X \rangle) \right]^2 \right\rangle \\ &= N \left\langle \left[\left(1 - \frac{1}{N} \right) (X_i - \langle X \rangle) - \frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^N (X_j - \langle X \rangle) \right]^2 \right\rangle \\ &\stackrel{\text{no. corr.}}{=} N \left[\left(1 - \frac{1}{N} \right)^2 \langle (X_i - \langle X \rangle)^2 \rangle + \frac{1}{N^2} \sum_{\substack{j=1 \\ j \neq i}}^N \langle (X_j - \langle X \rangle)^2 \rangle \right] \\ &= \frac{(N-1)^2}{N} \sigma^2 + \frac{N-1}{N} \sigma^2 \\ &= \sigma^2 \frac{[(N-1) + 1](N-1)}{N} = (N-1) \sigma^2 \end{aligned} \quad (1.13)$$

Thus,

$$\sigma_{est}^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2 \quad (1.14)$$

is an unbiased estimator for the variance σ^2 if the data is uncorrelated.

Above estimator systematically underestimates σ^2 for autocorrelated data ($\frac{1}{N}$ -effect).

Better: $\sigma_{est,imp}^2 = \frac{\sum_i (X_i - \bar{X})^2}{N - \tau}$.

Step (ii) variance of mean value \bar{X} :

$$\begin{aligned} \langle (\bar{X} - \langle X \rangle)^2 \rangle &= \left\langle \left(\frac{1}{N} \sum_{i=1}^N X_i - \langle X \rangle \right)^2 \right\rangle \\ &= \left\langle \left[\frac{1}{N} \sum_{i=1}^N (X_i - \langle X \rangle) \right]^2 \right\rangle \\ &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \langle (X_i - \langle X \rangle)(X_j - \langle X \rangle) \rangle \\ &\quad \left(\text{gen. Def.: } C_{A,B} = \frac{\langle (A - \langle A \rangle)(B - \langle B \rangle) \rangle}{\sigma_A \sigma_B} \right) \\ &= \frac{\sigma^2}{N^2} \sum_{i=1}^N \sum_{j=1}^N C_{i,j} \\ &\stackrel{\text{translationinv.}}{\approx} \frac{\sigma^2}{N^2} \sum_{i=1}^N \sum_{j=1}^N C_{i-j} \\ &= \frac{\sigma^2}{N^2} \sum_{i=1}^N \sum_{k=i-N}^{i-1} C_k \\ &\approx \frac{\sigma^2}{N^2} \sum_{i=1}^N \sum_{k=-\infty}^{\infty} C_k \\ &= \frac{\sigma^2}{N} \tau \end{aligned} \quad (1.15)$$

$$\text{autocorrelation time } \tau = \sum_{k=-\infty}^{\infty} C_k \stackrel{C_k = C_{-k}}{=} 1 + 2 \sum_{k=1}^{\infty} C_k$$

$$C_k \approx \frac{\frac{1}{N-k-1} \sum_{i=1}^{N-k} (X_i - \bar{X})(X_{i+k} - \bar{X})}{\frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2} \quad \text{autocorrelation function}$$

Beispiel Datenanalyse: Vorgegeben sind 6 Zeitreihen LINK, diese sollen analysiert werden, z.B. mit dem Statistik-Tool stats_v1.4.

1.2 Metropolis Monte-Carlo Methode

Generische Aufgabenstellung: Berechnung von Integral (oder Summe), z.B. über Teilgebiet des \mathbb{R}^d .

$$I = \int_V d^d r f(\vec{r}); \quad V \subset \mathbb{R}^d \text{ kompakt + endlich} \quad (1.16)$$

Deterministischer Lösungsansatz: wähle regelmäßiges Gitter, z.B. isotrop mit Schrittweite h

$$\vec{r}_{\vec{n}} = h\vec{n} \quad (\vec{n} \in \mathbb{Z}^d) \quad (1.17)$$

und approximiere Integral durch diskrete Approximation (numerische Integration):

$$I \approx h^d \sum_{\vec{r}_{\vec{n}} \in V} f(\vec{r}_{\vec{n}}) \quad (1.18)$$

Problem:

der relative Fehler ist proportional zur Schrittweite h^s (z.B. h oder h^2), der Aufwand t skaliert aber mit der Anzahl der Funktionsauswertungen:

$$t \approx \frac{V}{h^d} \propto h^{-d} \Rightarrow \Delta I \propto h^2 \propto t^{-\frac{2}{d}} \quad (1.19)$$

\uparrow
 fällt für große d
 nur sehr langsam ab
 bei Rechenzeit $t \rightarrow \infty$

Beispiel 2a: Bei Integration 2. Ordnung und $d = 200$ muss man für Halbierung des Fehlers $2^{100} \approx 10^{30}$ mal länger rechnen.

Alternative: *stochastischer Ansatz*

(i) *Monte Carlo (MC) ohne Gewichtung (simple Monte Carlo)*

$$\Delta I \approx I_N = V \frac{1}{N} \sum_{i=1}^N f(\vec{r}_i) \quad (1.20)$$

mit zufällig, gleichverteilt gezogenen Koordinaten (d.h. konstante Wahrscheinlichkeitsdichte für mögliche Koordinaten):

$$p(\vec{r}_i) = \frac{1}{V} \quad \text{mittels Standard-Zufallszahlgenerator} \quad (1.21)$$

Falls die Varianz σ_f^2 von f auf V existiert (und endlich ist), konvergiert die MC-Auswertung für lange Zeiten laut dem Gesetz der großen Zahl:

$$\begin{aligned} \Delta I_N &= \frac{\sigma_f}{\sqrt{N}} \xrightarrow{N \rightarrow \infty} 0 \\ \Delta I_N &\approx t^{-\frac{1}{2}} \end{aligned} \quad (1.22)$$

Schon bei $d \geq 5$ konvergiert die MC-Lösung also schneller als ein Integrationsverfahren mit Fehler $\propto h^2$.

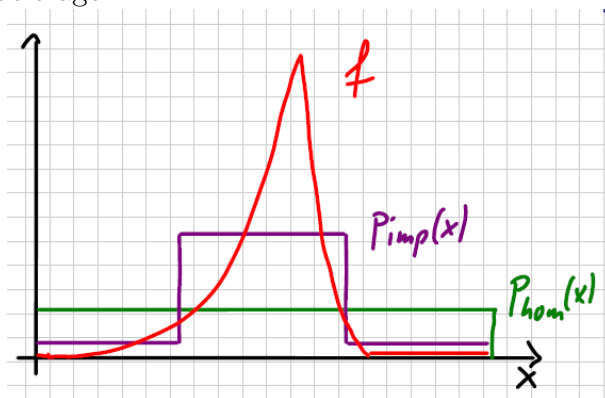
Beispiel 2b: Für Beispiel 2a ($d = 200$) lässt sich der Fehler (wie in allen Anwendungen) durch 4 Mal höhere Rechenzeit halbieren (statt Faktor 10^{30}).

Problem:

die Varianz σ_f^2 ist häufig groß, insbesondere in Anwendungen der Statistischen Physik (Schwankungen des Boltzmann-Faktors $\exp -\beta E_i$ um viele Größenordnungen).

(ii) Monte Carlo mit Gewichtung („importance sampling“)

Idee: konzentriere Stichproben \vec{r}_i auf Bereiche, die besonders stark zum Integral beitragen.



Man wählt also eine Verteilungsfunktion $p(x)$, die $|f(x)|$ approximiert.

Formal: zerlege Integrand $f(x)$ in Wahrscheinlichkeitsanteil $p(x)$ und Observablenanteil $o(x)$:

$$I = \int_V d^d r f(\vec{r}) = \int_V d^d r p(\vec{r}) o(\vec{r}) \quad (1.23)$$

$$\approx V \frac{1}{N} \sum_{i=1}^N o(\vec{r}_i); \quad \vec{r}_i \text{ zufällig mit Wahrscheinlichkeitsv. } p(\vec{r}_i) \quad (1.24)$$

beachte: Normierung von $p(\vec{r})$ muss bekannt sein!

$$\Delta I_N = \frac{\sigma_o}{\sqrt{N}}; \quad \sigma_o \ll \sigma_f \quad \text{für geeignete Wahl von } o(\vec{r}) \quad (1.25)$$

Problem:

Wie realisiert man inhomogene W.-Verteilungen?

- Variablen-Transformation
- Box-Müller für Gauß-Verteilung
- Von Neumann rejection method ineffizient
- ??? (keine allg. Methode verfügbar!)

(iii) Markow-Ketten-Monte Carlo, Metropolis Algorithmus

Idee: erzeuge Kette $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n$ von Zuständen, wobei die Wahrscheinlichkeitsverteilung $p(\vec{r}_n)$ asymptotisch gegen eine gewünschte Verteilung $p(\vec{r})$ konvergiert.

Definition: Ein *Stochastischer Prozess* ist eine Familie von Zufallsvariablen $(X_t)_{t \in I}$, wobei die Indexmenge geordnet ist: $I = \{t_0 < t_1 < \dots\}$. Für eine abzählbare Menge I (typisch $I = \mathbb{N}$) heißt der Prozess *zeitdiskret*, sonst *zeitstetig* (z.B. für $I = \mathbb{R}_0^+$).

Bemerkung: Die Wertemengen der Zufallsvariablen können in den Fällen diskret oder stetig sein.

Wichtig:: Zur vollständigen Charakterisierung eines Stochastischen Prozesses müssen neben den Wahrscheinlichkeitsverteilungen für festes t (1-Punkt-Verteilungen)

$$p_n(X_{t_n}) \equiv p(X_n, t_n) \quad (1.26)$$

i.A. auch alle n-Punkt-Verteilungen spezifiziert werden, z.B. $p(x_n, t_n; x_{n-1}, t_{n-1}; \dots; x_1, t_1)$

Beispiel 3a: Wiederholter Münzwurf; bei „Zahl“ steigt der Kontostand des Spielers um 1, bei „Wappen“ sinkt er um 1.

$$\begin{aligned} p_0(x) &= \delta_{x,0} \\ p_1(x) &= \frac{1}{2}(\delta_{x,-1} + \delta_{x,1}) \\ p_2(x) &= \frac{1}{2}(\delta_{x,-2} + 2\delta_{x,0} + \delta_{x,2}) \equiv p(x_2, 2) \\ &\dots \\ \text{aber: } p(x_n, n | x_{n-1}, n-1) &= \frac{1}{2}(\delta_{x_n, x_{n-1}-1} + \delta_{x_n, x_{n-1}+1}) \\ &\quad \uparrow \\ &\quad \text{bedingte W'} \end{aligned} \quad (1.27)$$

(Allgemein gilt: $p(x_n, n; x_m, m) = p(x_n, n | x_m, m)p(x_m, m)$)

Beispiel 3b: ein modifiziertes Spiel, in dem nach jedem Wurf das Vorzeichen des Kontostandes invertiert wird, hat war die gleiche 1-Punkt-Verteilung wie Bsp. 3a, aber andere n-Punkt-Verteilungen für $n \geq 2$.

Definition: Einen stochastischen Prozess, in dem die bedingte W'-verteilung nur vom letzten bekannten Zustand abhängt, nennt man *Markov-Prozess*:

$$p(x_n, t_n | x_{i_1}, t_{i_1}; \dots; x_{i_k}, t_{i_k}) = p(x_n, t_n | x_{i_s}, t_{i_s}) \quad (1.28)$$

für $t_{i_s} = \max_{1 \leq j \leq k} \{t_{i_j}\}$

Beispiel : Zufallsbewegungen (Random Walks) wie Bsp. 3a sind Markov-Prozesse; auch 3b ist Markov-Prozess.

Ein Markov-Prozess hat also ein minimales „Gedächtnis“; Prozesse ohne Gedächtnis nennt man unkorreliert.

Wichtiger Spezialfall: Markov-Prozesse mit stationären (zeitlich translationsinvarianten) Übergangswahrscheinlichkeiten:

$$p(x_n, t_n | x_k, t_k) = p_t(x_n | x_k); \quad t = t_n - t_k \quad (1.29)$$

Jetzt: zeitdiskreter Fall mit $t_n = n$ und stationären Übergangswahrscheinlichkeiten. Betrachte speziell einzelnen Zeitschritt: $p_1(x_{n+1} | x_n) \equiv W(x_{n+1} | x_n)$ „Übergangsrate“ Es gilt die *Master-Gleichung*:

$$p_{n+1}(x) = p_n(x) + \sum_{x'} W(x | x') p_n(x') - \sum_{x'} W(x', x) p_n(x) \quad (1.30)$$

Die W -verteilung ist genau dann *stationär*, falls sich für jedes x die Zu- und Abflüsse der Waage halten.

Hinreichende Bedingung für die Stationarität einer gewünschten Gleichgewichtsverteilung $p_{eq}(x)$ ist die

$$\text{Detailliertes Gleichgewicht: } \frac{W(x | x')}{W(x' | x)} = \frac{p_{eq}(x)}{p_{eq}(x')} \quad \forall x, x' \quad (1.31)$$

Markov-Ketten-Monte-Carlo-Algorithmen konstruieren Markov-Ketten, für die (i) eine vorgegebene W -verteilung stationär ist (detailliertes Gleichgewicht wird erfüllt) und die (ii) ergodisch sind (alle Zustände mit $p_{eq}(x) > 0$ sind erreichbar).

Vereinfachung:: betrachte in jedem Zeitschritt nur „Umgebung“ von gegebenem Zustand x . Teile dazu auf:

$$W(x' | x) = W_{\text{Vorschlag}}(x' | x) W_{\text{Akzeptanz}}(x' | x) \quad (1.32)$$

und wähle Vorschlagsw' symmetrisch: $W_{\text{Vorschlag}}(x' | x) = W_{\text{Vorschlag}}(x | x')$

Beispiel : Intervall/Hyperkubus im kontinuierlichen Fall, z.B. $W_{\text{Vorschlag}}(x' | x) = \frac{1}{a} \Theta(\frac{a}{2} - (x - x'))$ oder Einzel-Spin-Flip im Ising Modell

Dann reicht detailliertes Gleichgewicht für die Akzeptanzraten aus. Möglichkeiten:

1. $W_{\text{Akzeptanz}}(x' | x) = \min \left\{ 1, \frac{p_{eq}(x')}{p_{eq}(x)} \right\}$ **Metropolis-Regel**
2. $W_{\text{Akzeptanz}}(x' | x) = \frac{p_{eq}(x')}{p_{eq}(x) + p_{eq}(x')}$ **heat bath**

Dabei hat der Metropolis-Algorithmus den Vorteil der höchstmöglichen Akzeptanzrate, der heat bath Algorithmus stellt dagegen instantan das lokale Gleichgewicht her.



Allgemeiner Metropolis-Algorithmus (Vorschlagsw' nicht notwendig symmetrisch):

$$W_{\text{Akzeptanz}}(x'|x) = \min\{1, r\}; \quad r = \frac{p_{\text{eq}}(x')}{p_{\text{eq}}(x)} \frac{W_{\text{Vorschlag}}(x|x')}{W_{\text{Vorschlag}}(x'|x)} \quad (1.33)$$

Beachte: Im Grenzfall optimaler Vorschlagsw' werden alle Vorschläge akzeptiert.

Typische Realisierung des Metropolis-Algorithmus in der (klassischen) Statistischen Physik:

Boltzmannngewicht $p_i \propto e^{-\beta E_i}$ für Zustand i mit Energie E_i ($\beta = (k_B T)^{-1}$)

0. Initialisiere Konfiguration

1. Wähle Teilchen (in Kontinuum oder auf Gitter) bzw. Spin n aus ($1 \leq n \leq N$) - entweder zufällig oder bei jedem Durchgang (sweep) alle in fester Reihenfolge.
2. Schlage Veränderung von Teilchen n vor: Umklappen oder Verdrehen des Spins, Hüpfen oder Verschieben des Teilchens; berechne Energiedifferenz ΔE

$$r = e^{-\beta \Delta E} \frac{W_{\text{Vorschlag}}(x_{\text{neu}}|x_{\text{alt}})}{W_{\text{Vorschlag}}(x_{\text{alt}}|x_{\text{neu}})}$$

3. Akzeptiere Update mit W' $\min\{1, r\}$, sonst behalte alte Konfiguration.
4. Messe Observablen (z.B. Ausgabe in Datei)

wenn # sweeps erreicht:

5. Berechne Mittelwerte mit Standardabweichung, Verteilungen, ...

Wichtig:: (i) Die ersten Durchgänge, bei denen die W' -verteilung noch zu weit von der Stationarität entfernt ist, dürfen nicht für Observablenmittelwerte verwendet werden.

Typische Wahl: 1%-10% warm-up sweeps.

(ii) Zustandssumme/Freie Energie nicht messbar!

1.3 Statistische Physik im kanonischen Ensemble

Zustandssumme: $\text{Spur}\{e^{-\beta \mathcal{H}}\}$; $\beta = \frac{1}{k_B T}$

Hamiltonoperator \mathcal{H} , z.B. für Ising-Modell (mit Magnetfeld):

$$\mathcal{H}_{\text{Ising}} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - \mu_B B \sum_i \sigma_i \quad (1.34)$$

Im klassischen Fall bzw. für diagonalisierbares \mathcal{H} lässt sich die Zustandssumme vereinfachen:

$$Z = \sum_{\text{alle Zustände}} e^{-\beta \mathcal{H}_i} \quad (1.35)$$

Allg. Mittel/Erwartungswerte:

$$\langle O \rangle = \frac{1}{Z} \sum_i O_i e^{-\beta \mathcal{H}_i} \quad (1.36)$$

Helmholtzsche Freie Energie:

$$F = -k_B T \ln Z \quad (1.37)$$

Innere Energie:

$$\begin{aligned} E = \langle \mathcal{H} \rangle &= \frac{\sum_i \mathcal{H}_i e^{-\beta \mathcal{H}_i}}{\sum_i e^{-\beta \mathcal{H}_i}} = \frac{-\frac{\partial Z}{\partial \beta}}{Z} \\ &= -\frac{\partial(\ln Z)}{\partial \beta} = \frac{\partial(\beta F)}{\partial \beta} = F + \beta \frac{\partial F}{\partial \beta} \\ &\stackrel{\frac{\partial \beta}{\partial T} = \frac{1}{k_B T}}{=} F - T \frac{\partial F}{\partial T} = F - TS \end{aligned} \quad (1.38)$$

Entropie:

$$S = -\frac{\partial F}{\partial T} \quad (1.39)$$

Spezifische Wärme:

$$C = \frac{\partial E}{\partial T} = \frac{\partial E}{\partial \beta} \frac{\partial \beta}{\partial T} \stackrel{*}{=} \frac{\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2}{k_B T^2} \quad (1.40)$$

Sehr nützlich in MC-Kontext, erspart numerische Differentiation.
Magnetisierung (hier konkret für Ising-Modell):

$$m = \langle \mu_B \sum_i \sigma_i \rangle = \frac{1}{\beta} \frac{\partial(\ln Z)}{\partial \beta}; \quad M = \frac{m}{N \mu_B} \quad (1.41)$$

Phasenübergänge: Singularitäten thermodynamischer Größen (an Punkten/Linien/Flächen im Phasenraum)

Klassifikation nach Ehrenfest (F immer stetig)

Übergang 1. Ordnung: 1. Ableitung von F unstetig

Übergang 2. Ordnung: 1. Ableitung von F stetig, 2. Ableitung unstetig (ggf. auch divergent)

Beachte: In Systemen mit einer endlichen Zahl von Zuständen können keine Phasenübergänge auftreten:

$$Z = \sum_{s=1}^{s_{max} < \infty} e^{-\beta \mathcal{H}_s} \quad \text{ist analytisch.} \quad (1.42)$$

Daher ist bei MC-Untersuchungen zu Phasenübergängen i.A. eine sorgfältige Finite-Size-Analyse essentiell.

Kritische Exponenten charakterisieren Systeme in der Nähe von Übergängen 2. Ordnung, sind *universell*!

$$\left. \begin{array}{ll} \text{Magnetisierung} & m = m_0 \epsilon^\beta \\ \text{Suszeptibilität} & \chi = \chi_0 \epsilon^{-\gamma} \\ \text{Spez. Wärme} & C = C_0 \epsilon^{-\alpha} \\ \text{Korrelationslänge} & \xi = \xi_0 \epsilon^{-\gamma} \\ \text{Magnetisierung bei } T_c & m \propto B^{\frac{1}{\delta}} \end{array} \right\} \epsilon = \left| 1 - \frac{T}{T_c} \right| \quad (1.43)$$

1.4 Das Ising-Modell

Model of interacting quantum spins in a magnetic field, introduced as model for ferromagnetism by *Wilhelm Lenz* (\sim Runge-Lenz vector) and *Ernst Ising* (10.5.1900-11.5.1998) in Ising's PhD thesis (Hamburg 1924).

Ising later became teacher, emigrated via Luxemburg to the USA. After Onsager's solution (1944) he became professor at Bradley university, Peoria, Illinois, but never published another Journal article.

Hamiltonian of general Ising model (N lattice sites): *

$$\mathcal{H} = -\frac{1}{2} \sum_{i,j=1}^N J_{ij} \sigma_i \sigma_j - \mu_B B \sum_{i=1}^N \sigma_i \quad \sigma_i \in \{+1, -1\} \quad (1.44)$$

specifically for translation-invariant, isotropic nearest-neighbor interaction:

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - \mu_B B \sum_i \sigma_i \quad (1.45)$$

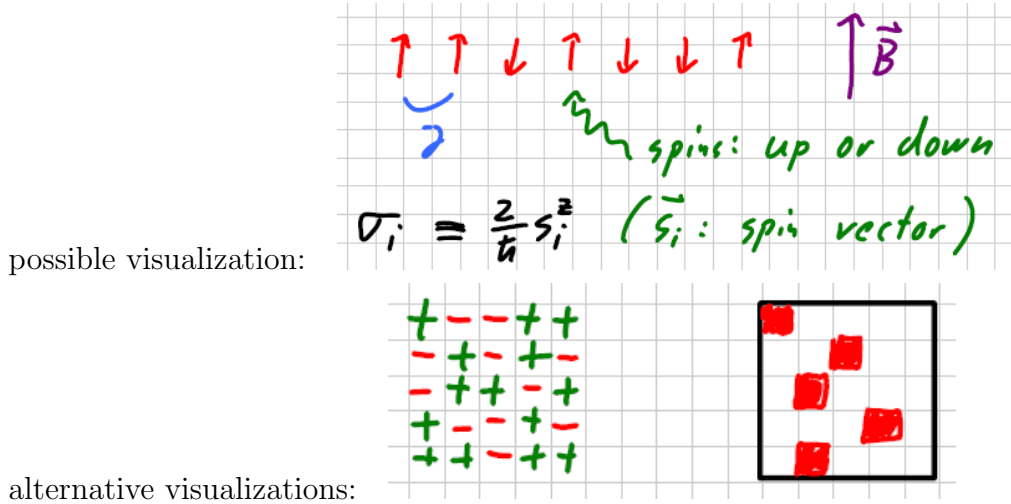
\uparrow
 sum over NN pairs,
 each pair counted once

- important special case: $B = 0$
- trivial single-spin (noninteracting) limit: $J = 0$
- ferromagnetic/antiferromagnetic coupling for $J < > 0$
- properties strongly dependent on lattice (i.e. also on dimensionality)

* more general spin Hamiltonian:

$$\mathcal{H} = -\frac{1}{2} \sum_{ij} \vec{S}_i \cdot \overleftrightarrow{I}_{ij} \cdot \vec{S}_j \quad (1.46)$$

(isotropic) Heisenberg model: $\overleftrightarrow{I}_{ij} = I_{ij} \mathbf{1}$
 Ising model: $(I_{ij})_{\mu\nu} = I_{ij} \delta_{\mu 3} \delta_{\nu 3}$
 use scaled interactions $J_{ij} = (\frac{1}{2}\hbar)^2 I_{ij}$



closely related: lattice gas model (lattice sites occupied/empty):

$$\mathcal{H}_{LGM} = U_{NN} \sum_{\langle ij \rangle} n_i n_j - \mu \sum_i n_i; \quad n_i \in \{0, 1\} \quad (1.47)$$

μ : chemical potential (in grand canonical ensemble)

1.4.1 Mean-field solution of the Ising model

Mean-field approximation: correlations of the form $\langle (\sigma_i - \langle \sigma_i \rangle)(\sigma_j - \langle \sigma_j \rangle) \rangle$ are neglected.

For homogeneous case ($\langle \sigma_i \rangle = \langle \sigma_j \rangle \equiv \langle \sigma \rangle = M$)

$$\begin{aligned} \mathcal{H} &= -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - B \sum_i \sigma_i \\ &\approx \mathcal{H}_{MF} = -J \sum_i \sigma_i \sum_{j \text{ NN. of } i} \langle \sigma \rangle - B \sum_i \sigma_i \\ &= -\underbrace{(B + qJ\langle \sigma \rangle)}_{B_{eff}} \sum_i \sigma_i \end{aligned} \quad (1.48)$$

q : coordination number (# of nearest neighbors)

effective single-spin Hamiltonian with self-consistency condition

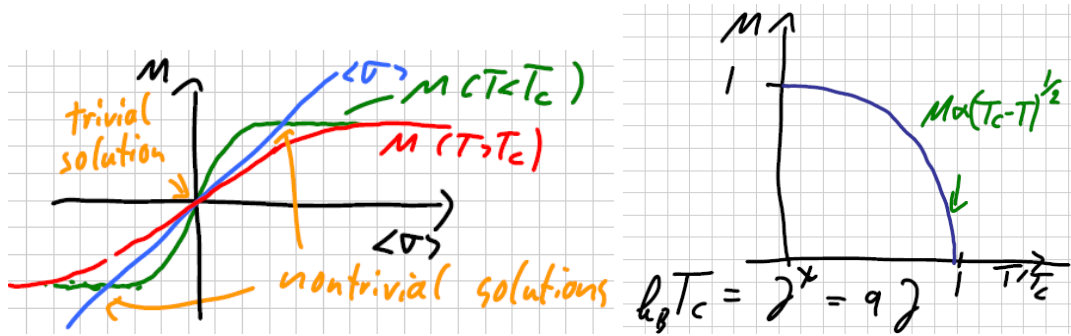
Rightarrow partition function:

$$\begin{aligned} Z &= \left(\sum_{\sigma_i = \pm 1} \exp \left(\sigma_i \frac{B + qJ\langle \sigma \rangle}{k_B T} \right) \right)^N \\ &= \left(2 \cosh \frac{B + qJ\langle \sigma \rangle}{k_B T} \right)^N \end{aligned} \quad (1.49)$$

=> magnetization:

$$M = \langle \sigma \rangle = \tanh \frac{B + qJ\langle \sigma \rangle}{k_B T} \quad (1.50)$$

graphical solution:



1.4.2 Ising model: solution in 1 dimension

(i) simple case: open chain, no magnetic field

$$\begin{aligned}\mathcal{H} &= -J \sum_{i=1}^{N-1} \sigma_i \sigma_{i+1} = -J \sum_{i=2}^N \sigma_{i-1} \sigma_i \\ Z &= \sum_{\sigma_1=\pm 1} \sum_{\sigma_2=\pm 1} \dots \sum_{\sigma_N=\pm 1} e^{-\beta \mathcal{H}\{\sigma_e\}}\end{aligned}\quad (1.51)$$

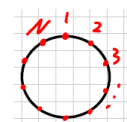
Trick: introduce new variables $\{s_i\}$ with

$$s_1 = \sigma_1; s_i = \sigma_i \sigma_{i-1} \text{ for } i \geq 2 \quad (\Rightarrow \sigma_i = \prod_{j=1}^i s_j) \quad (1.52)$$

$$\begin{aligned}\Rightarrow \mathcal{H} &= -J \sum_{i=2}^N \\ Z &= \left(\sum_{s_1=\pm 1} \right) \left(\sum_{s_2=\pm 1} e^{-\beta J s_2} \right) \left(\sum_{s_3=\pm 1} e^{-\beta J s_3} \right) \dots \left(\sum_{s_N=\pm 1} e^{-\beta J s_N} \right) \\ &= 2(2 \cosh(\beta J))^{N-1} = 2^N [\cosh(\beta J)]^{N-1} \\ \Rightarrow E(\beta) &= -\frac{\partial \ln Z}{\partial \beta} = -(N-1) \tanh(\beta J)\end{aligned}\quad (1.53)$$

$$E(T) = -(N-1)J \tanh\left(\frac{J}{k_B T}\right)$$

∞ often differentiable for all $0 < T < \infty \rightarrow$ no finite-T phase transition



(ii) more relevant case: periodic boundary conditions (+magnetic field)

$$\mathcal{H} = -J \sum_{i=1}^N \sigma_i \sigma_{i+1} - \mu_B B \sum_{i=1}^N \sigma_i \quad (1.54)$$

Treat using bond transfer matrices (2×2)

$$\Rightarrow \quad Z = (p^+)^N + (p^-)^N \text{ where} \quad (1.55)$$

$$p^\pm = e^{\beta J} \cosh(\beta \mu_B B) \pm \sqrt{e^{2\beta J} \sinh^2(\beta \mu_B B) + e^{-2\beta J}}$$

magnetization (in thermodynamic limit):

$$M = \frac{\sinh(\beta \mu_B B)}{\sqrt{\sinh^2(\beta \mu_B B) + e^{-4\beta J}}}; \quad m = N \mu_B M \quad (1.56)$$

no finite-temperature magnetism: $m \xrightarrow{B \rightarrow 0} 0$ for $T > 0$
 (but ground state fully polarized: $m = \pm N \mu_B$ for $T = 0$)

On the basis of these results, Ising and Lenz discarded the model as irrelevant for (finite T) magnetism.

1.4.3 Ising model on the 2d square lattice

Combination of high- and low-temperature expansions (*Kramers, Wannier, 1941*)
 using duality

$$\Rightarrow \frac{J}{k_B T_c} = \frac{1}{2} \operatorname{arcsinh}(1) = \frac{1}{2} \ln(\sqrt{2} + 1) \approx 0.4407 \quad (1.57)$$

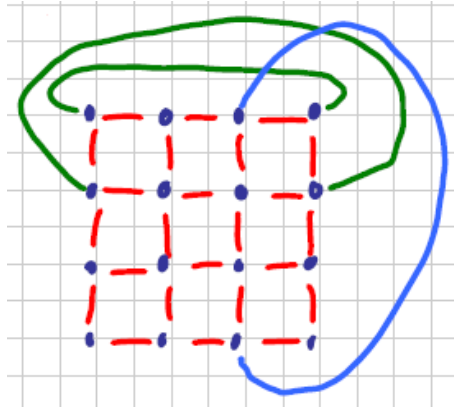
$$T_c \approx 2.2692 \frac{J}{k_B} \quad (1.58)$$

Critical exponents: $\alpha = 0$, $\beta = \frac{1}{8}$, $\gamma = \frac{7}{4}$, $\delta = 15$
 For details, see Thermodynamics script by *P. van Dongen*.

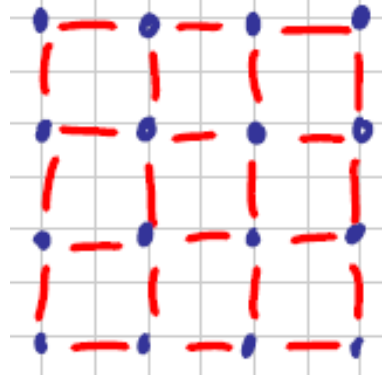
1.4.4 Ising model: simulation

1st consideration: Hamiltonian for finite-size system, i.e. *boundary conditions*
 BILD

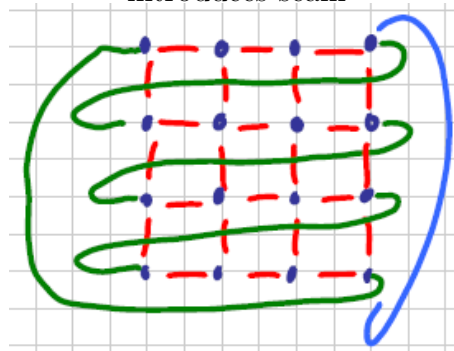
1. conventional choice: periodic boundary conditions \Rightarrow all sites equivalent



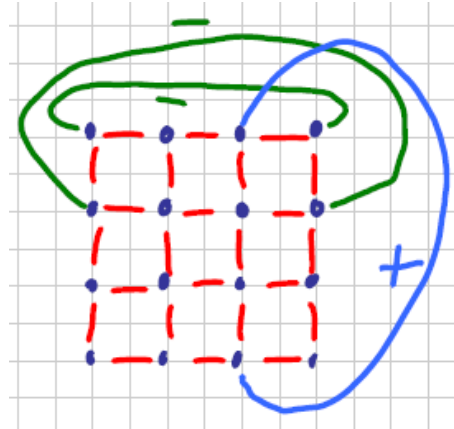
2. other extreme: open/free edge boundary conditions \Rightarrow inner + surface sites



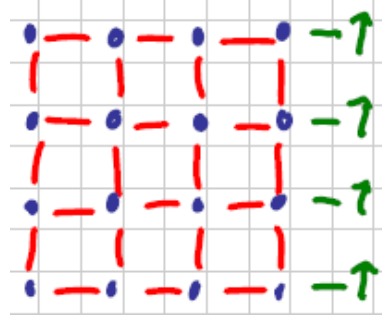
3. screw periodic boundary conditions easy to implement on $1d$ vector \Rightarrow introduces seam



4. antiperiodic boundary conditions: flip sign of interaction along some boundaries \Rightarrow generate odd number of domain walls (e.g. 1) at $T \rightarrow 1$



5. fixed or mean-field boundary conditions: boundary sites couple to external medium



Possibilities can be combined/mixed ...

1.4.5 Metropolis importance sampling Monte Carlo scheme

0. choose initial spin configuration
1. select site i
2. calculate $\Delta E = E\{\sigma_i \sigma_i \rightarrow -\sigma_i\} - E\{\sigma_i\}$
3. Metropolis step:
 - if $\Delta E < 0$: accept move
 - else generate random number $r \in [0, 1)$
 - if $r < \exp[-\frac{\Delta E}{k_B T}]$: flip spin ($\sigma_i \rightarrow -\sigma_i$)
 - else: keep state
4. after warm-up: measure
compute internally or print out for external analysis:

$$\begin{aligned}
 n_{sum} &+ 1 \\
 m_{sum} &+ \text{mag}\{\sigma_i\} \\
 E_{sum} &+ E\{\sigma_i\} \\
 &\vdots
 \end{aligned} \tag{1.59}$$

5. enough sweeps ($N = L^d$ attempted spin flips)?
 - yes: compute averages + error bars (if done internally)
 - no: return to step 1.

note: ΔE depends only on nearest neighbors of site i , specifically only on number of upspins in neighborhood (and σ_i) \rightarrow look-up table possible

1.4.6 Ising-Modell (NN-Wechselwirkung) - Kritische Temperaturen

dim	lattice	q	$\frac{k_B T_c}{J}$	$\frac{k_B T_c}{Jq}$
1	chain/ring	2	0	0
2	honeycomb	3	~ 1.52	~ 0.5
	square	4	2.269	0.57
	triangular	6	~ 3.64	~ 0.61
3	diamond	4	~ 2.704	~ 0.68
	cubic	6	~ 4.512	~ 0.75
	bcc	8	~ 6.35	~ 0.79
	fcc	12	~ 9.79	~ 0.82
4	hypercubic	8	~ 6.68	~ 0.84
∞	“	∞	∞	1.0

Tabelle 1.1: Most values taken from Peter Meyer. PhD Thesis, University of Derby (2000)

Note:

- phase transition for all lattices in $d > 1$
- $T_c \xrightarrow{q \rightarrow \infty} \frac{Jq}{k_B}$ (coordination number q)

1.4.7 Kritische Exponenten des Ising-Modells in d=3

$$\begin{aligned} \alpha &= 0.110(1); & \beta &= 0.3265(3); & \gamma &= 1.2372(5); \\ \delta &= 4.789(2); & \nu &= 0.6301(4); & \eta &= 0.0364(5); & \omega &= 0.84(4) \end{aligned}$$

Pelissetto, Vicari, Physics Reports 386, 549 (2002)

(zum Vergleich: *mean field*, d.h. $d \geq 4$: $\alpha = 0, \beta = \frac{1}{2}, \gamma = 1, \nu = \frac{1}{2}$ und nochmal $d = 2$: $\alpha = 0, \beta = \frac{1}{8}, \gamma = \frac{7}{4}, \delta = 15$)

1.5 Finite-size scaling

Ziel des *finite-size scaling (FSS)* ist die Extrapolation des kritischen Verhaltens im thermodynamischen Limes aus den (nichtsingulären) Eigenschaften endlicher Systeme.

Grundhypothese: für den singulären Anteil der Freien Energie ist nahe dem Phasenübergang 2. Ordnung die *relevante Längenskale durch die Korrelationslänge ξ gegeben*, nicht durch mikroskopische Längenskalen (mittlerer Teilchenabstand, Gitterperiode etc.). Charakteristisch für die Größe eines hyperkubischen Systems mit

Volumen $V = L^d$ ist also das Verhältnis $\frac{L}{\xi}$ von Ausdehnung zu Korrelationslänge und es gilt für die Dichte (des singulären Anteils) der Freien Energie in führender Ordnung:

$$f^{(s)}(L, T) = \frac{1}{V} F^{(s)}(L, T) \approx \frac{1}{V} \tilde{\Psi}\left(\frac{L}{\xi(T)}\right), \quad \text{wobei} \quad (1.60)$$

$$\xi(T) = \xi_0 \epsilon^{-\nu}; \quad \epsilon = \left|1 - \frac{T}{T_c}\right|$$

\Rightarrow dimensionslose Länge:

$$\frac{L}{\xi} = \xi_0^{-1} \epsilon^\nu L \quad (1.61)$$

Reskalierung:

$$\left(\frac{L}{\xi} \frac{\xi_0}{L_0}\right)^{\frac{1}{\nu}} = \epsilon \left(\frac{L}{L_0}\right)^{\frac{1}{\nu}} \quad (1.62)$$

$$\Rightarrow f^{(s)}(L, T) \approx L^{-d} Y\left(\overset{\text{nonuniversal}}{\downarrow} \underset{\text{universal function}}{\uparrow} C, \epsilon L^{\frac{1}{\nu}}, 0\right)$$

Für die Kopplung an das Magnetfeld nimmt man an, dass dieses effektiv durch eine Potenz von ξ verstärkt wird, da (für das ferromagnetische Modell) Bereiche mit Radius ξ schon nahezu parallel ausgerichtet sind:

$$\Rightarrow f^{(s)}(L, T, B) = L^{-d} Y[C, \epsilon L^{\frac{1}{\nu}}, C_2 B L^{\frac{\beta+\gamma}{\nu}}] \quad (1.63)$$

V. Privman, M.E. Fisher, PRB 30, 322 (1984)

Dabei sind alle Exponenten und die Funktion Y nur von der Universalitätsklasse abhängig (insbesondere von den Dimensionen von Raum und Ordnungsparameter). Es lassen sich nun Skalierungsgleichungen für die Observablen ableiten, z.B. (geschrieben für $L_0 \equiv 1$):

$$\begin{aligned} M(L, T) &= \tilde{M}(\epsilon L^{\frac{1}{\nu}}) L^{-\frac{\beta}{\nu}} \\ \chi(L, T) &= \tilde{\chi}(\epsilon L^{\frac{1}{\nu}}) L^{\frac{\gamma}{\nu}} \\ C(L, T) &= \tilde{C}(\epsilon L^{\frac{1}{\nu}}) L^{\frac{\alpha}{\nu}} \end{aligned} \quad (1.64)$$

Hier wurden teilweise Skalenbeziehungen zwischen den Exponenten benutzt:

$$\begin{aligned} \alpha + 2\beta + \gamma &= 2 & \text{Rushbrooke} & \quad \gamma = \beta(\delta - 1) & \text{Widom} \\ d\nu &= 2 - \alpha & \text{Josephson} & \quad \gamma = \nu(2 - \eta) & \text{Fisher} \end{aligned}$$

Die Funktionen $\tilde{M}, \tilde{\chi}, \tilde{C}$ etc. lassen sich wiederum durch Ableitungen der universellen Funktion Y und die nichtuniversellen Koeffizienten C_1, C_2 ausdrücken.

Beachte:

1. die Skalengesetze gelten nur asymptotisch für $\epsilon \rightarrow 0$ und $B \rightarrow 0$, in der Praxis können die Korrekturen (corrections to scaling) signifikant sein.
2. Im Allgemeinen kommen noch reguläre Anteile hinzu.
3. Phasenübergänge 1. Ordnung erfordern weitergehende Analyse

Praktische Anwendung: Um kritische Exponenten zu testen und nichtuniverselle Koeffizienten zu bestimmen, müsste man laut FSS skalierte Observablen gegen skalierte Parameter auftragen, z.B.

$$\tilde{M}_L(x), \text{ wobei } \tilde{M}_L = M_L L^{\frac{\beta}{\nu}}; x = \epsilon L^{\frac{1}{\nu}} \quad (1.65)$$

Dabei sollten die Kurven asymptotisch kollabieren:

$$\tilde{M}(x) = \lim_{L \rightarrow \infty} \tilde{M}_L(x) \quad (1.66)$$

jeweils separat für $T < T_c$ und $T > T_c$.

Problem:

1. die Abbildung $T \rightarrow \epsilon = |1 - \frac{T}{T_c}|$ erfordert die Kenntnis von T_c (nichtuniversell)
2. Exponenten i.A. unbekannt oder nur ungenau bekannt.

Lösung: Bestimmt T_c aus Observablen, deren Betrag nicht von L renormiert wird, die also nur im Argument von L abhängen. Beispiel: aus

$$\langle m^2 \rangle \propto (\underbrace{L^d}_{m=L^d M} L^{-\frac{\beta}{\nu}})^2 \text{ und } \langle m^4 \rangle \propto (L^d L^{-\frac{\beta}{\nu}})^4 \quad (1.67)$$

konstruiere skalierungsfreien Quotienten

$$\frac{\langle m^4 \rangle}{\langle m^2 \rangle^2} \propto (L^{d-\frac{\beta}{\nu}})^{4-2 \cdot 2} = L^0 \quad (1.68)$$

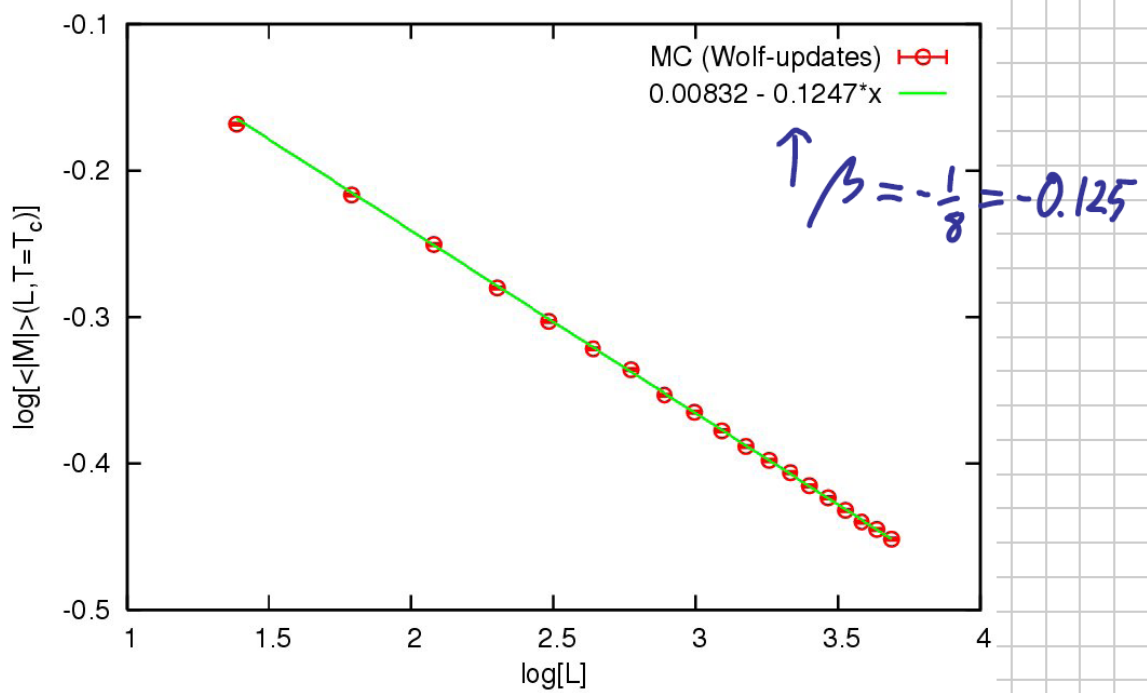
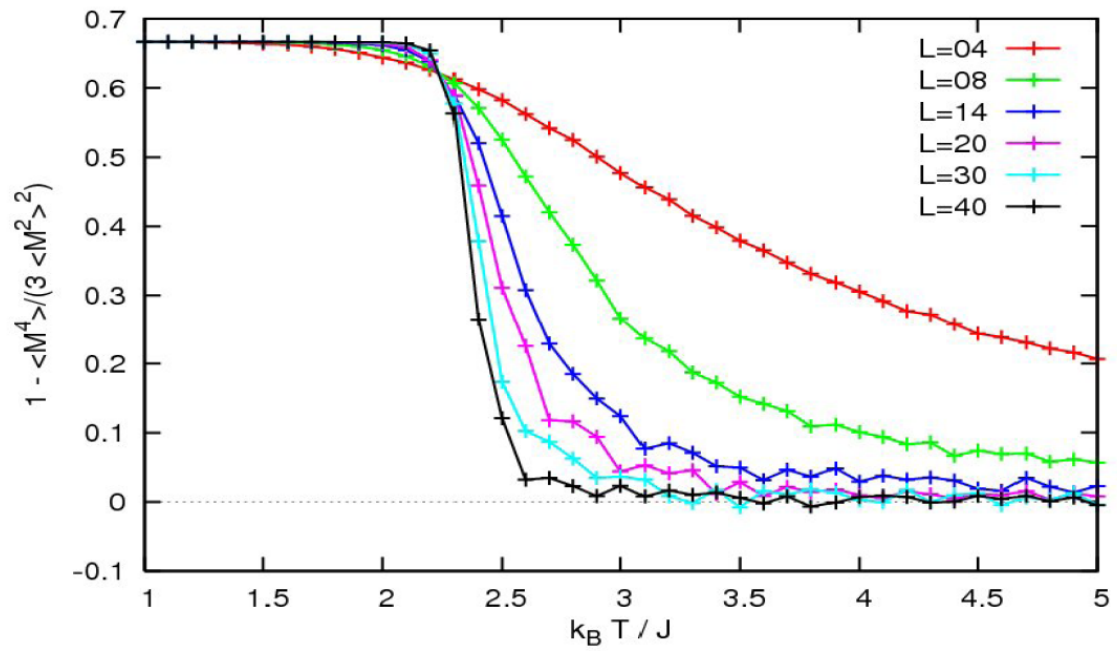
bzw. die *Binder-Kumulante* $U_4 = 1 - \frac{\langle m^4 \rangle}{4\langle m^2 \rangle^2}$ mit $U_4(L, T) = \tilde{U}_4(x)$

Trägt man nun Kurven $U_4(T)$ für verschiedenes L gemeinsam auf, schneiden sich diese asymptotisch für $T = T_c$ (und werden abseits mit verschiedenen Faktoren L horizontal gedehnt).

Beispiel für 2D Ising-Modell:

$$\text{Es gilt: } U_4 \xrightarrow{L \rightarrow \infty} \begin{cases} \frac{2}{3} & \text{für } T < T_c \\ U^* \approx 0.61 & \text{“ } T = T_c \\ 0 & \text{“ } T > T_c \end{cases}$$

Sobald T_c bekannt ist, können Exponenten z.B. aus logarithmischen Auftragungen bestimmt werden, hier für β :

Binder's cumulant (10^5 sweeps)

1.6 Cluster-MC-Algorithmen

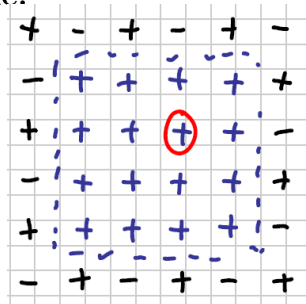
bisher: lokale Monte-Carlo-Schritte (single spin flips)

Problem:

Autokorrelationszeiten sind groß nahe T_c , divergieren (bei T_c) für $L \rightarrow \infty$: *critical slowing down*

Quantitativ: $\tau \propto \xi^z$ mit $z \approx 2.125$
↑
 dynamischer kritischer Exponent

Ursache:



große Cluster gleich ausgerichteter Spins sind stabil gegen lokale spin flips im Inneren, nur an der Oberfläche sind Akzeptanzraten signifikant:

Kosten für spin-flip im Inneren eines

großen Clusters: $\Delta E = 4DJ$

Kosten am Rand: $\Delta E = 4J$

Bei $T \approx T_c$ ändert sich die Gesamtkonfiguration daher in erster Linie durch Verschiebungen der Cluster Grenzen, was näherungsweise dem Wandern einer Hyperfläche ($D-1$ Dimensionen) in D Dimensionen bzw. einem random walk in einer Dimension entspricht. Für unabhängige Konfigurationen müssen sich die Domänenwände um eine Distanz ξ bewegt haben, was idealisiert ξ^2 Single-spin-flips erfordert ($\tau \propto \xi^z$ mit $z = 2$). Genauer: $z \approx 2.125$

Aber: der kritische Exponent ist spezifisch für die Monte-Carlo-Prozedur und kann verbessert werden!

Idee: ganze Cluster können (fast) frei umgeklappt werden ($\Delta E \propto \ln(N_{\text{Cluster}})$) \rightarrow Cluster-Updates!

Vorüberlegung: welche Cluster sollen Umklapp-Kandidaten sein? Jeweils alle zusammenhängenden Spins gleicher Orientierung? *Nein*, sonst könnten Cluster nur wachsen, jede Simulation würde zur vollen Polarisierung getrieben (Verletzung aller Gleichgewichtsbeziehungen)

Also: wie Cluster auswählen, nach welchen Regeln Spins invertieren?

Abbildung des Ising Modells (allgemeiner: q-state Potts model) auf ein Perkulations-Modell: *P.W. Kasteleyn and C.M. Fortuin, J. Phys. Soc. Jpn. Suppl. 26s, 11 (1969);*

C.M. Fortuin and P.W. Kasteleyn, Physica (Utrecht) 57,536 (1972), LINK

$$\begin{aligned}
 Z &= \sum_{\{\sigma_i\}} \exp[\beta \sum_{i,j} \sigma_i \sigma_j] \quad (J = 1) \\
 &= \sum_{\{\sigma_i\}} \prod_{\langle ij \rangle} e^{\beta \sigma_i \sigma_j} \\
 &= \sum_{\{\sigma_i\}} \prod_{\langle ij \rangle} e^{\beta [(1-p) + p \delta_{\sigma_i \sigma_j}]}; \quad p \equiv 1 - e^{-2\beta}
 \end{aligned} \tag{1.69}$$

Um uns von der Korrektheit zu überzeugen, müssen wir nur die möglichen Werte $\sigma_i \sigma_j$ einsetzen:

$$\begin{aligned}
 \sigma_i \sigma_j = 1 &\Rightarrow e^{\beta [(1-p) + p \delta_{\sigma_i \sigma_j}]} = e^{\beta} = e^{\beta \sigma_i \sigma_j} \\
 \sigma_i \sigma_j = -1 &\Rightarrow e^{\beta [(1-p) + p \delta_{\sigma_i \sigma_j}]} = e^{\beta(1-p)} = e^{\beta} [1 - (1 - e^{-2\beta})] \\
 &= e^{-\beta} = e^{\beta \sigma_i \sigma_j}
 \end{aligned} \tag{1.70}$$

Beachte: $0 \leq p \leq 1$ „riecht“ nach Wahrscheinlichkeit.

Wir formen Z weiter um; mit $\sum_{n=0}^1 [a \delta_{n,0} + b \delta_{n,1}] = a + b$ gilt:

$$Z = \sum_{\{\sigma_i\}} \sum_{\{n_{ij}\}} \prod_{\langle ij \rangle} e^{\beta [(1-p) \delta_{n_{ij},0} + p \delta_{\sigma_i \sigma_j} \delta_{n_{ij},1}]} \tag{1.71}$$

Zusätzlich zu den Spin-Variablen σ_i wurden hier Bond-Variablen n_{ij} eingeführt. Letztere können nur zwischen (Nachbar-) Spins gleicher Orientierung „gesetzt“ ($n_{ij} = 1$) sein - mit Wahrscheinlichkeit p .

Beispiel 2 Spins:

σ_1	σ_2	n_{12}	Bild	Beitrag zu Z	Bei gleichen Spins: relative W' für bond ist p
+1	+1	0	$\uparrow\uparrow$	$e^{\beta(1-p)}$	
+1	+1	1	$\uparrow\uparrow$	$e^{\beta} p$	
+1	-1	0	$\uparrow\downarrow$	$e^{\beta(1-p)}$	
+1	-1	1	$\uparrow\downarrow$	0	

Bei nicht-verbundenen Spins gleicher Beitrag: Spin-flip „kostenlos“

Letzter Fall: kein Beitrag \rightarrow darf nicht auftreten

1.6.1 Swendson-Wang Cluster-Update-Algorithmus

0. Initialisierung der Spin-Konfiguration (z.B. zufällig)

1. Bond-update: für alle Nächst-Nachbar-Paare $\langle i, j \rangle$, setze

$$n_{ij} = \begin{cases} 1 & \text{falls } \sigma_i = \sigma_j \text{ und } r < p = 1 - e^{-2\beta J} \\ 0 & \text{sonst} \end{cases}$$

2. Spin-update: Finde alle CLuster durch Bonds verbundener Spins und weise ihnen je zufällig neue Orientierung zu.
3. Messung von Observablen (nach Warm-up)
4. Weiter mit 1.

R.H. Swendsen, J.-S. Wang, Phys. Rev. Lett. 58, 86 (1987)

Man kann das Spin-Update (für feste Bond-Konfiguration) mit einem modifizierten Ising-Modell assoziieren, in dem die Paarwechselwirkung jeweils entweder verschwindet (für $n_{ij} = 0$) oder unendlich stark ferromagnetisch ist (für $n_{ij} = 1$).

Nichttrivialer Implementationsschritt: Finden der S-W-Cluster. Einfacher SW-bond-update Algorithmus mit Rekursion:

1. \forall spins i : setze `visited`[i] = 0

2. \forall spins i :

setze Variable σ_{neu} zufällig
`flip_cluster`(i, σ_{neu})

```

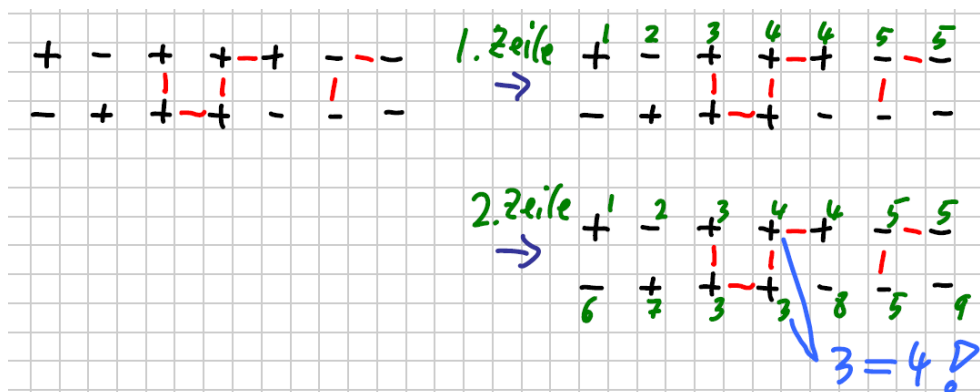
routine flip_cluster( $i, \sigma_{neu}$ )
  if visited[ $i$ ] = 0
     $\sigma[i] = \sigma_{neu}$ 
    visited[ $i$ ] = 1
     $\forall$  Nachbarn  $j$  von  $i$ 
      flip_cluster( $j, \sigma_{neu}$ )

```

Effizienter Algorithmus ohne Rekursion: „cluster multiple labeling technique“

J. Hoshen, R. Kopelman, Phys. Rev. B 14, 3438 (1976)

Prinzip: führe Cluster-Label für jeden Spin ein, das angibt, zu welchem Cluster er gehört, weise am Ende jedem Spin die richtige (Cluster-spezifische) Orientierung zu. Arbeite systematisch: zeilen+spaltenweise. etc. Dabei erhält jeder Spin, der mit einem schon besuchten Spin verbunden ist, dessen Clusternummer, sonst eine neue:



Problem:

häufig stellt sich nachträglich heraus, dass Spins mit verschiedenen Clusternummern doch verbunden sind \rightarrow Umbenennung + Buchhaltung nötig.

Effizienter und einfacher: *Wolff Cluster-Update* wie Swendsen Wang, jedoch mit Schritt

3.' Wähle zufällig einen Spin i und flippe den zugehörigen Cluster.

Vorteil: bevorzugte Updates großer Cluster.

Beide Cluster-Algorithmen reduzieren die Autokorrelationszeiten gegenüber Single-spin-flip dramatisch: $\tau \propto \xi^z$ mit $z = 0.2 - 0.5$

Besonderheit bei Swendsen-Wang: Observablen können über alle möglichen Einstellungen der Clusterspins gemittelt werden (nicht nur die, welche für die folgende Iteration benutzt wird), um die statistischen Fehler zu reduzieren.

Beispiel Suszeptibilität (ohne Subtraktion von $\langle |M| \rangle^2$):

$$\begin{aligned}
 \chi &\propto L^{2d} \langle M^2 \rangle = \left\langle \left(\sum_{i=1}^N \sigma_i \right)^2 \right\rangle \\
 &= \left\langle \left(\sum_c N_c \sigma_c \right)^2 \right\rangle \\
 &\quad \text{Summe über alle Cluster } c \text{ mit } N_c \text{ Spins, Orientierung } \sigma_c \\
 &= \left\langle \left(\sum_c N_c \sigma_c \right) \left(\sum_{c'} N_{c'} \sigma_{c'} \right) \right\rangle \\
 &= \left\langle \sum_{\{\sigma_c = \pm 1\}} \left(\sum_c N_c \sigma_c \right) \left(\sum_{c'} N_{c'} \sigma_{c'} \right) \right\rangle \\
 &= \left\langle \sum_c N_c^2 \right\rangle \quad \text{„improved estimator“}
 \end{aligned} \tag{1.72}$$

Beispiel Energie:

$$\langle \mathcal{H} \rangle = \left\langle \sum_c \langle \mathcal{H}_c \rangle \right\rangle = -J \left\langle \sum_{\langle ij \rangle} n_{ij} \right\rangle \tag{1.73}$$

Kapitel 2

Exakte Diagonalisierung

Das in Kapitel 1 betrachtete Ising-Modell ist untypisch für quantenmechanische Modelle, da die natürliche Basis $\{\sigma_i^2\}$ bereits Eigenbasis ist, also alle Energieeigenwerte im Prinzip direkt angegeben werden können. Schon bei der minimalen Erweiterung, dem *transversalen Ising-Modell* (mit Magnetfeld senkrecht zur Quantisierungsachse, z.B. B_x), gilt dies nicht mehr.

Im Allgemeinen sind Hamilton-Operatoren von Vielteilchensystemen in naheliegenden Basen des Hilbertraums also nicht diagonal, so dass schon die Bestimmung der Grundzustandsenergie höchst nichttrivial ist, umso mehr die Bestimmung aller Eigenenergien und Eigenvektoren oder von thermischen Erwartungswerten, z.B. via $Z = \text{Tr} \{e^{-\beta \mathcal{H}}\} \rightarrow \sum_i e^{-\beta E_i}$ mit Eigenzuständen i .

Allgemeinster Ansatz: *exakte Diagonalisierung von endlichen Systemen*. Dabei unterscheidet man die *vollständige Diagonalisierung* von einer teilweisen Diagonalisierung, typischerweise unter Benutzung des *Lanczos-Algorithmus*, bei dem nur extremale Eigenwerte und ggf. Eigenvektoren bestimmt werden. Eine weitere wichtige Unterscheidung bei Eigenwertproblemen (für hermitesche Matrizen) ist die zwischen dicht und dünn besetzten Matrizen; letztere ergeben sich insbesondere für Gittermodelle mit kurzreichweitiger Wechselwirkung.

2.1 Matrixdarstellung des Heisenberg-Modells

Das Heisenberg-Modell wurde 1928 von *Werner Heisenberg* als Modell für Ferromagnetismus eingeführt und 1931 von *Bethe* im 1-dimensionalen Fall gelöst („Bethe-Ansatz“, erst 1966 von *Yang+Yang* bestätigt).

Wir schreiben es (mit einer von Bethe und z.B. Baxter um den Faktor 2 abweichenden Normierung) analog zum Ising-Modell:

$$\mathcal{H}_0 = -J \sum_{\langle i,j \rangle} \vec{\sigma}_i \cdot \vec{\sigma}_j, \quad (2.1)$$

wobei $\langle i,j \rangle$ die Summe über alle nächst-Nachbar-Paare von Spins bezeichnet und $\vec{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ mit $\sigma_i^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; $\sigma_i^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$; $\sigma_i^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Beachte: in der Literatur wird meist $\vec{S}_i \equiv \vec{\sigma}_i$ verwendet (obwohl in unserer Nomenklatur $\vec{S}_i = \frac{\hbar}{2} \vec{\sigma}_i$).

Verallgemeinertes NN-Heisenberg-Modell:

$$\mathcal{H}_0 = \sum_{\langle i,j \rangle} [-J_x \sigma_i^x \sigma_j^x - J_y \sigma_i^y \sigma_j^y - J_z \sigma_i^z \sigma_j^z] \quad (2.2)$$

$J_x = J_y = 0 \rightarrow$ Ising-Modell

$J_x = J_y \neq 0 \rightarrow$ XXY-Modell

In allen Fällen lässt sich ein Magnetfeld ankoppeln: $\mathcal{H} = \mathcal{H}_0 - \mu_B \vec{B} \sum_i \vec{\sigma}_i$

Zur leichteren Handhabung formt man die nichtdiagonalen Elemente noch um. Mit

$$2\sigma^{(+)} := \sigma^x + i\sigma^y = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}; \quad 2\sigma^{(-)} := \sigma^x - i\sigma^y = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix} \quad (2.3)$$

$$\begin{aligned}\sigma_i^{(+)}\sigma_j^{(-)} &= \frac{1}{4}(\sigma_i^x + i\sigma_i^y)(\sigma_j^x - i\sigma_j^y) \\ &= \frac{1}{4}[\sigma_i^x\sigma_j^x + \sigma_i^y\sigma_j^y - i(\sigma_i^x\sigma_j^y - \sigma_i^y\sigma_j^x)]\end{aligned}\quad (2.4)$$

$$\begin{aligned}\sigma_i^{(-)}\sigma_j^{(+)} &= \frac{1}{4}[\sigma_i^x\sigma_j^x + \sigma_i^y\sigma_j^y + i(\sigma_i^x\sigma_j^y - \sigma_i^y\sigma_j^x)] \\ \text{folgt: } \vec{\sigma}_i \cdot \vec{\sigma}_j &= \sigma_i^z\sigma_j^z + \sigma_i^x\sigma_j^x + \sigma_i^y\sigma_j^y \\ &= \sigma_i^z\sigma_j^z + 2(\sigma_i^{(+)}\sigma_j^{(-)} + \sigma_i^{(-)}\sigma_j^{(+)})\end{aligned}\quad (2.5)$$

Damit erhalten wir für das isotrope Heisenberg-Modell (bei dem die Quantisierungsachse z ggf. durch ein Magnetfeld festgelegt wird):

$$\mathcal{H} = -J \sum_{\langle ij \rangle} [\sigma_i^z\sigma_j^z + 2(\sigma_i^{(+)}\sigma_j^{(-)} + \sigma_i^{(-)}\sigma_j^{(+)})] - B \sum_i \sigma_i^z \quad (2.6)$$

Bei der Anwendung auf einen der Basisvektoren

$$\{\sigma_i^z\} = |\sigma_1^z\rangle \otimes |\sigma_2^z\rangle \otimes \dots \otimes |\sigma_N^z\rangle; \quad \sigma_i \in \{\uparrow, \downarrow\} \equiv \{1, -1\} \quad (2.7)$$

liefert jeder elementare Operator ein Vielfaches eines Basisvektors mit Faktor $c \in \{-1, 0, 1\}$

$$\begin{aligned}\sigma_j^z|\sigma_1^z, \dots, \uparrow, \dots, \sigma_N^z\rangle &= |\sigma_1^z, \dots, \uparrow, \dots, \sigma_N^z\rangle \\ \sigma_j^z|\sigma_1^z, \dots, \downarrow, \dots, \sigma_N^z\rangle &= -|\sigma_1^z, \dots, \downarrow, \dots, \sigma_N^z\rangle \\ \sigma_j^{(+)}|\sigma_1^z, \dots, \uparrow, \dots, \sigma_N^z\rangle &= 0|\sigma_1^z, \dots, \uparrow, \dots, \sigma_N^z\rangle = \vec{0} \\ \sigma_j^{(+)}|\sigma_1^z, \dots, \downarrow, \dots, \sigma_N^z\rangle &= |\sigma_1^z, \dots, \uparrow, \dots, \sigma_N^z\rangle \text{ nicht-diagonal} \\ \sigma_j^{(-)}|\sigma_1^z, \dots, \uparrow, \dots, \sigma_N^z\rangle &= |\sigma_1^z, \dots, \downarrow, \dots, \sigma_N^z\rangle \text{ nicht-diagonal} \\ \sigma_j^{(-)}|\sigma_1^z, \dots, \downarrow, \dots, \sigma_N^z\rangle &= 0|\sigma_1^z, \dots, \downarrow, \dots, \sigma_N^z\rangle = \vec{0}\end{aligned}\quad (2.8)$$

Beispiel 2 Spins (offene Randbedingung):

$$H \rightarrow -J \begin{pmatrix} \uparrow\uparrow & \uparrow\downarrow & \downarrow\uparrow & \downarrow\downarrow \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} \uparrow\uparrow \\ \uparrow\downarrow \\ \downarrow\uparrow \\ \downarrow\downarrow \end{matrix} \quad \text{Block-diagonal!} \quad (2.9)$$

Allgemein haben 2×2 -Matrizen der Form $\begin{pmatrix} A & B \\ B & A \end{pmatrix}$ die Eigenvektoren $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ zum EW $A + B$ und $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ zum EW $A - B$. Unser Modell hat also die Energie-Eigenwerte $E_1 = E_2 = E_3 = -J$ (Spin-triplett) und $E_4 = 3J$ (Spin-Singulett); die Vielfachheit hätte man aus der SU(2)-Invarianz von \mathcal{H} direkt folgern können. Es kommt auf das Vorzeichen von J an, welches die Grundzustandsenergie ist, nur für das antiferromagnetische Modell ($J < 0$) ist der Grundzustand eindeutig.

Allgemeine Bestimmung von Energie-EW und EV? Schwierig!

2.2 Mathematischer Exkurs: Eigenwertprobleme

Literatur:

- Stoer, Bulirsch, Numerische Mathematik 2, Springer
- Press, Teukolsky, Vetterling, Flannery, NUmberical Recipes (in C), Cambridge University Press

Einführung: Eine $N \times N$ -Matrix A hat den (*Rechts-*) *Eigenvektor* (EV) $\vec{x} \neq \vec{0}$ zum *Eigenwert* (EW) λ , falls gilt:

$$\begin{aligned} A\vec{x} &= \lambda\vec{x} \Leftrightarrow (A - \lambda\mathbf{1}_N)\vec{x} = \vec{0} \\ &\Rightarrow |A - \lambda\mathbf{1}| := \det(A - \lambda\mathbf{1}) = 0 \end{aligned} \quad (2.10)$$

Notwendige Bedingung für die Existenz eines EV zum EW λ_0 ist also eine Nullstelle des *charakteristischen Polynoms* $P(\lambda)$ an der Stelle λ_0 . Umgekehrt existiert für jede solche Nullstelle λ_i mindestens 1 zugehöriger EV. Folglich existiert z.B. im Fall von N verschiedenen Nullstellen eine vollständige Basis von (i.A. nicht orthogonalen) Rechts-EVs (bei mehrfachen Nullstellen können dagegen i.A. zusätzlich *Hauptvektoren* auftreten).

Definition: Eine $N \times N$ -Matrix $A = (a_{ij})$ heißt

reell,	falls	$a_{ij}^* = a_{ij}$	$\forall 1 \leq i, j \leq N$
symmetrisch,	"	$A^T = A \Leftrightarrow a_{ji} = a_{ij}$	"
hermitesch,	"	$A^\dagger = A \Leftrightarrow a_{ji}^* = a_{ij}$	"
orthogonal,	"	$A^T A = A A^T = \mathbf{1}$	
unitär	"	$A^\dagger A = A A^\dagger = \mathbf{1}$	
normal	"	$A^\dagger A = A A^\dagger$	

Dabei heisst $A^T = (a_{ji})$ die *Transponierte* von A , $A^\dagger = (a_{ji}^*)$ die *hermitesch Konjugierte* von A .

Offensichtlich gilt für reelle Matrizen: symmetrisch $\hat{=}$ hermitesch, orthogonal $\hat{=}$ unitär. Neben hermiteschen und unitären Matrizen sind z.B. auch *schiefhermitesche* Matrizen ($a_{ji}^* = -a_{ij}$) normal.

Unitäre Matrizen lassen die *Länge* $|\vec{x}| = \sqrt{\vec{x} \cdot \vec{x}}$ eines Vektors invariant:

$$|\vec{x}|^2 = \vec{x}^T \vec{x} = \vec{x}^T A^T A \vec{x} = (A\vec{x}) \cdot (A\vec{x}) = |A\vec{x}|^2 \quad (2.11)$$

Weiterhin heisst eine $N \times N$ -Matrix $A = (a_{ij})$

singulär,	falls	$\det A = 0$
diagonal,	"	$A = \text{diag}\{\vec{d}\} \Leftrightarrow a_{ij} = d_i \delta_{ij} \forall i, j$
tridiagonal,	"	$a_{ij} = 0$ für $ i - j > 1$
obere Dreiecksmatrix,	"	$a_{ij} = 0$ für $i > j$
Hessenbergmatrix,	"	$a_{ij} = 0$ für $i > j + 1$

Für eine nichtsinguläre $N \times N$ -Matrix T heisst die Abbildung $A \rightarrow T^{-1}AT$ eine Ähnlichkeitstransformation. A und $T^{-1}AT$ heissen *ähnlich*.

2.2.1 Wichtige allgemeine Aussagen

1. Die EW einer Dreiecksmatrix sind die Diagonalelemente (Beweis: $p(\lambda) = \prod_i (a_{ii} - \lambda)$).
2. Eine Ähnlichkeitstransformation lässt das char. Polynom (und damit alle Eigenwerte) invariant:

$$\begin{aligned} \det |T^{-1}AT - \lambda \mathbf{1}| &= \det |T^{-1}AT - T^{-1}\lambda \mathbf{1}T| \\ &= \det |T^{-1}(A - \lambda \mathbf{1})T| \\ &= (\det |T|)^{-1} \det |A - \lambda \mathbf{1}| \det |T| \\ &= \det |A - \lambda \mathbf{1}| \end{aligned} \quad (2.12)$$

3. *Satz von Schur*: Jede $N \times N$ -Matrix A ist *unitär ähnlich* zu einer oberen Dreiecksmatrix R , d.h. für jedes A existiert eine unitäre Matrix U

$$\text{mit} \quad U^\dagger AU = R; \quad R = \begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix} \quad (2.13)$$

4. Jede $N \times N$ -Matrix A ist ähnlich zu ihrer *Jordan-Normalform*, einer Block-Diagonalmatrix mit Blöcken der Gestalt $C_\nu^\lambda = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$

5. Eine $N \times N$ -Matrix A ist genau dann normal, wenn es eine unitäre Matrix U gibt mit

$$U^{-1}AU = U^\dagger AU = D = \text{diag}\{(\lambda_1, \lambda_2, \dots, \lambda_n)\}, \quad (2.14)$$

d.h. (nur) normale Matrizen haben eine vollständige Basis von orthonormalen Eigenvektoren, den Spalten von U .

6. Alle EW einer hermiteschen Matrix A sind reell (Beweis: $(U^\dagger AU)^\dagger = U^\dagger A^\dagger U = U^\dagger AU$ ist wieder hermitesch, insbesondere für U nach 5.; jede hermitesche Diagonalmatrix ist jedoch reell.)

In der Physik interessieren wir uns vorwiegend für EWs und EVs von hermiteschen Matrizen; viele Lösungsverfahren sind jedoch allgemeiner (z.B. Tridiagonalmatrix \rightarrow Hessenbergmatrix).

Selbst für hermitesche $N \times N$ -Matrizen ist das EW-Problem höchst nichttrivial: für $N \geq 5$ gibt es keine allgemeine analytische Lösung, d.h. kein Verfahren, das in endlich vielen Schritten exakte Eigenwerte liefert.

Bevor wir praktische Methoden zur Lösung von Eigenwert-Problemen behandeln, wollen wir noch einen scheinbaren Widerspruch auflösen. Für eine $N \times N$ *Zufallsmatrix* Z (z.B. alle z_{ij} unabhängige gleichverteilte Zufallszahlen aus Intervall $[0, 1]$)

erwarten wir mit Wahrscheinlichkeit $P_1 = 1$, dass alle N Nullstellen von $p(\lambda)$ verschieden sind (da die Unterräume mit $\lambda_i = \lambda_j$ für ein Paar $i \neq j$ vom Maß 0 sind). Laut der letzten Vorlesung existiert dann eine vollständige Basis von Eigenvektoren $\{\vec{x}_1, \dots, \vec{x}_N\}$. Andererseits ist $Z^\dagger Z - Z Z^\dagger \stackrel{i.A.}{\neq} 0$, mit $P_2 = 1$ ist Z also nicht-normal, d.h. nicht diagonalisierbar!

Lösung: Die EV $\vec{x}_1, \dots, \vec{x}_N$ sind i.A. nur Rechts-EV und nicht orthogonal. Jeder Rechts-EV \vec{x}_i^R ist i.A. nur zu Links-EV \vec{x}_j^L zu verschiedenem EW orthogonal:

$$\begin{aligned} \lambda_i \vec{x}_i^L \cdot \vec{x}_j^R &= \vec{x}_i^{L^T} A \vec{x}_j^R = \vec{x}_i^{L^T} \vec{x}_j^R \lambda_j = \lambda_j \cdot \vec{x}_i^L \cdot \vec{x}_j^R \\ &\Rightarrow (\lambda_i - \lambda_j) \vec{x}_i^L \cdot \vec{x}_j^R = 0 \quad \forall i, j \end{aligned} \quad (2.15)$$

Beispiel :

1. $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \Rightarrow \lambda_1 = \lambda_2 = 2$, einziger EV $\vec{x}_1^R = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
2. $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \Rightarrow \lambda_1 = 1; \lambda_2 = 2; \vec{x}_1^R = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \vec{x}_2^R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$
 $\vec{x}_1^L = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \vec{x}_2^L = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

2.3 Reduktion des Hilbertraums (vor ED)

Ein effizienter Einsatz von ED-Methoden auf Physikalische Probleme setzt in der Regel eine Reduktion des Hilbertraums, also der Dimension der jeweils zu diagonalisierenden Hamilton-Matrizen voraus. Dazu werden die Basisvektoren nach Symmetrien oder anderen Erhaltungsgrößen klassifiziert bzw. wird eine neue geeignete Basis gesucht.

Grundprinzip: Falls der Operator \hat{P} mit dem Hamiltonoperator $\hat{\mathcal{H}}$ vertauscht, $[\hat{\mathcal{H}}, \hat{P}] = 0$, so gibt es eine gemeinsame Eigenbasis. Konsequenz: $\hat{\mathcal{H}}$ koppelt keine \hat{P} -Eigenzustände zu verschiedenen \hat{P} -Eigenwerten; folglich kann $\hat{\mathcal{H}}$ in den einzelnen \hat{P} -Eigenräumen separat diagonalisiert werden. Bei mehreren Symmetrien $\hat{P}, \hat{Q}, \hat{R}, \dots$, kann die Dimension für die ED-Schritte weiter reduziert werden.

Außer der Reduktion des numerischen Aufwands hilft die Symmetrieklassifikation auch bei der physikalischen Interpretation und bei der Grenzwertbildung $N \rightarrow \infty$. Symmetrien des Heisenberg-Modells (in Ising-Basis):

- z -Komponente des Gesamtspins $m_z = \sum_{i=1}^N \sigma_i^z$ EW: $\{-N, -N+2, \dots, N\}$
- Spin-Umkehr $\sigma_i^z \rightarrow -\sigma_i^z \quad \forall 1 \leq i \leq N$ EW: ± 1
- Spiegelung $\sigma_i^z \leftrightarrow \sigma_{N-i}^z \quad \forall 1 \leq i \leq N$ EW: ± 1

- Translation $\sigma_i^z \rightarrow \sigma_{i+1}^z \quad \forall 1 \leq i \leq N$ (nur bei periodischen Randbedingungen)
EW: $e^{\frac{2\pi i k}{N}}; 0 \leq k < N$
- Diagonalspiegelung (nur $d > 1$), z.B. $\sigma_{ij}^z \rightarrow \sigma_{ji}^z$ EW: ± 1

Dabei hilft die Spinumkehr höchstens im $m_z = 0$ -Sektor.

Beispiel Heisenberg-Modell, N=2: Hier Spiegelung $\hat{=}$ Translation, daher egal, ob offene oder periodische Randbedingungen

$m_z = 2$: einziger Repräsentant $|2, 1\rangle = \uparrow\uparrow$
damit automatisch \mathcal{H} -EV $\mathcal{H}|2, 1\rangle = -2J|2, 1\rangle$ auch EV bezüglich Spiegelung (und Translation)

$m_z = 0$: Repräsentanten $\uparrow\downarrow$ und $\downarrow\uparrow$
hier Spin-Umkehr $\hat{=}$ Spiegelung ($\hat{=}$ Translation)
symm. Lösung ($q = 0$): $|0, 1\rangle = \frac{1}{\sqrt{2}}(\uparrow\downarrow + \downarrow\uparrow)$
antisymm. Lösung ($q = \pi$): $|0, 2\rangle = \frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow)$

$$H = J \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix} \begin{matrix} \uparrow\downarrow \\ \downarrow\uparrow \end{matrix} \quad \begin{matrix} \mathcal{H}|0, 1\rangle = 6J|0, 1\rangle \\ \mathcal{H}|0, 2\rangle = -2J|0, 2\rangle \end{matrix} \quad (2.16)$$

$m_z = -2$: $|-2, 1\rangle = \downarrow\downarrow; \mathcal{H}|-2, 1\rangle = -2J|2, 1\rangle$

Damit ergibt sich für die unitäre Transformationsmatrix ($J = 1$):

$$\begin{aligned} U^{-1} H U &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 2 & 4 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & \frac{6}{\sqrt{2}} & -\frac{2}{\sqrt{2}} & 0 \\ 0 & \frac{6}{\sqrt{2}} & \frac{2}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \end{aligned} \quad (2.17)$$

Beispiel Heisenberg-Modell, N=4: $m_z = 0$:

$$\begin{array}{ll} \uparrow\downarrow\uparrow\downarrow & \downarrow\uparrow\downarrow\uparrow \\ 1. \text{ alternierend } p = 0; & \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} |0, 1\rangle \\ & p = \pi & \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix} |0, 2\rangle \end{array}$$

2. phasensepariert

		$\uparrow\uparrow\downarrow\downarrow$	$\downarrow\uparrow\uparrow\downarrow$	$\downarrow\downarrow\uparrow\uparrow$	$\uparrow\downarrow\downarrow\uparrow$	
$p = 0$	$\frac{1}{2}$	1	1	1	1	$ 0, 3\rangle$
$p = \frac{\pi}{2}$	$\frac{1}{\sqrt{2}}$	1	0	-1	0	$ 0, 4\rangle$
		0	1	0	-1	$ 0, 5\rangle$
$p = \pi$	$\frac{1}{2}$	1	-1	1	-1	$ 0, 6\rangle$

Hamilton-Matrix für $m_z = 0$:

$$H \hat{=} -J \begin{pmatrix} \uparrow\downarrow\uparrow\downarrow & \downarrow\uparrow\uparrow\downarrow & \uparrow\uparrow\downarrow\downarrow & \downarrow\uparrow\downarrow\uparrow & \downarrow\downarrow\uparrow\uparrow & \uparrow\downarrow\downarrow\uparrow \\ -4 & 0 & 2 & 2 & 2 & 2 \\ 0 & -4 & 2 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \uparrow\downarrow\uparrow\downarrow \\ \downarrow\uparrow\uparrow\downarrow \\ \uparrow\uparrow\downarrow\downarrow \\ \downarrow\uparrow\downarrow\uparrow \\ \downarrow\downarrow\uparrow\uparrow \\ \uparrow\downarrow\downarrow\uparrow \\ \uparrow\downarrow\downarrow\uparrow \end{matrix} \quad (2.18)$$

Impuls $p = 0$:

$$\begin{aligned} \Rightarrow \mathcal{H}|0, 1\rangle &= -J \frac{1}{\sqrt{2}} \begin{pmatrix} -4 \\ -4 \\ 4 \\ 4 \\ 4 \\ 4 \end{pmatrix} = 4J|0, 1\rangle - 4\sqrt{2}J|0, 3\rangle \\ H|0, 3\rangle &= -\frac{J}{2} \begin{pmatrix} 8 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = -4\sqrt{2}J|0, 1\rangle \\ \Rightarrow \mathcal{H} &\hat{=} J \begin{pmatrix} 4 & -4\sqrt{2} \\ -4\sqrt{2} & 0 \end{pmatrix} \end{aligned} \quad (2.19)$$

$$\begin{aligned} p(\lambda) &= (\lambda - 4)\lambda - 32 = \lambda^2 - 4\lambda - 32 \\ \lambda_{\pm} &= 2 \pm \sqrt{36} = 2 \pm 6 = \begin{cases} 8 \\ -4 \end{cases} \end{aligned} \quad (2.20)$$

EW8:

$$\begin{aligned} \begin{pmatrix} 4 & 4\sqrt{2} \\ 4\sqrt{2} & 8 \end{pmatrix} \vec{x} = 0 &\Rightarrow \vec{x} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix} \\ &\Rightarrow \sqrt{\frac{2}{3}}|0, 1\rangle - \frac{1}{\sqrt{3}}|0, 3\rangle \text{ ist EV zum EW } 8J. \end{aligned} \quad (2.21)$$

EW-4:

$$\begin{aligned} \begin{pmatrix} -8 & 4\sqrt{2} \\ 4\sqrt{2} & -4 \end{pmatrix} \vec{x} = 0 &\Rightarrow \vec{x} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix} \\ &\Rightarrow \frac{1}{\sqrt{3}}|0, 1\rangle + \sqrt{\frac{2}{3}}|0, 3\rangle \text{ ist EV zum EW } -4J \end{aligned} \quad (2.22)$$

$p = \pi$:

$$\begin{aligned}\mathcal{H}|0, 2\rangle &= -\frac{J}{\sqrt{2}} \begin{pmatrix} -4 \\ 4 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 4J|0, 2\rangle \quad \text{EV} \\ \mathcal{H}|0, 6\rangle &= -\frac{J}{2} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0J|0, 6\rangle \quad \text{EV}\end{aligned}\tag{2.23}$$

$p = \frac{\pi}{2}, \frac{3\pi}{2}$:

$$\begin{aligned}\mathcal{H}|0, 4\rangle &= -\frac{J}{2} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0J|0, 4\rangle \quad \text{EV} \\ \mathcal{H}|0, 5\rangle &= -\frac{J}{2} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0J|0, 5\rangle \quad \text{EV}\end{aligned}\tag{2.24}$$

$$\Rightarrow \text{EVs: } \frac{1}{2\sqrt{3}} \begin{pmatrix} \frac{2}{2} \\ -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

$$\begin{array}{cccccc} p=0 & p=0 & p=\pi & p=\pi & p=\frac{\pi}{2}, \frac{3\pi}{2} & p=\frac{\pi}{2}, \frac{3\pi}{2} \\ E=8J & E=-4J & E=4J & E=0J & E=0J & E=0J \end{array}$$

$m_z = 4$:

$$p = 0 \quad \uparrow\uparrow\uparrow\uparrow = |4, 1\rangle; \mathcal{H}|4, 1\rangle = -4|4, 1\rangle\tag{2.25}$$

$m_z = 2$:

		$\downarrow\uparrow\uparrow\uparrow$	$\uparrow\downarrow\uparrow\uparrow$	$\uparrow\uparrow\downarrow\uparrow$	$\uparrow\uparrow\uparrow\downarrow$		
	$p = 0$:	$\frac{1}{2}$	(1	1	1	1)	$ 2, 1\rangle$
cos	$p = \frac{\pi}{2}$:	$\frac{1}{\sqrt{2}}$	(1	0	-1	0)	$ 2, 2\rangle$
sin		$\frac{1}{\sqrt{2}}$	(0	1	0	-1)	$ 2, 3\rangle$
	$p = \pi$:	$\frac{1}{2}$	(1	-1	1	-1)	$ 2, 4\rangle$

$$\begin{aligned}p = 0 \quad \mathcal{H}|2, 1\rangle &= -\frac{J}{2} \begin{pmatrix} 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = -\frac{J}{2} \begin{pmatrix} 4 \\ 4 \\ 4 \\ 4 \end{pmatrix} = -4J|2, 1\rangle \\ p = \frac{\pi}{2}, \frac{3\pi}{2} \quad \mathcal{H}|2, 2\rangle &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0J|2, 2\rangle \quad \mathcal{H}|2, 3\rangle = 0J|2, 3\rangle \\ p = \pi \quad \mathcal{H}|2, 4\rangle &= -\frac{J}{2} \begin{pmatrix} -4 \\ 4 \\ -4 \\ 4 \end{pmatrix} = 4J|2, 4\rangle\end{aligned}\tag{2.26}$$

Damit haben wir das 4Spin-Heisenberg-Modell mit periodischen Randbedingungen vollständig analytisch gelöst - durch Ausnutzen der Symmetrien!

Beispiel Heisenberg-Modell, N=4, offene Randbedingung:

$$N = 4, S_z = 4 : \quad |4, 1\rangle = |\uparrow\uparrow\uparrow\uparrow\rangle; \quad \mathcal{H}|4, 1\rangle = -3J|4, 1\rangle$$

$$S_z = 2 :$$

$$\begin{array}{lcl} \text{symm} & |2, 1\rangle = \frac{1}{\sqrt{2}} & \begin{array}{cccc} \downarrow\uparrow\uparrow\uparrow & \uparrow\downarrow\uparrow\uparrow & \uparrow\uparrow\downarrow\uparrow & \uparrow\uparrow\uparrow\downarrow \\ (1 & 0 & 0 & 1) \end{array} \\ & |2, 2\rangle = \frac{1}{\sqrt{2}} & \begin{array}{cccc} (0 & 1 & 1 & 0) \end{array} \\ \text{asymm} & |2, 3\rangle = \frac{1}{\sqrt{2}} & \begin{array}{cccc} (1 & 0 & 0 & -1) \\ |2, 4\rangle = \frac{1}{\sqrt{2}} & (0 & 1 & -1 & 0) \end{array} \end{array} \quad (2.27)$$

$$\mathcal{H} \hat{=} -J \begin{array}{c} \begin{array}{cccc} \downarrow\uparrow\uparrow\uparrow & \uparrow\downarrow\uparrow\uparrow & \uparrow\uparrow\downarrow\uparrow & \uparrow\uparrow\uparrow\downarrow \\ \left(\begin{array}{cccc} 1 & 2 & 0 & 0 \\ 2 & -1 & 2 & 0 \\ 0 & 2 & -1 & 2 \\ 0 & 0 & 2 & 1 \end{array} \right) & \begin{array}{c} \downarrow\uparrow\uparrow\uparrow \\ \uparrow\downarrow\uparrow\uparrow \\ \uparrow\uparrow\downarrow\uparrow \\ \uparrow\uparrow\uparrow\downarrow \end{array} \end{array} \end{array} \quad (2.28)$$

symm:

$$\mathcal{H}|2, 1\rangle = -\frac{J}{\sqrt{2}} \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix} = -J|2, 1\rangle - 2J|2, 2\rangle \quad (2.29)$$

$$\mathcal{H}|2, 2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 \\ 1 \\ 1 \\ 2 \end{pmatrix} = 2J|2, 1\rangle - J|2, 2\rangle$$

$$\mathcal{H} \hat{=} J \begin{pmatrix} -1 & -2 \\ -2 & -1 \end{pmatrix}; \quad p(\lambda) = (\lambda + 1)(\lambda + 1) - 4$$

$$= \lambda^2 + 2\lambda - 3 \quad (2.30)$$

$$\Rightarrow \lambda_{\pm} = -1 \pm \sqrt{4} = \{1, -3\}$$

asymm:

$$\mathcal{H}|2, 3\rangle = -\frac{J}{\sqrt{2}} \begin{pmatrix} 1 \\ 2 \\ -2 \\ -1 \end{pmatrix} = -J|2, 3\rangle - 2J|2, 4\rangle \quad (2.31)$$

$$\mathcal{H}|2, 4\rangle = -\frac{J}{\sqrt{2}} \begin{pmatrix} 2 \\ -3 \\ 3 \\ -2 \end{pmatrix} = -2J|2, 3\rangle + 3J|2, 4\rangle$$

$$\begin{aligned}
\mathcal{H} &\hat{=} J \begin{pmatrix} -1 & -2 \\ -2 & 3 \end{pmatrix}; \quad p(\lambda) = (\lambda + 1)(\lambda - 3) - 4 \\
&= \lambda^2 - 2\lambda - 7 \\
&\Rightarrow \lambda_{\pm} = 1 \pm 2\sqrt{2}
\end{aligned} \tag{2.32}$$

(Basis abweichend von Vorlesung!)

 $s_z = 0$:

$$\begin{array}{ll}
\text{symm.} & |0, 1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \uparrow\uparrow\downarrow & \downarrow\downarrow\uparrow\uparrow & \uparrow\downarrow\downarrow\uparrow & \downarrow\uparrow\uparrow\downarrow & \uparrow\downarrow\uparrow\downarrow & \downarrow\uparrow\downarrow\uparrow \\ 1 & 1 & & & & \end{pmatrix} \\
& |0, 2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} & & 1 & 1 & & \end{pmatrix} \\
& |0, 3\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} & & & & 1 & 1 \end{pmatrix} \\
\text{asymm.} & |0, 4\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & & & & \end{pmatrix} \\
& |0, 5\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} & & 1 & -1 & & \end{pmatrix} \\
& |0, 6\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} & & & & 1 & -1 \end{pmatrix}
\end{array} \tag{2.33}$$

$$\mathcal{H} \hat{=} -J \begin{pmatrix} \uparrow\uparrow\downarrow & \downarrow\downarrow\uparrow\uparrow & \uparrow\downarrow\downarrow\uparrow & \downarrow\uparrow\uparrow\downarrow & \uparrow\downarrow\uparrow\downarrow & \downarrow\uparrow\downarrow\uparrow \\ 1 & 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & -1 & 0 & 2 & 2 \\ 0 & 0 & 0 & -1 & 2 & 2 \\ 2 & 0 & 2 & 2 & -3 & 0 \\ 0 & 2 & 2 & 2 & 0 & -3 \end{pmatrix} \begin{matrix} \uparrow\uparrow\downarrow \\ \downarrow\downarrow\uparrow\uparrow \\ \uparrow\downarrow\downarrow\uparrow \\ \downarrow\uparrow\uparrow\downarrow \\ \uparrow\downarrow\uparrow\downarrow \\ \downarrow\uparrow\downarrow\uparrow \\ \downarrow\uparrow\downarrow\uparrow \end{matrix} \tag{2.34}$$

symm.:

$$\begin{aligned}
\mathcal{H}|0, 1\rangle &= -\frac{J}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 2 \\ 2 \end{pmatrix} = -J(1|0, 1\rangle + 2|0, 3\rangle) \\
\mathcal{H}|0, 2\rangle &= -\frac{J}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ -1 \\ -1 \\ 4 \\ 4 \end{pmatrix} = -J(-|0, 2\rangle + 4|0, 3\rangle) \\
\mathcal{H}|0, 3\rangle &= -\frac{J}{\sqrt{2}} \begin{pmatrix} 2 \\ 2 \\ 4 \\ 4 \\ -3 \\ -3 \end{pmatrix} = -J(2|0, 1\rangle + 4|0, 2\rangle - 3|0, 3\rangle)
\end{aligned} \tag{2.35}$$

$$\begin{aligned}
\mathcal{H} &\hat{=} J \begin{pmatrix} -1 & 0 & -2 \\ 0 & 1 & -4 \\ -2 & -4 & 3 \end{pmatrix}; \quad p(\lambda) = (\lambda^2 - 1)(\lambda - 3) - 4(\lambda - 1) - 16(\lambda + 1) \\
&= \lambda^3 - 3\lambda^2 - \lambda + 3 - 4\lambda + 4 - 16\lambda - 16 \\
&= \lambda^3 - 3\lambda^2 - 21\lambda - 9
\end{aligned} \tag{2.36}$$

Um das allgemeine Verfahren zur Lösung von Polynomen 3. Grades zu umgehen, raten wir, dass eine schon gefundene Nullstelle wieder auftritt:

$$\begin{aligned}
p(\lambda) &= (\lambda + 3)(\lambda^2 - 6\lambda - 3) \\
&\Rightarrow \lambda_1 = -3 \\
&\lambda_{2,3} = 3 \pm 2\sqrt{3}
\end{aligned} \tag{2.37}$$

assym:

$$\begin{aligned}
 \mathcal{H}|0,5\rangle &= J|0,5\rangle \\
 \mathcal{H}|0,4\rangle &= -\frac{J}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 2 \\ -2 \end{pmatrix} = -J|0,4\rangle - 2J|0,6\rangle \\
 \mathcal{H}|0,6\rangle &= -\frac{J}{\sqrt{2}} \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \\ -3 \\ 3 \end{pmatrix} = -2J|0,4\rangle + 3J|0,6\rangle
 \end{aligned} \tag{2.38}$$

$$\begin{aligned}
 \mathcal{H} \hat{=} J \begin{pmatrix} -1 & -2 \\ -2 & 3 \end{pmatrix}; \quad p(\lambda) &= (\lambda+1)(\lambda-3) - 4 \\
 &= \lambda^2 - 2\lambda - 7 \\
 \lambda_{\pm} &= 1 \pm 2\sqrt{2}
 \end{aligned} \tag{2.39}$$

Damit ist auch das 4-Spin-Modell mit offenen Randbedingungen exakt gelöst.

Beispiel Heisenberg-Modell, $N = 3$ (periodisch):

$m_z = 3$: $|3,1\rangle = \uparrow\uparrow\uparrow$; $\mathcal{H}|3,1\rangle = -3J|3,1\rangle$

$m_z = 1$:

	Zustände	$\overbrace{\downarrow\uparrow\uparrow}^{\text{asymm}}$	$\overbrace{\uparrow\downarrow\uparrow}^{\text{symm}}$	$\overbrace{\uparrow\uparrow\downarrow}^{\text{asymm}}$	Kombination muss EV sein
Muss EV sein	$q = 0 \quad \frac{1}{\sqrt{3}}$	1	1	1	$ 1,1\rangle$
	$q = \frac{2\pi}{3} \quad \sqrt{\frac{2}{3}}$	1	$-\frac{1}{2}$	$-\frac{1}{2}$	$ 1,2\rangle$
	$\frac{1}{\sqrt{2}}$	0	1	-1	$ 1,3\rangle$

(2.40)

$$\mathcal{H} \hat{=} -J \begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix} \tag{2.41}$$

$$\begin{aligned}
 \mathcal{H}|1,1\rangle &= -\frac{J}{\sqrt{3}} \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = -3J|1,1\rangle \\
 \mathcal{H}|1,2\rangle &= -\sqrt{\frac{2}{3}} J \begin{pmatrix} -3 \\ \frac{3}{2} \\ \frac{3}{2} \end{pmatrix} = -3J|1,2\rangle \\
 \mathcal{H}|1,3\rangle &= -\frac{J}{\sqrt{2}} \begin{pmatrix} 0 \\ -3 \\ 3 \end{pmatrix} = 3J|1,3\rangle
 \end{aligned} \tag{2.42}$$

2.3.1 Vergleich mit allgemeinem vollständig verbundenem Modell

$$\begin{aligned}\mathcal{H} &= -J \sum_{i \neq j} \vec{\sigma}_i \cdot \vec{\sigma}_j = -4J \sum_{i \neq j} \vec{S}_i \vec{S}_j = -2J(\vec{S}_{ges}^2 - \sum_i \vec{S}_i^2) \\ &\stackrel{h=1}{=} -2J(\vec{S}_{ges}^2 - \frac{3}{4}N); \quad \vec{S}_{ges}^2 = S_{ges}(S_{ges} + 1);\end{aligned}\tag{2.43}$$

$S_{ges} \in \{0, 1, \dots, \frac{N}{2}\}$ für N gerade
 $S_{ges} \in \{\frac{1}{2}, \frac{3}{2}, \dots, \frac{N}{2}\}$ für N ungerade

N	S_{ges}	\vec{S}_{ges}^2	$\frac{3}{4}N$	E/J	
1	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{3}{4}$	0	nicht-WW.
2	0, 1	0, 2	$\frac{3}{2}$	3, -1	$N = 2$ offene Randb.
3	$\frac{1}{2}, \frac{3}{2}$	$\frac{3}{4}, \frac{15}{4}$	$\frac{9}{4}$	3, -3	$N = 3$ periodische Randb.
4	0, 1, 2	0, 2, 6	3	6, 2, -6	$N = 4$ “ + NNN-WW
5	$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$	$\frac{3}{4}, \frac{15}{4}, \frac{35}{4}$	$\frac{15}{4}$	6, 0, -10	

(2.44)

2.4 Numerische Verfahren zur vollständigen Bestimmung aller Eigenwerte und (optional) Eigenvektoren

2.4.1 Naive Strategie:

1. Bestimme charakteristisches Polynom $p(\lambda)$
2. Finde Nullstellen λ_i
3. Für jedes i : Löse $(A - \lambda_i \mathbf{1})\vec{x}_i = \vec{0}$

Problem:

1. Die direkte Bestimmung von $p(\lambda)$ z.B. über Entwicklung der Determinante nach Spalten ist ineffizient ($N!$ Operationen). Selbst die effizienteste Determinantenbestimmung ($\frac{1}{3}N^3$ Additionen und Multiplikationen bei LU-Zerlegung) wäre zu teuer, da sie für jedes λ wiederholt werden müsste.
2. Die Nullstellenbestimmung $p(\lambda) - \{\lambda_i\}$ ist schlecht konditioniert (dagegen ist die gesamte EW-Bestimmung von hermiteschen Matrizen gut konditioniert), d.h. kleine Rundungsfehler in den Koeffizienten von $p(\lambda)$ können zu großen Fehlern in den EW führen.
3. Nullstellensuche ist (im Fall mehrerer Nullstellen) schwierig, besonders für $(2n)$ -fache Nullstellen.

2.4.2 Von-Mises-Verfahren (wiederholte Multiplikation mit \mathcal{H}), ggf. mit Orthogonalisierung

Inverse Iteration (*Wielandt*). Multiplikation mit $(\mathcal{H} - \tilde{\lambda})^{-1}$ geschätzter Eigenwert $\tilde{\lambda}$: gut zur Verfeinerung von $\tilde{\lambda} \rightarrow \lambda$ und selektiven EV-Bestimmung.

Beide Verfahren sind ungeeignet zur Bestimmung vieler/aller EW/EV von großen Matrizen.

Moderne Verfahren zur Bestimmung von Eigenwerten und -vektoren beruhen auf Ähnlichkeitstransformationen mit unitären Matrizen, mit deren Hilfe die gegebene Matrix (teilweise oder vollständig) auf Diagonalgestalt transformiert wird. Dies kann direkt oder über den Zwischenschritt einer Tridiagonalmatrix erfolgen. Zuerst behandeln wir eine konzeptionell einfache direkte Methode.

2.4.3 Iterative Diagonalisierung mit Jacobi-Rotationen

Grundidee: die Diagonalisierung einer hermiteschen Matrix A soll erreicht werden, indem wiederholt einzelne Nicht-Diagonal-Elemente a_{pq} ($p \neq q$) eliminiert werden. Im Fall reeller Matrizen gelingt dies mit Hilfe der orthogonalen Rotationsmatrix (Jacobi-Matrix):

$$P_{pq} = \left(\begin{array}{c|ccc|c} \mathbf{1} & & & & \mathbf{0} & \mathbf{1} \\ \hline & c & 0 & \cdots & 0 & s \\ & 0 & & & & 0 \\ \mathbf{0} & \vdots & & \mathbf{1} & \vdots & \mathbf{0} \\ & 0 & & & & 0 \\ & -s & 0 & \cdots & 0 & c \\ \hline \mathbf{0} & & & \mathbf{0} & & \mathbf{1} \end{array} \right) \quad (2.45)$$

mit $c^2 + s^2 = 1 \Rightarrow c = \cos(\phi); s = \sin(\phi)$

Damit entspricht die Ähnlichkeitstransformation $A \rightarrow P_{pq}^T A P_{pq}$ einer Rotation in der pq -Ebene um den Winkel ϕ . Konkret ändert dabei die Linksmultiplikation einer Matrix M mit P_{pq}^T , $M \rightarrow M' = P_{pq}^T M$ nur die Zeilen p und q von M , die Rechtsmultiplikation $M \rightarrow M' = M P_{pq}$ entsprechend nur die Spalten p und q . Folglich unterscheidet sich $A' = P_{pq}^T A P_{pq}$ von A höchstens in den Elementen:

$$A' = \left(\begin{array}{cccccc} & a'_{1p} & & a'_{1q} & & \\ & \vdots & & \vdots & & \\ a'_{p1} & \cdots & a'_{pp} & \cdots & a'_{pq} & \cdots & a'_{pN} \\ & \vdots & & \vdots & & & \\ a'_{q1} & \cdots & a'_{qp} & \cdots & a'_{qq} & \cdots & a'_{qN} \\ & \vdots & & \vdots & & & \\ & a'_{Np} & & a'_{Nq} & & & \end{array} \right) \quad (2.46)$$

Mit der Symmetrie von A erhalten wir:

$$\begin{aligned}
 a'_{pp} &= c^2 a_{pp} + s^2 a_{qq} - 2sca_{pq} \\
 a'_{qq} &= s^2 a_{pp} + c^2 a_{qq} + 2sca_{pq} \\
 a'_{pq} &= (c^2 - s^2)a_{pp} + sc(a_{pp} - a_{qq}) \\
 a'_{rp} &= ca_{rp} - sa_{rq} \quad \text{für } r \neq p, r \neq q \\
 a'_{rq} &= ca_{rq} + sa_{rp} \quad \text{für } r \neq p, r \neq q
 \end{aligned} \tag{2.47}$$

Verifikation: Berechne zunächst $\tilde{A} = AP_{pq}$

$$\begin{aligned}
 \tilde{a}_{ir} &= a_{ir} \quad \text{für } r \neq p, r \neq q \\
 \tilde{a}_{ip} &= ca_{ip} - sa_{iq} \\
 \tilde{a}_{iq} &= sa_{ip} + ca_{iq}
 \end{aligned} \tag{2.48}$$

führe jetzt die zweite Multiplikation durch:

$$\begin{aligned}
 a'_{ri} &= \tilde{a}_{ri} \quad \text{für } r \neq p, r \neq q \\
 a'_{pi} &= c\tilde{a}_{pi} - s\tilde{a}_{qi} \\
 a'_{qi} &= s\tilde{a}_{pi} + c\tilde{a}_{qi}
 \end{aligned} \tag{2.49}$$

$$\begin{aligned}
 \Rightarrow a'_{pp} &= c\tilde{a}_{pp} - s\tilde{a}_{qp} = c(ca_{pp} - sa_{pq}) - s(ca_{qp} - sa_{qq}) \\
 &\stackrel{A=A^T}{=} c^2 a_{pp} + s^2 a_{qq} - 2sca_{pq}
 \end{aligned} \tag{2.50}$$

analog:

$$\begin{aligned}
 a'_{qq} &= s^2 a_{pp} + c^2 a_{qq} + 2sca_{pq} \\
 a'_{pq} &= c\tilde{a}_{pq} - s\tilde{a}_{qq} = c(sa_{pp} + ca_{pq}) - s(sa_{qp} + ca_{qq}) \\
 &\stackrel{A=A^T}{=} (c^2 - s^2)a_{pq} + sc(a_{pp} - a_{qq}) \\
 a'_{rp} &= \tilde{a}_{rp} = ca_{rp} - sa_{rq} \quad \text{für } r \neq p, r \neq q \\
 a'_{rq} &= \tilde{a}_{rq} = sa_{rp} + ca_{rq} \quad \text{für } r \neq p, r \neq q
 \end{aligned} \tag{2.51}$$

Um den Eintrag mit Koordinaten p, q zu eliminieren, fordern wir:

$$0 \stackrel{!}{=} a'_{pq} \Leftrightarrow \frac{c^2 - s^2}{2sc} = \frac{a_{qq} - a_{pp}}{2a_{pq}} \equiv \Theta \tag{2.52}$$

$$\text{Es gilt: } \Theta = \frac{c^2 - s^2}{2sc} = \frac{\cos^2(\phi) - \sin^2(\phi)}{2 \sin(\phi) \cos(\phi)} = \frac{\cos(2\phi)}{\sin(2\phi)} = \cot(2\phi)$$

Vor der Rotation ist Θ bekannt. Wir könnten also theoretisch setzen: $\phi = \frac{1}{2 \arctan(\Theta)}$; $c = \cos(\phi)$, $s = \sin(\phi)$ und dann mit (2.47) A' ausrechnen.

Stattdessen formt man in der Praxis um:

$$t \equiv \frac{s}{c} (= \tan(\phi)) \Rightarrow \Theta = \frac{1 - t^2}{2t} \Leftrightarrow t^2 + 2t\Theta - 1 = 0 \tag{2.53}$$

und löst die quadratische Gleichung numerisch stabil nach der betragskleineren Lösung auf:

$$t = \begin{cases} \frac{\text{sign}(\Theta)}{|\Theta| + \sqrt{\Theta^2 + 1}} & \text{falls } \Theta^2 < \text{MAX} \\ \frac{1}{2\Theta} & \text{sonst} \end{cases} \quad (2.54)$$

$$c = \frac{1}{\sqrt{t^2 + 1}}; \quad s = tc$$

Um Rundungsfehler zu minimieren, vermeidet man es, verschiedene Diagonalelemente (die i.A. betragsgroß sind) zu addieren. Daher benutzt man statt (2.47) z.B.

$$0 = a'_{pq} = (c^2 - s^2)a_{pq} + sca_{pp} - sca_{qq} \quad (2.55)$$

$$\Leftrightarrow a_{qq} = a_{pp} + \frac{c^2 - s^2}{sc} a_{pq}$$

$$\begin{aligned} \Rightarrow a'_{pp} &= c^2 a_{pp} + s^2 a_{qq} - 2sca_{pq} \\ &= \underbrace{(c^2 + s^2)}_{=1} a_{pp} + \underbrace{\frac{s}{c}}_{=t} \underbrace{(c^2 - s^2 - 2c^2)}_{=-1} a_{pq} \\ &= a_{pp} - ta_{pq} \end{aligned} \quad (2.56)$$

Insgesamt erhält man mit $\tau = \frac{s}{1+c}$ die Aktualisierungsgleichungen:

$$a'_{pp} = a_{pp} - ta_{pq} \quad (2.57)$$

$$a'_{qq} = a_{qq} + ta_{pq} \quad (2.58)$$

$$a'_{pq} = 0 \quad (2.59)$$

$$a'_{rp} = a_{rp} - s(a_{rq} + \tau a_{rp}) \quad \text{für } r \neq p, r \neq q \quad (2.60)$$

$$a'_{rq} = a_{rq} + s(a_{rp} - \tau a_{rq}) \quad \text{für } r \neq p, r \neq q \quad (2.61)$$

Damit ist die Jacobi-Rotation vollständig definiert, die ein Nebendiagonalelement „vernichtet“: $a_{pq} \rightarrow 0$. Allerdings können gleichzeitig neue nichtverschwindende Einträge a'_{rp}, a'_{rq} erzeugt werden.

Frage: wurde überhaupt ein Fortschritt erzielt?

Antwort: *ja!* Betrachte als Maß für die „Nichtdiagonalität“ die Summe der Quadrate der Nichtdiagonalelemente: $S = \sum_{r \neq s} |a_{rs}|^2$. Aus (2.47) folgt für $r \neq p, r \neq q$

$$\begin{aligned} |a'_{rp}|^2 + |a'_{rq}|^2 &= c^2 a_{rp}^2 + s^2 a_{rq}^2 - 2sca_{rp}a_{rq} \\ &\quad + c^2 a_{rq}^2 + s^2 a_{rp}^2 + 2sca_{rp}a_{rq} = a_{rp}^2 + a_{rq}^2 \end{aligned} \quad (2.62)$$

Mit $|a_{pq}|^2 = 0 = a_{pq}^2 - a_{pq}^2$ und $A^T = A, A'^T = A'$ gilt also: $S' = \sum_{r \neq s} |a'_{rs}|^2 = S - 2a_{pq}^2$. Bei iterativer Anwendung ist also die Folge S_n streng monoton fallend (bei jeweiliger Wahl von p, q mit $|a_{pq}| > 0$), d.h. das Jacobi-Verfahren führt asymptotisch auf eine Diagonalmatrix, die alle Eigenwerte von A enthält.

Da die Multiplikation mit einer orthogonalen Matrix die Länge jedes Zeilen- bzw. Spaltenvektors invariant lässt (für $M \rightarrow MP$ bzw. $M \rightarrow P^T M$) bleibt die Summe der Quadrate aller Elemente konstant. Es folgt: $|a'_{pp}|^2 + |a'_{qq}|^2 = a_{pp}^2 + a_{qq}^2 + 2a_{pq}$

Die Eigenvektoren erhält man als Spalten der Gesamttransformationsmatrix: $V = \mathbf{1}P_1P_2P_3\dots$ mit den Updates

$$V' = VP_i \Leftrightarrow \begin{cases} v'_{rs} = v_{rs} & \text{für } s \neq p, s \neq q \\ v'_{rp} = cv_{rp} - sv_{rq} \\ v'_{rq} = sv_{rp} + cv_{rq} \end{cases} \quad (2.63)$$

Anwendungsstrategie: offensichtlich ist es wünschenswert, jeweils große Elemente a_{pq} zu eliminieren. Eine vollständige Suche ist jedoch viel zu teuer.

Lösung in Numerical Recipes: normalerweise werden in einem *sweep* nacheinander Rotationen P_{pq} für alle Paare p, q mit $p < q$ versucht, jedoch in den ersten 3 sweeps Rotationen übersprungen, falls

$$|a_{pq}| > \frac{1}{5} \frac{S_0}{N^2}; \quad S_0 = \sum_{r < s} |a_{rs}| \quad (2.64)$$

also für $|a_{pq}| \gtrsim \frac{1}{10} |a_{rs}|$.

Später werden nur bei extrem kleinen $|a_{pq}|$ vereinfachte Updates gemacht ($a'_{pq} = 0$, andere Elemente unverändert), wenn sich dadurch die Eigenwerte nicht ändern.

Typisch sind 6-10 sweeps für Konvergenz auf Maschinengenauigkeit erforderlich, d.h. $3N^2 - 5N^2$ Jacobi-Rotationen. Gesamtaufwand: ca. $20N^3$ Operationen (Multiplikationen bzw. Additionen, die Quadratwurzeln beeinflussen nur den N^2 -Term).

Der beschriebene Jacobi-Algorithmus ist einfach und absolut stabil. Trotz Effizienz-nachteilen gegenüber der Kombination Householder- und QR-Verfahren wird er daher für die vollständige Diagonalisierung mäßig großer Matrizen empfohlen.

2.4.4 Vollständige Tridiagonalisierung symmetrischer Matrizen nach Givens

Das Givens-Verfahren zur Tridiagonalisierung verwendet Jacobi-Rotationen um symmetrische (oder hermitesche) Matrizen Spalte für Spalte zu tridiagonalisieren. Konkret eliminiert man mit P_{pq} ($p < q$) das Element $a_{p-1,q}$ (und aus Symmetriegründen $a_{q,p-1}$):

$$\begin{aligned} 0 &\stackrel{!}{=} a'_{p-1,q} = ca_{p-1,q} + s \overset{\text{Nebendiagonalelement}}{\downarrow} a_{p-1,p} \\ \text{falls } a_{p-1,p} &= 0 \Rightarrow c = 0, s = 1 \\ \text{sonst } \frac{s}{c} &= \tan(\phi) - \frac{a_{p-1,q}}{a_{p-1,p}} \end{aligned} \quad (2.65)$$

Da für die Bearbeitung der Spalte r nur Jacobi-Matrizen P_{pq} mit $r < p < q$ eingesetzt werden, bleibt diese von der Spaltenkopplung unbeeinflusst. Die Zeilenkopplung wiederum erfasst nur das Nebendiagonalelement sowie das ausgesuchte Element.

Somit bleiben alle einmal erzeugten Nullelemente außerhalb der tridiagonalen erhalten; das Verfahren benötigt höchstens $\frac{(N-1)(N-2)}{2}$ Schritte.

Gesamtordnung (ohne Transformationsmatrix): $\frac{4}{3}N^3$

Beispiel 5×5 -Matrix:

$$\begin{array}{ccccc}
* & * & * & * & * \\
\boxed{*} & * & * & * & * \\
* & * & * & * & * \\
* & * & * & * & * \\
* & * & * & * & *
\end{array}
\begin{array}{c}
\begin{array}{ccccc}
* & * & 0 & * & * \\
\boxed{*} & * & * & * & * \\
0 & * & * & * & * \\
\boxed{*} & * & * & * & * \\
* & * & * & * & *
\end{array}
\begin{array}{c}
\begin{array}{ccccc}
* & * & 0 & 0 & * \\
\boxed{*} & * & * & * & * \\
0 & * & * & * & * \\
0 & * & * & * & * \\
0 & * & * & * & *
\end{array}
\begin{array}{c}
\begin{array}{ccccc}
* & * & 0 & 0 & 0 \\
* & * & * & 0 & * \\
0 & \boxed{*} & * & * & * \\
0 & * & * & * & * \\
0 & \boxed{*} & * & * & *
\end{array}
\begin{array}{c}
\begin{array}{ccccc}
* & * & 0 & 0 & 0 \\
* & * & * & 0 & 0 \\
0 & * & * & * & * \\
0 & 0 & \boxed{*} & * & * \\
0 & \boxed{*} & * & * & *
\end{array}
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{ccccc}
* & * & 0 & 0 & 0 \\
* & * & * & * & * \\
0 & \boxed{*} & * & * & * \\
0 & * & * & * & * \\
0 & * & * & * & *
\end{array}
\begin{array}{c}
\begin{array}{ccccc}
* & * & 0 & 0 & 0 \\
* & * & * & 0 & 0 \\
0 & * & * & * & * \\
0 & 0 & \boxed{*} & * & * \\
0 & \boxed{*} & * & * & *
\end{array}
\begin{array}{c}
\begin{array}{ccccc}
* & * & 0 & 0 & 0 \\
* & * & * & 0 & 0 \\
0 & * & * & * & * \\
0 & 0 & \boxed{*} & * & * \\
0 & \boxed{*} & * & * & *
\end{array}
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{ccccc}
* & * & 0 & 0 & 0 \\
* & * & * & 0 & 0 \\
0 & * & * & * & 0 \\
0 & 0 & * & * & * \\
0 & 0 & 0 & * & *
\end{array}
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{ccccc}
* & * & 0 & 0 & 0 \\
* & * & * & 0 & 0 \\
0 & * & * & * & 0 \\
0 & 0 & * & * & * \\
0 & 0 & 0 & * & *
\end{array}
\end{array}
\end{array}
\quad (2.66)$$

2.4.5 Vollständige Tridiagonalisierung symmetrischer Matrizen mit Householder-Transformationen

In der Praxis verwendet man zur vollständigen Tridiagonalisierung statt der Givens-Methode die um den Faktor 2 effizienteren Transformationen mit Householder-Matrizen der Form

$$P_{\vec{w}} = \mathbf{1} - 2\vec{w}\vec{w}^T; \quad |\vec{w}|^2 = \vec{w}^T\vec{w} = 1 \quad (2.67)$$

Die Matrizen $P_{\vec{w}}$ sind offensichtlich für jeden (normierten) Vektor \vec{w} symmetrisch, außerdem auch orthogonal:

$$\begin{aligned}
P_{\vec{w}}^T P_{\vec{w}} &= P_{\vec{w}}^2 = (\mathbf{1} - 2\vec{w}\vec{w}^T)(\mathbf{1} - 2\vec{w}\vec{w}^T) \\
&= \mathbf{1} - 4\vec{w}\vec{w}^T + 4\vec{w}\underbrace{\vec{w}\vec{w}^T}_{=1}\vec{w}^T = \mathbf{1}
\end{aligned} \quad (2.68)$$

Um die Normierung vorläufig offenlassen zu können, schreiben wir

$$P_{\vec{u}} = \mathbf{1} - \frac{\vec{u}\vec{u}^T}{H}; \quad H = \frac{1}{2}|\vec{u}|^2 \quad (2.69)$$

Jetzt partitionieren wir die zu tridiagonalisierende Matrix A und die 1. Transformationsmatrix P_1 als

$$A = \left(\begin{array}{c|c} a_{11} & \vec{x}^T \\ \hline \vec{x} & * \end{array} \right); \quad P_1 = \left(\begin{array}{c|c} 1 & \vec{0} \\ \hline \vec{0} & \vec{P}_1 \end{array} \right) \quad (2.70)$$

Der Vektor \vec{x} besteht also aus den Nicht-diagonalelementen der ersten Spalte von A . Im ersten Schritt der Tridiagonalisierung soll dieser auf den ersten Einheitsvektor (im Raum \mathbb{R}^{N-1}) abgebildet werden.

Dazu wählen wir: $\vec{u} = \vec{x} \mp |x|\vec{e}_1$

$$\left(\Rightarrow H = \frac{1}{2}|\vec{u}|^2 = \frac{1}{2}[|\vec{x}|^2 + |\vec{x}| \mp 2|x|x_1] = |x|^2 \mp |x|x_1 \right) \quad (2.71)$$

$$\begin{aligned} \Rightarrow \tilde{P}_1 \vec{x} &= \vec{x} - \frac{1}{H} \vec{u} (\vec{u}^T \cdot \vec{x}) \\ &= \vec{x} - \frac{1}{H} [\vec{x}^T \vec{x} \mp |x| \overbrace{\vec{e}_1^T \vec{x}}^{x_1}] = \vec{x} - \vec{u} = \pm |x| \vec{e}_1 \end{aligned} \quad (2.72)$$

Damit hat $\vec{x}' = \tilde{P}_1 \vec{x}$ tatsächlich die gewünschte Form; beachte, dass $|\vec{x}'| = |\vec{x}|$, die Norm ist also erhalten. Damit hat $P_1 A$ die Gestalt

$$P_1 A = \left(\begin{array}{c|c} a_{11} & \vec{x}^T \\ \hline \mp |x| & \\ 0 & \\ \vdots & \\ 0 & \end{array} \right) \quad (2.73)$$

Wegen der Blockgestalt von P bleibt die erste Spalte auch bei Rechtsmultiplikation mit P erhalten:

$$A^{(2)} = P_1 A P_1 = \left(\begin{array}{c|cccc} a_{11} & \mp |x| & 0 & \cdots & 0 \\ \hline \mp |x| & & & & \\ 0 & & & & \\ \vdots & & * & & \\ \vdots & & & & \\ 0 & & & & \end{array} \right) \quad (2.74)$$

Weitere Schritte mit $P_2 = \left(\begin{array}{c|c} 1 & \\ \hline & \tilde{P}_2 \end{array} \right)$ usw. lassen alle vorher tridiagonalisier-

ten Spalten (und Zeilen) invariant (d.h. P_q lässt die r -ten Spalten für $r < q$ invariant), so dass $Q = P_1 P_2 \dots P_{N-2} A$ tridiagonalisiert.

In der Praxis erfolgen die Aktualisierungen $A^{(n)} \rightarrow A^{(n+1)}$ im n -ten Schritt wie folgt (hier $A \equiv A^{(n)}$):

$$1. \sigma = \text{sign}(a_{n+1,n}) \sum_{i=n+1}^N |a_{in}|^2$$

$$2. \vec{u} = (\overbrace{0, \dots, 0}^n, a_{n+1,n} + \sigma, a_{n+2,1}, \dots, a_{N,n})^T$$

3. $H = \frac{1}{2}|\vec{u}|^2$
4. $\vec{p} = \frac{1}{H}A\vec{u}$
5. $K = \frac{1}{2H}\vec{u}^T\vec{p}$
6. $\vec{q} = \vec{p} - K\vec{u}$
7. $\vec{A}' = \vec{A} - \vec{q}\vec{u}^T - \vec{u}\vec{q}^T$

Allerdings vertauschen die Numerical Recipes und z.B. EISPACK die Reihenfolge, statt von links oben wird dort die Matrix von rechts unten aus tridiagonalisiert.

Endergebnis einer Tridiagonalisierung nach Givens, Householder oder Lanczos (siehe 2.6) sind

1. der Diagonalenvektor $\vec{d} = (a_{11}^{(N-1)}, a_{22}^{(N-1)}, \dots, a_{NN}^{(N-1)})$
2. der Nebendiagonalvektor, z.B. $\vec{e} = (0, a_{12}^{(N-1)}, a_{23}^{(N-1)}, \dots, a_{N-1,N}^{(N-1)})$
3. ggf. die Transformationsmatrix Q

Für den nächsten Schritt, die iterative Diagonalisierung, benötigen wir ein Verfahren, das die Tridiagonalgestalt erhält. Das oben besprochene Jacobi-Verfahren kommt daher nicht infrage.

2.5 Bestimmung aller Eigenwerte und (optional) Eigenvektoren von symmetrischen Tridiagonalmatrizen

Die hier besprochenen Verfahren zur iterativen Diagonalisierung ergänzen die in 2.4 und 2.6 behandelten Tridiagonalisierungsverfahren.

2.5.1 Bestimmung der EVs aus charakteristischem Polynom

Das charakteristische Polynom einer hermiteschen Tridiagonalmatrix A lässt sich leicht rekursiv auswerten ($p(\lambda) \equiv p_N(\lambda)$):

$$\begin{aligned} p_1(\lambda) &= a_{11} - \lambda \\ p_n(\lambda) &= (a_{nn} - \lambda)p_{n-1}(\lambda) - |a_{n,n-1}|^2 p_{n-2}(\lambda) \quad (n > 1) \end{aligned} \quad (2.75)$$

Dabei wurde die Konvention benutzt: $p_0(\lambda) = 1$

Diese lassen sich sogar differenzieren:

$$\begin{aligned} p'_1(\lambda) &= \frac{d}{d\lambda} p_1(\lambda) = -1 \\ p'_n(\lambda) &= -p_{n-1}(\lambda) + (a_{nn} - \lambda)p'_{n-1}(\lambda) - |a_{n,n-1}|^2 p'_{n-2}(\lambda) \end{aligned} \quad (2.76)$$

was den Einsatz des Newton-Verfahrens zur Nullstellensuche ermöglicht, z.B. in den Gerschgorin-Grenzen

$$\lambda_n \geq \min_i \{a_{ii} - \sum_{j \neq i} |a_{ij}|\}; \quad \lambda_n \leq \max_i \{a_{ii} + \sum_{j \neq i} |a_{ij}|\} \quad (2.77)$$

Allerdings ist die EW-Bestimmung via $p(\lambda)$ (wie in 2.4 besprochen) nicht stabil.

2.5.2 Iterative Diagonalisierung einer Tridiagonalmatrix mittels QR- bzw. QL-Zerlegung

Jede reelle Matrix A kann in der Form

$$A = QR; \quad Q^T Q = \mathbf{1}; \quad R = \begin{pmatrix} * & & * \\ & * & \\ 0 & & * \end{pmatrix} \quad (2.78)$$

also in das Produkt einer orthogonalen Matrix und einer oberen Dreiecksmatrix zerlegt werden (alternativ ist auch $A = QL$ mit unterer Dreiecksmatrix L möglich).

Dann ist $A' = RQ = Q^T A Q$ unitär ähnlich zu A . Man kann zeigen, dass die Folge $A^{(n)}$ mit sequentiellen QR -Updates gegen eine Diagonalmatrix konvergiert. Dabei gilt für $r \neq s$ und $\lambda_r < \lambda_s$:

$$a_{rs}^{(n)} \sim \left(\frac{\lambda_r}{\lambda_s} \right)^n \quad (2.79)$$

Im Allgemeinen (volle symmetrische Matrix A) würde jede Iteration $\mathcal{O}(n^3)$ Operationen benötigen, was prohibitiv teuer ist.

In der Anwendung auf Tridiagonalmatrizen lässt sich Q aus Jacobi-Rotationen aufbauen:

$$\begin{aligned} A = \begin{pmatrix} * & * & & \\ * & * & * & \\ & * & \ddots & \ddots \\ & & \ddots & \ddots & * \\ & & & \ddots & \ddots & * \end{pmatrix} &\xrightarrow{P_{12}^T} \begin{pmatrix} * & * & * & & \\ 0 & * & * & & \\ & * & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots & * \\ & & & \ddots & \ddots & \ddots & * \end{pmatrix} &\xrightarrow{P_{23}^T} \begin{pmatrix} * & * & * & & \\ 0 & * & * & * & \\ & 0 & * & * & \\ & & * & \ddots & \ddots & * \\ & & & \ddots & \ddots & \ddots & * \end{pmatrix} \\ &\longrightarrow \dots \xrightarrow{P_{N-1,N}^T} \begin{pmatrix} * & * & * & & \\ 0 & * & * & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & * \\ & & \ddots & \ddots & \ddots & \ddots & * \\ & & & 0 & * & * \\ & & & & 0 & * \end{pmatrix} = R; \quad P_{n,n+1} = \begin{pmatrix} \mathbf{1}_{n-1} & & \\ & c & s \\ & -s & c \\ & & & \mathbf{1}_{N-n-1} \end{pmatrix} \end{aligned} \quad (2.80)$$

$$Q^T = P_{N-1,N}^T \dots P_{1,2}^T \quad (2.81)$$

Dabei sind die Bedingungen an die Koeffizienten c, s wieder anders als bei der Jacobi-Diagonalisierung oder der Givens-Tridiagonalisierung, z.B. für P_{12} :

$$\begin{aligned} a'_{12} &= ca_{12} - sa_{22} \stackrel{!}{=} 0 \\ \text{falls } a_{22} = 0 : c &= 0, s = 1, \text{ sonst: } \frac{s}{c} = \tan(\phi) = \frac{a_{12}}{a_{22}} \end{aligned} \quad (2.82)$$

Auf den ersten Blick sieht es so aus, als würde der QR-Schritt die Tridiagonalität zerstören. Multiplikation mit Q führt allerdings auf:

$$\begin{aligned} R \xrightarrow{\cdot P_{12}} & \begin{pmatrix} * & * & * & & \\ * & * & * & \ddots & \\ & 0 & \ddots & \ddots & \ddots \\ & & \ddots & \ddots & * \\ & & & \ddots & * \\ & & & & 0 & * \end{pmatrix} \xrightarrow{\cdot P_{23}} & \begin{pmatrix} * & * & * & & \\ * & * & * & \ddots & \\ & * & * & \ddots & \ddots \\ & & 0 & \ddots & \ddots & * \\ & & & \ddots & \ddots & * \\ & & & & 0 & * \end{pmatrix} \\ & \longrightarrow \dots \xrightarrow{\cdot P_{N-1,N}} & \begin{pmatrix} * & * & * & & \\ * & * & * & \ddots & \\ & \ddots & \ddots & \ddots & \ddots \\ & & \ddots & \ddots & * & * \\ & & & \ddots & * & * \\ & & & & * & * \end{pmatrix} \end{aligned} \quad (2.83)$$

(Beachte, dass die Rechtsmultiplikation mit $P_{n,n+1}$ nur die Spalten n und $n+1$ „durchmischt“ und alle anderen invariant lässt.)

Damit ist $A' = Q^T A Q$ obere Hessebergmatrix.

Andererseits ist $(A')^T = Q^T A^T Q = Q^T A Q = A'$ symmetrisch, also wieder tridiagonal!

Aspekte der praktischen Anwendung:

1. Ändere Reihenfolge:

$A' = P_{N-1,N}^T \dots (P_{23}^T (P_{12}^T A P_{12}) P_{23}) \dots P_{N-1,N}$ Für Konjunktionen sind kompakte Updates (nach $*$) möglich.

2. Aber: jeweils Speicherung eines Außertridiagonalelements nötig.

3. Parameter t bzw. ϕ für $P_{n-1,n}$ hängt von Form der Matrix nach „halbem Update“ (nur Multiplikation mit $P_{n-1,n-1}^T$, nicht volle Konjugation) ab!
→ ziehen von Wurzel nötig

2.6 Partielle Tridiagonalisierung mit dem Lanczos-Algorithmus

Bisher: vollständige (Tri-)Diagonalisierung großer Matrizen A (M -dimensionale Basis)

Problem:

Speicherplatzbedarf ($2M^2$)

Rechenzeit ($\propto M^3$)

Alternativansatz: extrahiere nur „relevanteste Teile“ der Matrix, behandle diese weiter

(analog zu Symmetrien \rightarrow Blockdiagonalform bzw. Eigenräumen der Lösung)

Wie? - Finde „relevante“ Vektoren V_i , berechne

$$(\tilde{A})_{ij} = \langle V_i | A | V_j \rangle \longrightarrow \dim \tilde{A} = \dim \{V_j\} \stackrel{=V}{\ll} M \quad (2.84)$$

Falls (i) $AV_i \in \mathcal{V} \forall i$ und (ii) $\{V_i\}$ linear unabhängig, dann sind alle EW von \tilde{A} auch EW von A ; die den EV von \tilde{A} entsprechenden Linearkombinationen der V_i sind EV von A .

Krylov-Raum: $K(A, V) = \{V, AV, A^2V, \dots\}$

Jede Folge von Krylov-Vektoren zu einer Matrix A mit EV $\{\lambda_i\}_{i=1}^p$ wird nach spätestens $p + 1$ Schritten linear abhängig, da

$$(A - \lambda_1)(A - \lambda_2) \dots (A - \lambda_p) = 0 \quad (2.85)$$

(bekannter: $p(A) = 0$ für char. Polynom $p(\lambda)$ von A) **Beweis:** es existiert unitäre Matrix U mit

$$U^T A U = \text{diag}(\lambda_1, \dots, \lambda_M) \equiv D \quad (2.86)$$

$$\Rightarrow A^n V = U^T D^n U V \quad (2.87)$$

Korrolar: für zwei orthogonale EV V_1, V_2 zum selben EW λ von A gilt:

$$\begin{aligned} V_1 &\notin K(A, V_2) \\ V_2 &\notin K(A, V_1) \end{aligned} \quad (2.88)$$

Ein Krylov-Raum enthält also zu jedem EW von A höchstens einen EV.

2.6.1 Lanczos-Tridiagonalisierung

1. Wähle Startvektor \vec{v}_i (zufällig oder z.B. orthogonal zu bestimmten Vektoren).

2. Konstruiere Folge im Krylov-Raum $K(A, \vec{v}_1)$:

$$\begin{aligned} \vec{v}_2 &= C_2(A\vec{v}_1 - \tilde{A}_{11}\vec{v}_1) \\ \vec{v}_{n+1} &= C_{n+1}(A\vec{v}_n - \tilde{A}_{nn}\vec{v}_n - \tilde{A}_{n,n-1}\vec{v}_{n-1}) \quad \text{für } n > 1 \end{aligned} \quad (2.89)$$

Falls $\tilde{A}_{ij} \neq 0 \Rightarrow \vec{v}_j$ hat Überlapp mit $A\vec{v}_i$

aber: $A\vec{v}_i \in \langle \vec{v}_1, \vec{v}_2 \rangle$
 folglich: falls $\vec{v}_n \perp \vec{v}_1, \vec{v}_n \perp \vec{v}_2$ für $n > 2$
 dann ist $\tilde{A}_{1n} = 0$ “ “

Tatsächlich sind die Lanczos-Vektoren orthogonal:

$$\begin{aligned}
 \vec{v}_1 \cdot \vec{v}_2 &= C_2(\underbrace{\vec{v}_1 A \vec{v}_1}_{=\tilde{A}_{11}} - \underbrace{\tilde{A}_{11}}_{=1} \underbrace{\vec{v}_1 \vec{v}_1}_{=1}) = 0 \\
 \vec{v}_n \cdot \vec{v}_{n+1} &= C_{n+1}(\vec{v}_n A \vec{v}_n - \underbrace{\tilde{A}_{nn}}_{=1} \underbrace{\vec{v}_n \vec{v}_n}_{=0 \text{ n.V.}} - \underbrace{\tilde{A}_{nn-1}}_{=0 \text{ n.V.}} \underbrace{\vec{v}_n \vec{v}_{n-1}}_{=0 \text{ n.V.}}) = 0 \\
 \vec{v}_{n-1} \vec{v}_{n+1} &= C_{n+1}(\vec{v}_{n-1} A \vec{v}_n - \underbrace{\tilde{A}_{nn-1}}_{=1} \vec{v}_{n-1} \vec{v}_n - \underbrace{\tilde{A}_{nn-1}}_{=1} |\vec{v}_{n-1}|^2) = 0 \\
 \vec{v}_{n-2} \vec{v}_n &= C_{n+1} \vec{v}_{n-2} \underbrace{A \vec{v}_n}_{\in K(A, \vec{v}_n)} = 0
 \end{aligned} \tag{2.90}$$

Praktische Implementierung

1. Formulierung mit allen Basisvektoren \vec{v}_j (nach Iteration j):

Initialisierung: $\vec{v}_1 \leftarrow$ Zufallsvektor, normiert
 $\vec{v}_0 \leftarrow \vec{0}$
 $\beta_1 \leftarrow 0$

Schleife: $\vec{w}_j \leftarrow A\vec{v}_j - \beta_j \vec{v}_{j-1}$
 $\alpha_j \leftarrow \vec{w}_j \vec{v}_j$
 $\vec{w}_j \leftarrow \vec{w}_j - \alpha_j \vec{v}_j$
 $\beta_{j+1} \leftarrow |\vec{w}_j|$
 $\vec{v}_{j+1} \leftarrow \frac{\vec{w}_j}{\beta_{j+1}} \Leftarrow$ Normierung

$$\tilde{A} = \begin{pmatrix} \alpha_1 & \beta_2 & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \beta_3 & \alpha_3 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \alpha_{m-1} & \beta_m \\ & & & & \beta_m & \alpha_m \end{pmatrix} \tag{2.91}$$

2. Speichersparende Implementierung:

Initialisierung: \vec{v} Zufallsvektor, normiert
 $\vec{v}' = \vec{0}$
 $\beta_i = 0$
 $j = 0$

Schleife: $j \leftarrow j + 1$
 $\vec{w} \leftarrow A\vec{v} - \beta_j \vec{v}'$
 $\alpha_j \leftarrow \vec{w} \cdot \vec{v}$
 $\vec{w} \leftarrow \vec{w} - \alpha_j \vec{v}$
 $\beta_{j+1} \leftarrow |\vec{w}|$
 $\vec{v}' \leftarrow \vec{v}$
 $\vec{v} \leftarrow \frac{\vec{w}}{\beta_{j+1}}$

Kapitel 3

Quantum Monte Carlo (QMC) simulations

There are many different QMC methods; common feature: Monte Carlo evaluation of high-dimensional integrals relevant for quantum systems.

3.1 Classification of QMC methods

1. Systems:

- continuum systems (in real space)
- lattice systems (in particular: electrons)
- single impurity model (relevant for DMFT)
- quantum Heisenberg model ...

2. regime:

- finite temperatures (canonical ensemble)
- ground state (projective QMC)
- variational MC (upper bound on E_{GS})

3. approach:

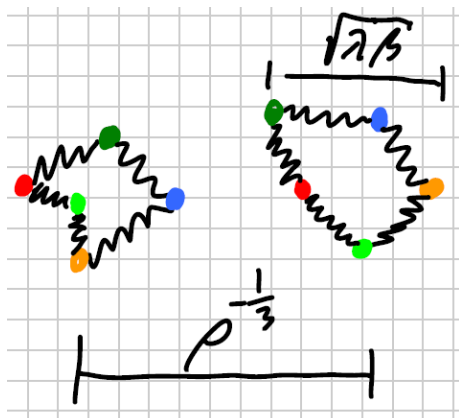
- wave function
- density matrix
- auxiliary field

4. statistics:

- „boltzmannons“ (distinguishable particles)
- bosons
- fermions

3.2 Path Integral Monte Carlo (PIMC)

Characterization:



- continuum system
- finite temperature
- based on density matrix
- works best for boltzmannons, good for bosons, fermions require fixed-node approximation

- employs analogy between path integrals and polymers
- interaction only between beads of different ring polymers at equal time“ (equal color)
- beads connected by springs

Foundation: quantum statistical mechanics In terms of *eigenstates* $|\Phi_i\rangle$, thermal expectation values of observables may be written as

$$\langle \hat{O} \rangle = Z^{-1} \sum_i \langle \Phi_i | \hat{O} | \Phi_i \rangle e^{-\beta E_i} \quad (3.1)$$

where $\beta = \frac{1}{k_B T}$, the partition function is

$$Z = \sum_i e^{-\beta E_i} \quad (3.2)$$

and

$$E_i = \langle \Phi_i | \hat{\mathcal{H}} | \Phi_i \rangle \quad (3.3)$$

Compare classical case:

$$\begin{aligned} Z_{class} &\propto \int d^{3N} r \int d^{3N} v e^{-\beta E(\{r\}, \{v\})} \\ E &= E_{pot} + E_{kin}; \quad E_{kin} = \sum_{i=1}^{3N} \frac{1}{2} m_i v_i^2; \quad E_{pot} = E_{pot}(\{r\}) \\ Z_{class} &\propto \left(\int d^{3N} r e^{-\beta E_{pot}(\{r\})} \right) \left(\int d^{3N} v e^{-\beta \frac{1}{2} m v^2} \right)^{3N} \\ &= Z_{pot} \left(\frac{2\pi}{\beta m} \right)^{\frac{3N}{2}} \end{aligned} \quad (3.4)$$

Kinetic part has trivial solution \rightarrow MC for static problem

For an arbitrary basis (in which the Hamiltonian is in general not diagonal):

$$\begin{aligned} \langle \hat{O} \rangle &= \frac{1}{Z} \sum_{\alpha} \langle \alpha | \hat{O} e^{-\beta \hat{\mathcal{H}}} | \alpha \rangle = \frac{1}{Z} \text{tr}(\hat{O} e^{-\beta \hat{\mathcal{H}}}) \\ &= \frac{1}{Z} \sum_{\alpha} \sum_{\alpha'} \langle \alpha | \hat{O} | \alpha' \rangle \langle \alpha' | e^{-\beta \hat{\mathcal{H}}} | \alpha \rangle \end{aligned} \quad (3.5)$$

with $Z = \sum_{\alpha} \langle \alpha | e^{-\beta \hat{\mathcal{H}}} | \alpha \rangle = \text{tr}(e^{-\beta \hat{\mathcal{H}}})$

Central operator: *density matrix* $e^{-\beta \hat{\mathcal{H}}}$

Now choose *position basis with labeled particles*.

Advantage: in this basis, all matrix elements are *real* and *positive* (proof: later)

Position-space density matrix:

$$\rho(R, R', \beta) = \langle R | e^{-\beta \hat{\mathcal{H}}} | R' \rangle = \sum_i \Phi_i^*(R) \Phi_i(R') e^{-\beta \mathcal{E}_i} \quad (3.6)$$

with position representations of the eigenstates

$$\Phi_i(R') = \langle i | R' \rangle; \quad \Phi_i^*(R) = \langle R | i \rangle = \langle i | R \rangle^* \quad (3.7)$$

and position vectors $R = \{\vec{r}_1, \dots, \vec{r}_N\}$.

In $d = 3$ dimensions ρ is a function of $6N + 1$ variables. In position basis, the expectation values become:

$$\langle \hat{O} \rangle = Z^{-1} \int dR dR' \rho(R, R', \beta) \langle R' | \hat{O} | R \rangle \quad (3.8)$$

with $Z = \int dR \rho(R, R', \beta)$

Note: the product of two density matrices is a density matrix (all for the same Hamiltonian!):

$$e^{-\beta_1 \hat{\mathcal{H}}} e^{-\beta_2 \hat{\mathcal{H}}} = e^{-(\beta_1 + \beta_2) \hat{\mathcal{H}}} \quad (3.9)$$

- in this case, operators can be treated like numbers since $[\hat{\mathcal{H}}, \hat{\mathcal{H}}] = 0$
- the product matrix is associated with a *lower temperature*

Written for positions \rightarrow *convolution*

$$\rho(R_1, R_3; \beta_1 + \beta_2) = \int dR_2 \rho(R_1, R_2; \beta_1) \rho(R_2, R_3; \beta_2) \quad (3.10)$$

3.2.1 Discrete path integrals

Now consider $\hat{\mathcal{H}} = \hat{T} + \hat{V}$ with the usual kinetic contribution (nonrelativistic, no magnetic field):

$$\hat{T} = \sum_{n=1}^N -\frac{1}{2m} \hat{p}_n^2 \quad (\hat{p}_n = -i\hbar \vec{\nabla}_n \text{ in position basis}) \quad (3.11)$$

and the potential contribution:

$$\hat{V} = \int dR \underbrace{V(R)}_{=V(\vec{R})} |R\rangle \langle R| \quad (3.12)$$

The fundamental problem in quantum mechanics is that we know eigenbasis and density matrix only for either the kinetic *or* the potential part of the Hamiltonian, not both:

$$\underbrace{e^{-\beta \hat{\mathcal{H}}}}_{\text{needed, unknown}} \neq \underbrace{e^{-\beta \hat{T}}}_{\text{known (see below)}} \underbrace{e^{-\beta \hat{V}}}_{\text{known (see below)}} \quad (\text{for general } \beta) \quad (3.13)$$

But: asymptotic factorization at high temperatures (classical limit), i.e. small β .
More generally: good high-T approximations for density matrix!

Notation: imaginary time step $\Delta\tau = \frac{\beta}{M}$ ($M \in \mathbb{N}$) Exact discrete path integral (M equal steps):

$$\rho(R_0, R_M, \beta) = \int dR_1 \int dR_2 \dots \int dR_{M-1} \rho(R_0, R_1; \Delta\tau) \quad (3.14)$$

$$\rho(R_1, R_2; \Delta\tau) \dots \rho(R_{M-1}, R_M; \Delta\tau) \quad (3.15)$$

We will now derive an explicit approximation for this path integral.

Campbell-Baker-Hausdorff formula:

$$\exp \left[-\Delta\tau(\hat{T} + \hat{V}) + \underbrace{\frac{1}{2}\Delta\tau^2[\hat{T}, \hat{V}] + \mathcal{O}(\Delta\tau^3)}_{\text{negligible in limit of small } \Delta\tau} \right] = e^{-\Delta\tau\hat{T}} e^{-\Delta\tau\hat{V}} \quad (3.16)$$

→ Trotter formula:

$$e^{-\beta(\hat{T}+\hat{V})} = \lim_{M \rightarrow \infty} [e^{-\Delta\tau\hat{T}} e^{-\Delta\tau\hat{V}}]^M; \quad \Delta\tau = \frac{\beta}{M} \quad (3.17)$$

$$\approx [e^{-\Delta\tau\hat{T}} e^{-\Delta\tau\hat{V}}]^M \text{ for } \Delta\tau \text{ small} \quad (3.18)$$

Primitive approximation in position space for single time step:

$$\rho(R_0, R_2; \Delta\tau) = \int dR_1 \langle R_0 | e^{-\Delta\tau\hat{T}} | R_1 \rangle \langle R_1 | e^{-\Delta\tau\hat{V}} | R_2 \rangle \quad (3.19)$$

Potential operator is diagonal, therefore:

$$\langle R_1 | e^{-\Delta\tau\hat{V}} | R_2 \rangle = e^{-\Delta\tau V(R_1)} \delta(R_1 - R_2) \quad (*)$$

For kinetic energy operator, introduce eigenfunction expansion in a cube with side length L and pbc, denoted by integer vector $\vec{n} \in \mathbb{R}^{3N}$:

$$\Phi_{\vec{n}}(\vec{R}) = \langle \vec{n} | \vec{R} \rangle = L^{-\frac{3N}{2}} e^{-i\vec{k}_{\vec{n}} \cdot \vec{R}}; \quad \vec{k}_{\vec{n}} = \frac{2\pi\vec{n}}{L} \text{ plane waves} \quad (3.20)$$

with eigen values $\frac{\hbar^2}{2m} \vec{k}_{\vec{n}}^2 \equiv \lambda \vec{k}_{\vec{n}}^2$

$$\begin{aligned} \langle R_0 | e^{-\Delta\tau\hat{T}} | R_1 \rangle &= \sum_{\vec{n}} \langle R_0 | \vec{n} \rangle \langle \vec{n} | e^{-\Delta\tau\hat{T}} | \vec{n} \rangle \langle \vec{n} | R_1 \rangle \\ &= \sum_{\vec{n}} L^{-\frac{3N}{2}} e^{-i\vec{k}_{\vec{n}} \cdot \vec{R}_0} e^{-\Delta\tau \lambda \vec{k}_{\vec{n}}^2} L^{-\frac{3N}{2}} e^{+i\vec{k}_{\vec{n}} \cdot \vec{R}_1} \\ &= \sum_{\vec{n}} L^{-3N} \exp \left[-\Delta\tau \lambda \vec{k}_{\vec{n}}^2 - i\vec{k}_{\vec{n}} \cdot (\vec{R}_0 - \vec{R}_1) \right] \\ &\approx (4\pi\lambda\Delta\tau)^{-\frac{3N}{2}} \exp \left[-\frac{(\vec{R}_0 - \vec{R}_1)^2}{4\lambda\Delta\tau} \right] \end{aligned} \quad (\Delta)$$

↑

$$\sum_{\vec{n}} \rightarrow \int d^{3N}n, \text{ good for } \lambda\Delta\tau \ll L^2 \text{ then Gauss integral}$$

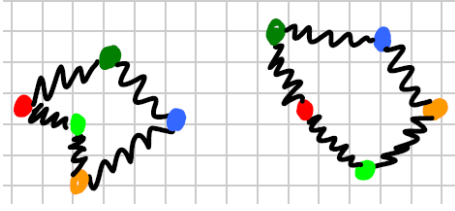
(exact kinetic density matrix: $\Theta_3(z, q) \dots$)

Insert (lowest order) Trotter approximation, using $*$ and Δ , into general discrete path integral:

$$\rho(\vec{R}_0, \vec{R}_M; \beta) = \int d\vec{R}_1 \int d\vec{R}_2 \dots \int d\vec{R}_{M-1} (4\pi\lambda\Delta\tau)^{-\frac{3NM}{2}} \quad (3.21)$$

$$\exp \left[- \sum_{m=1}^M \left(\frac{(\vec{R}_{m-1} - \vec{R}_m)^2}{4\lambda\Delta\tau} + \Delta\tau V(\vec{R}_m) \right) \right] \geq 0 \quad (3.22)$$

$$\text{primitive approximation for density matrix} \quad (3.23)$$



$Z = \int d\vec{R} \rho(\vec{R}, \vec{R}; \beta)$ corresponds to classical partition function of ring polymers with ideal springs within each polymer and interpolymer interaction only at equal times ($\hat{=}$ color).

More explicitly: position-dependent part of Z for pair interactions

$$V(\vec{R}_m) = \sum_{i=1}^N \sum_{j=i+1}^N v(\vec{r}_i(\tau_m) - \vec{r}_j(\tau_m)) \quad (3.24)$$

$$\text{is } \int d\vec{R}_1 \int d\vec{R}_2 \dots \int d\vec{R}_M e^{-\beta \tilde{V}(\{\vec{R}\})}$$

with

$$\tilde{V}(\{\vec{R}\}) = \sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{M} \sum_{m=1}^M v(\vec{r}_i(\tau_m) - \vec{r}_j(\tau_m)) \quad (3.25)$$

$$+ \sum_{i=1}^N \underbrace{\frac{M}{4\lambda\beta^2}}_{\frac{k}{2}} \sum_{m=1}^M (\vec{r}_i(\tau_{m-1}) - \vec{r}_i(\tau_m))^2 \quad (3.26)$$

3.2.2 Fractal nature

In the case of a smooth imaginary-time path $\vec{r}(\tau)$ we would expect $|\vec{r}(\tau_i) - \vec{r}(\tau_j)|^2 \leq v_{max}^2 |\tau_i - \tau_j|^2$ however, in limit $V \rightarrow 0$ we get

$$\langle |\vec{r}_i(\tau_m) - \vec{r}_i(\tau_{m-1})|^2 \rangle = 3(2\lambda\Delta\tau) \quad (3.27)$$

\rightarrow in the limit $\Delta\tau \rightarrow 0$, each polymer acquires infinite length! (smoothing + shading)

3.2.3 Role of temperature

In the polymer model, the finite size of each polymer is an entropic effect: for fixed spring constant k , the size would vanish for $T \rightarrow 0$.

However: for fixed M ; $k \propto T^2 \Rightarrow |\vec{r}_i(\tau_m) - \vec{r}_i(\tau_{m-1})|^2 \propto T^{-1} \rightarrow \infty$
 for fixed τ : $k \propto T \Rightarrow |\vec{r}_i(\tau_m) - \vec{r}_i(\tau_{m-1})|^2 \propto T^0 \rightarrow \text{const}$ and $M \rightarrow \infty$

in both cases: polymer diameter $\propto \sqrt{\lambda\beta} \rightarrow \infty$

3.2.4 Connection to Feynman-Kac path integral

For simplicity, write only for Z (analogous for ρ):

$$\begin{aligned}
 Z &= \left\langle (4\pi\lambda\Delta\tau)^{-\frac{3NM}{2}} \prod_{m=1}^M e^{-\frac{|\vec{R}_{m-1} - \vec{R}_m|^2}{4\lambda\Delta\tau}} e^{-\Delta\tau V(\vec{R}_m)} \right\rangle_{\text{uniform}} \\
 &= Z_{\text{free}} \left\langle \prod_{m=1}^M e^{-\Delta\tau V(\vec{R}_m)} \right\rangle_{\text{free}} \quad \{R_m\} \text{ generated according to } Z \text{ for } V = 0 \\
 &= Z_{\text{free}} \left\langle \exp \left[- \sum_{m=1}^M \Delta\tau V(\vec{R}_m) \right] \right\rangle_{\text{free}} \\
 &\xrightarrow{M \rightarrow \infty} Z_{\text{free}} \langle \exp[-\int_0^\beta V(\vec{R}(\tau)) d\tau] \rangle_{\text{free}}
 \end{aligned} \tag{3.28}$$

$$e^{-\beta(F - F_{\text{free}})} = \frac{Z}{Z_{\text{free}}} = \langle \exp[-\int_0^\beta V(\vec{R}(\tau)) d\tau] \rangle_{\text{free}} \tag{3.29}$$

3.2.5 Monte Carlo sampling of the partition function

Not possible: naive simple sampling (take all $\frac{r_i}{\tau_m}$ as independent random variables in L^3)

1. single particle displacement moves (\sim classical atoms)

- (a) select particle i
- (b) draw displacement $\vec{\delta}$ from uniform distribution inside cube $[-\frac{\Delta}{2}, \frac{\Delta}{2}]^3$
- (c) propose move $\vec{r}_i(\tau_m) \longrightarrow \vec{r}_i(\tau_m) + \vec{\delta}$ for all m
- (d) accept with probability $p = \min\{1, e^{-\beta\Delta V}\}$

- These moves are only usefull for not too low T
- Δ is adjusted for acceptance probability of 25% – 75%

2. single-particle single-slice moves

- i. select particle i and slice m
- (a) classic rule
 - ii. random displacement $\vec{\delta}$ uniformly from cube $[-\frac{\Delta}{2}, \frac{\Delta}{2}]^3$

- iii. propose move $\vec{r}_i(\tau_m) \longrightarrow \vec{r}_i(\tau_m) + \vec{\delta}$ for single m
- iv. accept with probability $p = \min\{1, e^{-\beta\Delta\tilde{V}}\}$
 - now also kinetic contributions to $\Delta\tilde{V}$
 - adjust $\Delta \approx \sqrt{\lambda\Delta\tau}$ to $\sim 50\%$ acceptance prob.

(b) free-particle sampling

In general, the probability density for position vector \vec{R}_m at fixed \vec{R}_i for $i \neq m$ can be written as

$$P[\vec{R}_m] \propto \rho(\vec{R}_{m-1}, \vec{R}_m; \Delta\tau) \rho(\vec{R}_m, \vec{R}_{m+1}; \Delta\tau) \quad (3.30)$$

specifically for $\hat{V} = 0$ we have

$$\begin{aligned} P[\vec{R}_m] &\propto \exp\left[-\frac{|\vec{R}_{m-1} - \vec{R}_m|^2}{4\lambda\Delta\tau}\right] \exp\left[-\frac{|\vec{R}_m - \vec{R}_{m+1}|^2}{4\lambda\Delta\tau}\right] \\ (x-a)^2 + (x-b)^2 &= 2x^2 - 2x(a+b) + a^2 + b^2 \\ &= 2\left[x^2 - 2x\frac{a+b}{2} + \left(\frac{a+b}{2}\right)^2\right] + \frac{(a-b)^2}{2} \\ \Rightarrow P[\vec{R}_m] &\propto \exp\left[-\frac{|\vec{R}_m - \vec{R}_{av}|^2}{2\lambda\Delta\tau}\right]; \quad \vec{R}_{av} = \frac{\vec{R}_{m-1} + \vec{R}_{m+1}}{2} \end{aligned} \quad (3.31)$$

in case of single particle move:

$$P[r_i(\tau_m)] \propto \exp\left[-\frac{|\vec{r}_i(\tau_m) - \vec{r}_i^{av}|^2}{2\lambda\Delta\tau}\right]; \quad \vec{r}_i^{av} = \frac{\vec{r}_i(\tau_{m+1}) + \vec{r}_i(\tau_{m-1})}{2} \quad (3.32)$$

This can be sampled directly (diffusion step)

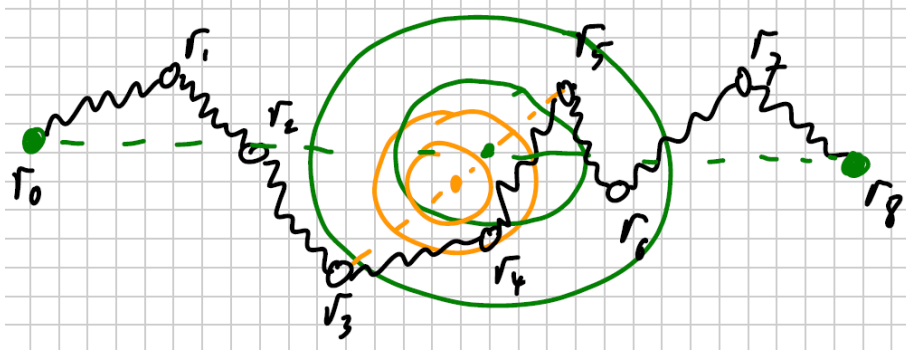
- ii. draw $\vec{\delta}$ from Gaussian distribution with width $\sqrt{\lambda\Delta\tau}$ (in each dimension)
- iii. propose move $\vec{r}_i(\tau_m) \rightarrow \frac{\vec{r}_i(\tau_{m+1}) + \vec{r}_i(\tau_{m-1})}{2} + \vec{\delta}$
- iv. accept with probability $p = \min\{1, e^{-\beta\Delta V}\}$
(this expression is valid for primitive approximation; generalize for better actions)

procedure analogous to diffusion Monte Carlo

$$|O\rangle = \lim_{\beta \rightarrow \infty} e^{-\beta\hat{\mathcal{H}}} |\Psi_0\rangle \approx \prod_{m=1}^M \left(\underset{\substack{\uparrow \\ \text{diffusion} \\ \text{of „walkers“}}}{e^{-\Delta\tau\hat{T}}} \underset{\substack{\uparrow \\ \text{birth/death} \\ \text{of „walkers“}}}{e^{-\Delta\tau\hat{V}}} \right) |\Psi_0\rangle \quad (3.33)$$

1) and 2) are in principle sufficient for distinguishable particles. However: single-slice moves inefficient for large M

2'. Multilevel sampling



On fine grid: \vec{r}_4 can not move very far, idea:

- consider coarse grid with $\tilde{\Delta}\tau = 4\Delta\tau$ first
- construct move iteratively
- keep track of intermediate proposal prob.
- correct everything in final acceptance rule

3. Exchange moves: for distinguishable particles, we must allow exchanges (discrete move - not possible in MD)



In practice: combine with multilevel moves.

3.2.6 Observables

- density, pair correlation function, structure factor: directly from \vec{R}_m
- already the energy is nontrivial

1. „direct“ or „Hamiltonian“ estimator:

$$E_{\mathcal{H}} = \frac{\text{Tr}(\hat{\mathcal{H}} e^{-\beta\hat{\mathcal{H}}})}{\text{Tr}(e^{-\beta\hat{\mathcal{H}}})} = \dots \quad (3.34)$$

2. thermodynamic estimator:

$$\begin{aligned} E_T &= -\frac{1}{Z} \frac{dZ}{d\beta} \\ &= \frac{1}{M} \sum_{m=1}^M \left\{ \frac{3N}{2\Delta\tau} - \frac{|\vec{R}_m - \vec{R}_{m-1}|^2}{4\lambda\Delta\tau^2} + V(\vec{R}_m) \right\} \end{aligned} \quad (3.35)$$

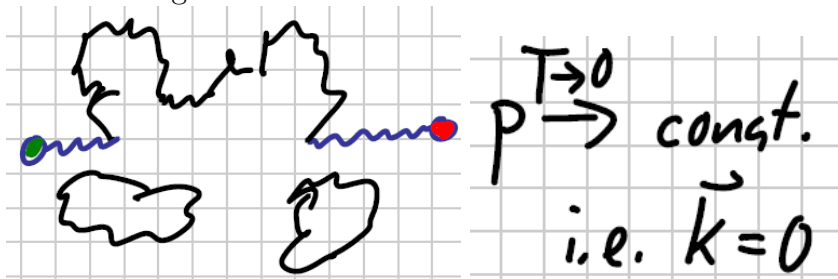
3. virial estimator:

$$E_V = \left\langle \dots - \frac{1}{2} \overset{\substack{\text{derivation from} \\ \text{center of mass}}}{\underset{\substack{\text{generalized force}}}{F^m}} \Delta_m \dots \right\rangle \quad (3.36)$$

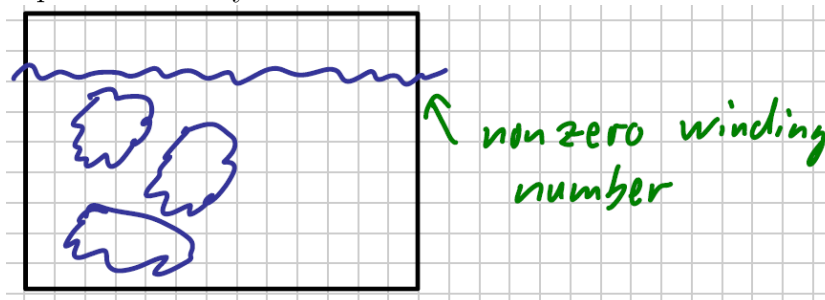
- condensate fraction
insert two free ends \vec{r}_0, \vec{r}_+
measure $\text{prob}(\vec{r}_0 - \vec{r}_+)$
no exchange:



with exchange:



- superfluid density



3.2.7 Application to vacancies in solid ^4He

see slides: [LINK](#)

3.3 Weltlinien-Quanten-Monte-Carlo-Methode (world line QMC)

Weltlinien-QMC ist ein konzeptionell relativ einfacher Zugang für Modelle wechselwirkender Quanten-Spins oder Fermionen, der erst Simulationen erlaubt und Basis

für effizientere Algorithmen (Schleifen-, Wurm-Algorithmen) ist.
Als Anwendungsbeispiel wählen wir das XXZ-Modell:

$$\mathcal{H} = \sum_{i=1}^L \mathcal{H}^{(i)}; \quad \mathcal{H}^{(i)} = \frac{J_x}{a} (S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+) + J_z (S_i^z + S_{i+1}^z) \quad (3.37)$$

mit $S_i^x = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $S_i^y = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $S_i^z = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

und periodischen Randbedingungen: $\vec{S}_{L+1} \equiv \vec{S}_1$

Basis des Hilbert-Raums sind $\sigma_i \in \{|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$

Mit $S^+ = S^x + iS^y$, $S^- = S^x - iS^y$ gilt:

$$\mathcal{H} = \frac{J_x}{2} \sum_{i=1}^L (S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+) + J_z \sum_{i=1}^L (S_i^z S_{i+1}^z) \quad (3.38)$$

Für 2 Teilchen mit offenen Randbed.:

$$\mathcal{H}_{12} \hat{=} \begin{pmatrix} \uparrow\uparrow & \uparrow\downarrow & \downarrow\uparrow & \downarrow\downarrow \\ \frac{J_z}{4} & 0 & 0 & 0 \\ 0 & -\frac{J_z}{4} & \frac{J_x}{2} & 0 \\ 0 & \frac{J_x}{2} & -\frac{J_z}{4} & 0 \\ 0 & 0 & 0 & \frac{J_z}{4} \end{pmatrix} \begin{matrix} \uparrow\uparrow \\ \uparrow\downarrow \\ \downarrow\uparrow \\ \downarrow\downarrow \end{matrix} \quad (3.39)$$

Damit ergeben sich für Eigenvektoren und Eigenwerte des 2-Spin-Systems:

\vec{S}^2	S^z	EV	EW	
1	1	$ \uparrow\uparrow\rangle$	$\frac{J_z}{4}$	
1	-1	$ \downarrow\downarrow\rangle$	$\frac{J_z}{4}$	
1	0	$\frac{1}{\sqrt{2}}(\uparrow\downarrow\rangle + \downarrow\uparrow\rangle)$	$\frac{J_x}{2} - \frac{J_z}{4}$	(*)
0	0	$\frac{1}{\sqrt{2}}(\uparrow\downarrow\rangle - \downarrow\uparrow\rangle)$	$-\frac{J_x}{2} - \frac{J_z}{4}$	

Grundidee von Weltlinien-Verfahren ist es nun, den Hamilton-Operator so aufzuteilen, dass jeder der Teile in kommutierende 2-Platz-Terme zerfällt. Bei Modellen, in denen \mathcal{H} maximal nächste Nachbarn auf einem hyperkubischen Gitter verbindet, sind das 2d Teile: $\mathcal{H} = \mathcal{H}_1^x + \mathcal{H}_2^x + \mathcal{H}_1^y + \mathcal{H}_2^y + \dots$

Wichtig: Notwendige Voraussetzung ist, dass die Systemlänge in jeder Richtung gerade ist: $L(\alpha) = 2n_\alpha$ Speziell im vorliegenden 1-d-Fall:

$$\mathcal{H} = \underbrace{\sum_{n=1}^{\frac{L}{2}} \mathcal{H}^{(2n)}}_{\mathcal{H}_1} + \underbrace{\sum_{n=1}^{\frac{L}{2}} \mathcal{H}^{(2n)}}_{\mathcal{H}_2} \quad (3.40)$$



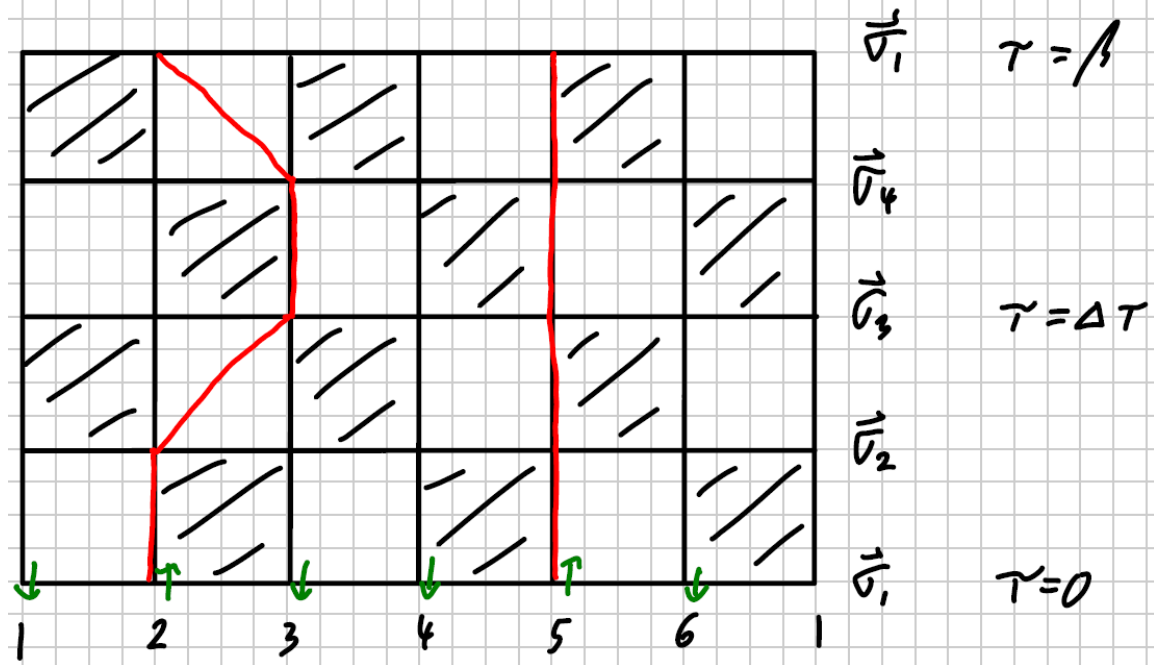
Diese Aufteilung wird nun in einer Trotter-Zerlegung der Zustandssumme benutzt ($\Delta\tau = \frac{\beta}{M}$):

$$\begin{aligned}
 Z &= \text{Sp}[e^{-\beta\mathcal{H}}] = \text{Sp}[(e^{-\Delta\tau\mathcal{H}})^M] \\
 &\approx \text{Sp}[(e^{-\Delta\tau\mathcal{H}_1}e^{-\Delta\tau\mathcal{H}_2})^M] \\
 &= \sum_{\vec{\sigma}_1 \dots \vec{\sigma}_2} \langle \vec{\sigma}_1 | e^{-\beta\mathcal{H}_1} | \vec{\sigma}_{2M} \rangle \langle \vec{\sigma}_{2M} | e^{-\beta\mathcal{H}_2} | \vec{\sigma}_{2M-1} \rangle \\
 &\quad \langle \vec{\sigma}_3 | e^{-\beta\mathcal{H}_1} | \vec{\sigma}_2 \rangle \langle \vec{\sigma}_2 | e^{-\beta\mathcal{H}_2} | \vec{\sigma}_1 \rangle
 \end{aligned} \tag{3.41}$$

mit dem Gesamtspinvektor $\vec{\sigma}_\tau$ zur imaginären Zeit τ :

$$\vec{\sigma}_\tau = (\sigma_{1,\tau}, \sigma_{2,\tau}, \dots, \sigma_{L,\tau}) \tag{3.42}$$

Graphische Repräsentation: Weltlinien-Evolution von \uparrow -Zuständen (bei Fermion-Simulationen: von besetzten Zuständen) in der imaginären Zeit:



$L = 6$ Spins, davon 2 up-Spins ($S^z = -1$)

$M = 2$ Diskretisierungsschritte der imaginären Zeit.

Wir führen noch einige Bezeichnungen ein:

Konfiguration: $w = (\vec{\sigma}_1, \vec{\sigma}_2, \dots, \vec{\sigma}_{2M})$

Gewicht einer Konfiguration:

$$\begin{aligned}
 \Omega(w) &= \langle \vec{\sigma}_1 | e^{-\beta\mathcal{H}_1} | \vec{\sigma}_{2M} \rangle \langle \vec{\sigma}_{2M} | e^{-\beta\mathcal{H}_2} | \vec{\sigma}_{2M-1} \rangle \dots \\
 &\quad \langle \vec{\sigma}_3 | e^{-\beta\mathcal{H}_1} | \vec{\sigma}_2 \rangle \langle \vec{\sigma}_2 | e^{-\beta\mathcal{H}_2} | \vec{\sigma}_1 \rangle
 \end{aligned} \tag{3.43}$$

Damit gilt: $Z = \sum_w \Omega(w)$

Die Matrixwlemente in $\Omega(w)$ faktorisieren:

$$\langle \vec{\sigma}_{\tau+1} | e^{-\Delta\tau\mathcal{H}_2} | \vec{\sigma}_\tau \rangle = \prod_{i=1}^{\frac{L}{2}} \langle \sigma_{2i,\tau+1}, \sigma_{2i+1,\tau+1} | e^{-\Delta\tau\mathcal{H}^{(2i)}} | \sigma_{2i,\tau}, \sigma_{2i+1,\tau} \rangle \tag{3.44}$$

Damit ist die Berechnung der Gewichte auf das 2-Platz-Problem zurückgeführt. Z.B. erhalten wir mit *:

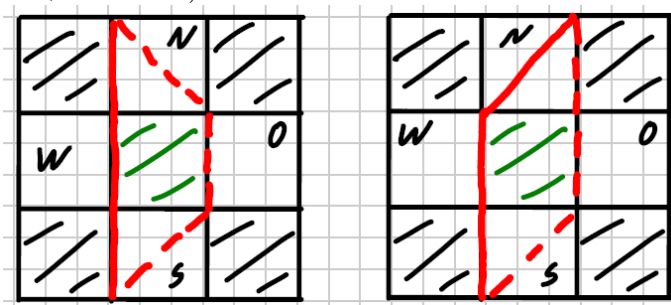
$$\begin{aligned}
 \langle \downarrow \uparrow | e^{-\Delta\tau \mathcal{H}_2} | \uparrow \downarrow \rangle &= \frac{1}{\sqrt{2}} \langle \downarrow \uparrow | e^{-\Delta\tau \mathcal{H}_2} \left[\frac{1}{\sqrt{2}} (| \uparrow \downarrow \rangle - | \downarrow \uparrow \rangle) + \frac{1}{\sqrt{2}} (| \uparrow \uparrow \rangle + | \downarrow \downarrow \rangle) \right] \\
 &= \frac{1}{\sqrt{2}} \langle \downarrow \uparrow | e^{-\Delta\tau (-\frac{J_z}{4} - \frac{J_x}{2})} \frac{1}{\sqrt{2}} (| \uparrow \downarrow \rangle - | \downarrow \uparrow \rangle) \\
 &\quad + e^{-\Delta\tau (-\frac{J_z}{4} + \frac{J_x}{2})} \frac{1}{\sqrt{2}} (| \uparrow \uparrow \rangle + | \downarrow \downarrow \rangle) \rangle \\
 &= -e^{\Delta\tau \frac{J_z}{4}} \sinh \left(\frac{\Delta\tau J_x}{2} \right)
 \end{aligned} \tag{3.45}$$

Analog erhält man für die Plaketten:

Beachte: negative Gewichte würden i.A. zu einem Vorzeichenproblem führen. Bei nächst-Nachbar-Hüpfen auf bipartiten Gittern ist die Anzahl der Diagonalplaketten jedoch in jeder Konfiguration gerade, das Vorzeichen also irrelevant.

Weltlinien-Algorithmus

1. initialisiere Weltlinien (z.B. senkrecht) für gewünschtes L, S_z
2. Wähle schattierte Plakette, schlage Update vor (möglich, falls schatt. Plakette AF, stationär)



3. Berechne Verhältnis der Gewichte aus 4 umgebenden weissen Plaketten für w_{alt}, w_{neu} :

$$r = \frac{\Omega(w_{neu})}{\Omega(w_{alt})} = \frac{(\boxed{N} \cdot \boxed{O} \cdot \boxed{S} \cdot \boxed{W})_{neu}}{(\boxed{N} \cdot \boxed{O} \cdot \boxed{S} \cdot \boxed{W})_{alt}} \tag{3.46}$$

Akzeptiere Update mit W' $p = \frac{r}{1+r}$ (heat bath)
 (Alternative: Metropolis - $p = \min\{r, 1\}$)

Berechnung von Observablen $\langle O \rangle = \frac{\text{Sp}[e^{-\beta \mathcal{H}} O]}{\text{Sp}[e^{-\beta \mathcal{H}}]}$

Für Observablen, die lokal in der imaginären Zeit sind, kann man schreiben: $\langle O \rangle = \frac{\sum_w \Omega(w) O(w)}{\sum_w \Omega(w)}$

Die Kunst ist es nun, einen Schätzer $O(w)$ zu bestimmen, der (i) korrekt ist und (ii) möglichst kleine Varianz hat.

3.3.1 Berechnung der Energie

$$\begin{aligned} \langle \mathcal{H} \rangle &= \frac{1}{Z} \text{Sp} \left[(e^{-\Delta\tau \mathcal{H}_1} e^{-\Delta\tau \mathcal{H}_2})^M (\mathcal{H}_1 + \mathcal{H}_2) \right] \\ &= \frac{1}{Z} \text{Sp} \left[(e^{-\Delta\tau \mathcal{H}_1} e^{-\Delta\tau \mathcal{H}_2})^{M-1} (e^{-\Delta\tau \mathcal{H}_1} \mathcal{H}_1 e^{-\Delta\tau \mathcal{H}_2} + e^{-\Delta\tau \mathcal{H}_1} e^{-\Delta\tau \mathcal{H}_2} \mathcal{H}_2) \right] \end{aligned} \quad (3.47)$$

Durch Einschieben von Einheitsoperatoren $1 = \sum_\sigma |\vec{\sigma}\rangle \langle \vec{\sigma}|$ erhalten wir:

$$\begin{aligned} \langle \mathcal{H} \rangle &= \frac{1}{Z} \sum_{\vec{\sigma}_1, \dots, \vec{\sigma}_{2M}} \left[\langle \vec{\sigma}_1 | e^{-\Delta\tau \mathcal{H}_1} | \vec{\sigma}_{2M} \rangle \dots \langle \vec{\sigma}_3 | e^{-\Delta\tau \mathcal{H}_1} | \vec{\sigma}_2 \rangle \langle \vec{\sigma}_2 | e^{-\Delta\tau \mathcal{H}_2} \mathcal{H}_2 | \vec{\sigma}_2 \rangle \right. \\ &\quad \left. + \langle \vec{\sigma}_1 | e^{-\Delta\tau \mathcal{H}_1} | \vec{\sigma}_{2M} \rangle \dots \langle \vec{\sigma}_3 | e^{-\Delta\tau \mathcal{H}_1} \mathcal{H}_1 | \vec{\sigma}_2 \rangle \langle \vec{\sigma}_2 | e^{-\Delta\tau \mathcal{H}_2} | \vec{\sigma}_1 \rangle \right] \\ &= \frac{1}{Z} \sum_{\vec{\sigma}_1, \dots, \vec{\sigma}_{2M}} \langle \vec{\sigma}_1 | e^{-\Delta\tau \mathcal{H}_1} | \vec{\sigma}_{2M} \rangle \dots \langle \vec{\sigma}_3 | e^{-\Delta\tau \mathcal{H}_1} | \vec{\sigma}_2 \rangle \langle \vec{\sigma}_2 | e^{-\Delta\tau \mathcal{H}_2} | \vec{\sigma}_1 \rangle \\ &\quad \underbrace{\left[\frac{\langle \vec{\sigma}_3 | e^{-\Delta\tau \mathcal{H}_1} \mathcal{H}_1 | \vec{\sigma}_2 \rangle}{\langle \vec{\sigma}_3 | e^{-\Delta\tau \mathcal{H}_1} | \vec{\sigma}_2 \rangle} + \frac{\langle \vec{\sigma}_2 | e^{-\Delta\tau \mathcal{H}_2} \mathcal{H}_2 | \vec{\sigma}_1 \rangle}{\langle \vec{\sigma}_2 | e^{-\Delta\tau \mathcal{H}_2} | \vec{\sigma}_1 \rangle} \right]}_{E(w)} \\ &= \frac{\sum_w \Omega(w) E(w)}{\sum_w \Omega(w)} \equiv \{E(w)\}_w \end{aligned} \quad (3.48)$$

Die Matrixelemente lassen sich als Summen über (weisse) Plaketten schreiben, z.B.:

$$\begin{aligned} \frac{\langle \vec{\sigma}_2 | e^{-\Delta\tau \mathcal{H}_2} \mathcal{H}_2 | \vec{\sigma}_1 \rangle}{\langle \vec{\sigma}_2 | e^{-\Delta\tau \mathcal{H}_2} | \vec{\sigma}_1 \rangle} &= \sum_{i=1}^{\frac{L}{2}} \frac{\langle \sigma_{2i, \tau+1}, \sigma_{2i+1, \tau+1} | e^{-\Delta\tau \mathcal{H}^{(2i)}} \mathcal{H}^{(2i)} | \sigma_{2i, \sigma}, \sigma_{2i+1, \tau} \rangle}{\langle \sigma_{2i, \tau+1}, \sigma_{2i+1, \tau+1} | e^{-\Delta\tau \mathcal{H}^{(2i)}} | \sigma_{2i, \sigma}, \sigma_{2i+1, \tau} \rangle} \\ &= \sum_{i=1}^{\frac{L}{2}} -\frac{\partial}{\partial \Delta T} \left[\underbrace{\langle \sigma_{2i, \tau+1}, \sigma_{2i+1, \tau+1} | e^{-\Delta\tau \mathcal{H}^{(2i)}} | \sigma_{2i, \sigma}, \sigma_{2i+1, \tau} \rangle}_w \right] \end{aligned} \quad (3.49)$$

Zuerst berechnen wir das neue Matrixelement direkt (d.h. gemäß erster Zeile) für $\begin{array}{|c|c|} \hline \blacksquare & \blacksquare \\ \hline \end{array}$:

$$\begin{aligned}
 \langle \uparrow\uparrow | e^{-\Delta\tau \mathcal{H}_2} \mathcal{H}_2 | \uparrow\uparrow \rangle &= \frac{1}{\sqrt{2}} \langle \uparrow\uparrow | e^{-\Delta\tau \mathcal{H}_2} \left[\frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) + \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \right] \\
 &= \frac{1}{\sqrt{2}} \langle \uparrow\uparrow | \left[e^{-\Delta\tau(-\frac{J_z}{4} - \frac{J_x}{2})} \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \right. \\
 &\quad \left. + e^{-\Delta\tau(-\frac{J_z}{4} + \frac{J_x}{2})} \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \right] \\
 &= -e^{\Delta\tau \frac{J_z}{4}} \left[\sinh\left(\frac{\Delta\tau J_x}{2}\right) \left(-\frac{J_z}{4}\right) + \cosh\left(\frac{\Delta\tau J_x}{2}\right) \left(-\frac{J_x}{2}\right) \right]
 \end{aligned} \tag{3.50}$$

Allerdings ist es praktischer, die logarithmische Ableitung für die Berechnung des Quotienten zu benutzen:

Configuration	H_2	H_1
$\begin{array}{ c c } \hline \square & \square \\ \hline \end{array}$	$e^{-\Delta\tau J_z/4}$	$-J_z/4$
$\begin{array}{ c c } \hline \square & \blacksquare \\ \hline \end{array}$	$e^{\Delta\tau J_z/4} \cosh(\Delta\tau J_x/2)$	$-J_z/4 - \frac{J_x}{2} \tanh(\Delta\tau \frac{J_x}{2})$
$\begin{array}{ c c } \hline \blacksquare & \square \\ \hline \end{array}$	$e^{\Delta\tau J_z/4} \sinh(\Delta\tau J_x/2)$	$-J_z/4 - \frac{J_x}{2} \coth(\Delta\tau \frac{J_x}{2})$
$\begin{array}{ c c } \hline \blacksquare & \blacksquare \\ \hline \end{array}$		

Handwritten notes: $E_{loc} = -\frac{\partial}{\partial \Delta\tau} \ln[\omega]$

Insgesamt erhalten wir den Erwartungswert der Energie für ein Sample (eine Konfiguration w) als Summe über die Plaketten in den unteren beiden Reihen:

$$\langle \mathcal{H} \rangle_{w, \text{einzel}} = \sum_{\substack{\text{Plakette } p \text{ mit} \\ \text{Zeiten } \tau_1, \tau_2 \\ \text{bzw. } \tau_2, \tau_3}} E_{loc}(p) \tag{3.51}$$

Um die Varianz zu reduzieren, dürfen wir wegen der Äquivalenz der imaginären Zeiten über alle Plaketten summieren. Für N Messungen (z.B. nach jeweils einem vollen Sweep, bei dem alle schwarzen Plaketten betrachtet wurden) erhalten wir:

$$\langle \mathcal{H} \rangle = \frac{1}{N} \sum_{\substack{\text{sweeps} \\ s=1}}^N \frac{1}{M} \sum_{\substack{\text{alle} \\ \text{Platten } p}} E_{loc}(p, s) \tag{3.52}$$

NN-Spinkorrelation: $\langle \sigma_i^2 \sigma_{i+1}^2 \rangle = \frac{1}{N} \sum_s \frac{1}{M} \sum_p C_{loc}(p, s)$ mit

$$\begin{aligned}
 C_{loc} &= 1 \text{ für } p = \begin{array}{|c|c|} \hline \square & \blacksquare \\ \hline \end{array} \\
 C_{loc} &= -1 \text{ sonst}
 \end{aligned}$$

3.4 WL-QMC mit Loop-Updates

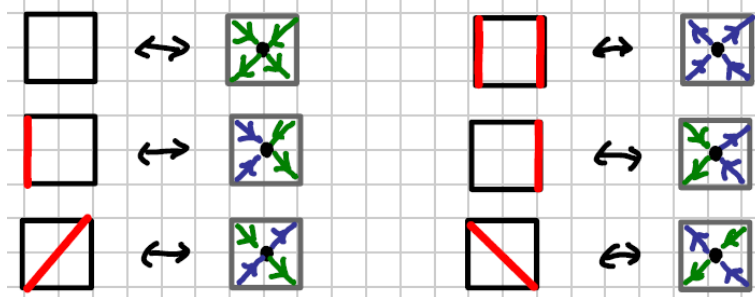
Probleme des in 3.3 vorgestellten Weltlinien-QMC-Verfahrens mit lokalen Updates:

1. konstante Anzahl der Weltlinien ($\rightarrow S^z$ konstant im XXZ-Modell, n konstant in Hubbard-Modellen)
2. Autokorrelationszeiten divergieren für $\Delta\tau \rightarrow 0$ ($\Rightarrow \Delta\tau$ -Extrapolation schwierig+teuer)

Beide Nachteile werden in Schleifen (=Loop)-Algorithmen überwunden.

3.4.1 Abbildung auf 6-Vertex-Modell

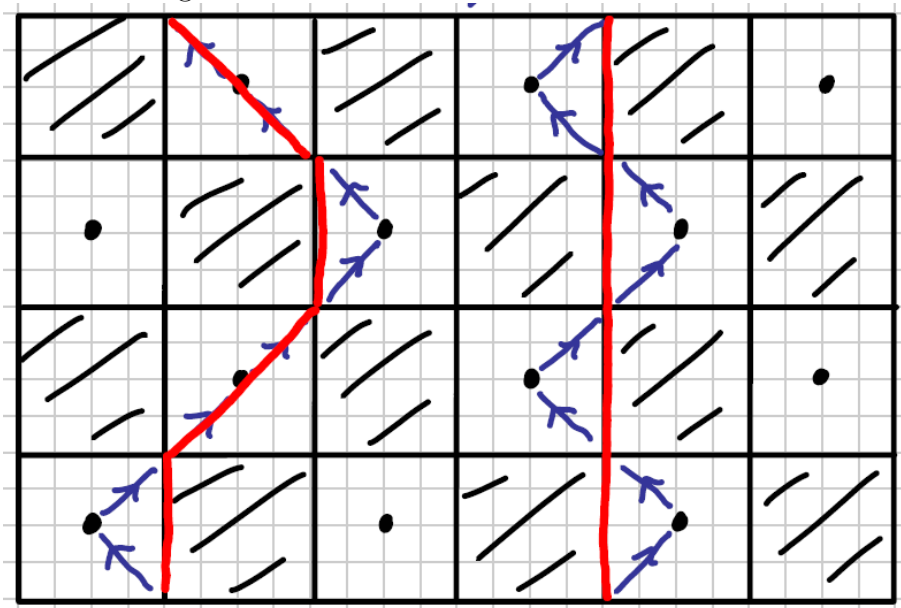
Im ersten Schritt werden die Weltlinien-Konfigurationen auf Konfigurationen des 6-Vertex-Modells abgebildet. Dieses besteht aus Vertices auf einem um 45 Grad rotierten Gitter (Mittelpunkt der weissen Plaketten), die divergenzfrei durch Pfeile verbunden sind. Auf der Ebene der Plaketten gilt folgende Zuordnung:



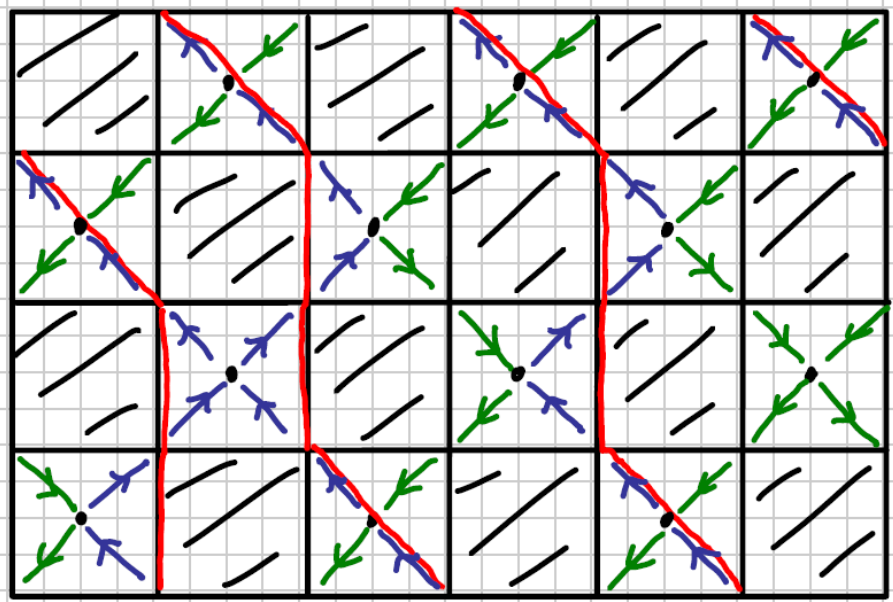
Grundsätzlich gilt also, dass aufwärtsgerichtete Pfeile zu Weltlinien korrespondieren.

Anwendungsbeispiel (der Klarheit halber sind beide Systeme übereinander gezeichnet):

1. nur aufwärtsgerichtete Pfeile



4. Spin-Flip der Schleife und Weltlinien-Update:



Mit dem Update hat sich also (i) die Zahl der Weltlinien erhöht und (ii) ist jetzt die Windungszahl (in L -Richtung) endlich.

Im Prinzip könnte man ein Loop-QMC-Verfahren wie im Beispiel beschrieben implementieren:

1. Initialisierung etc.
2. Für Konfiguration w : wähle Loop zufällig $\Rightarrow w_{neu}$
3. Berechne $\frac{\Omega(w_{neu})}{\Omega(w)}$
4. Akzeptiere z.B. nach heat bath-Regel
5. Messe Observablen
6. zurück zu 2.

Problem:

- Schritt 3. ist i.A. sehr teuer
- Akzeptanzrate in 4. ist oft niedrig

Lösung: Verlagere Akzeptanzschritt auf Konstruktion der Schleife (analog zur Clusterkonstruktion bei Fortuin-Kasteleyn). Führe dazu neue Variablen ein, sogenannte Graphen, die gegebenen WL- bzw Vertex-Konfigurationen mit einer zu bestimmenden Wahrscheinlichkeit zugewiesen werden. Dabei definiert ein Satz von zulässigen Graphen eindeutig Schleifen, analog den F-K-Clustern im Ising-Fall. Umgekehrt sollen zu einem gegebenen Graphen alle kompatiblen Vertex-Konfigurationen gleichwahrscheinlich sein.

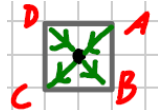
Konkret muss man in diesem Ansatz maximal 4 Plaketten-Graphen einführen:



Dabei gilt die Kompatibilitätstabelle:

	①	②	③	④
↔	✓	—	✓	✓
↔	✓	—	✓	✓
↔	—	✓	✓	✓
↔	✓	✓	✓	—
↔	✓	✓	✓	—

Zur Logik: die Linien in den Graphen sind ungerichtet und müssen möglichen Vertex-Pfaden entsprechen. Dies gilt allerdings nicht für den eingefrorenen Graphen, der dem Umklappen aller Spins entspricht und immer kompatibel ist. Als Beispiel betrachten wir die 1. Zeile:



Es gibt Pfade $A \rightarrow B, A \rightarrow C, D \rightarrow B, D \rightarrow C$, aber z.B. keine Verbindung $A \rightarrow D$ oder $D \rightarrow A$ — kein Graph

Formalismus: für die betrachtete Plakette sei S eine Spinkonfiguration, G ein Graph und $W(S)$ das in 3.3 bestimmte Gewicht zu S .

Es sollen nun Gewichte $W(S, G)$ für jede kompatible Konfiguration $\{S, G\}$ bestimmt werden, so dass gilt:

1. $\sum_G W(S, G) W(S)$
2. $W(S, G) = W(S', G) \forall S, S'$ kompatibel mit G

Wenn dieses lineare Gleichungssystem gelöst ist, lautet der Algorithmus wie folgt:

1. Initialisiere Weltlinien

2. Weise jeder (weissen) Plakette mit Spinkonfiguration S einen Graphen G zu mit W :

$$P(S \rightarrow \{S, G\}) = \frac{W(S, G)}{W(S)} \quad (3.53)$$

3. Identifiziere alle Schleifen (jeweils verbundene Linien). Lege für jede Schleife die Spinrichtung neu fest (jeweils $W' \frac{1}{2}$).
4. Messe Observablen
5. zurück zu 2.

Dies entspricht dem Swendsen-Wang-Algorithmus.

Alternative: Wolld-artige Modifikation

- 3'. Wähle einen Spin und flippe die assoziierte Schleife

Offensichtlich erfüllen beide Varianten das detaillierte Gleichgewicht.

Die Gewichte $W(S, G)$ sind i.A. nicht eindeutig festgelegt. Es empfiehlt sich, die gefrorenen Graphen zu minimieren, da sonst die Cluster zu groß werden (und im Fall eines einzigen Clusters effektiv nichts passiert).

Speziell für das Heisenberg-Modell (XXZ-Modell mit $J_x = J_z$) können wir auf den gefrorenen Graphen ganz verzichten und auch den Diagonalgraphen eliminieren:

$$e^{\Delta\tau J/4} \cosh(\Delta\tau J/2) \equiv W_1 = W_{1,1} + W_{1,2}$$

$$e^{\Delta\tau J/4} \sinh(\Delta\tau J/2) \equiv W_2 = W_{2,2}$$

$$e^{-\Delta\tau J/4} \equiv W_3 = W_{3,1}$$

Dabei folgt aus der Gleichverteilungsbedingung:

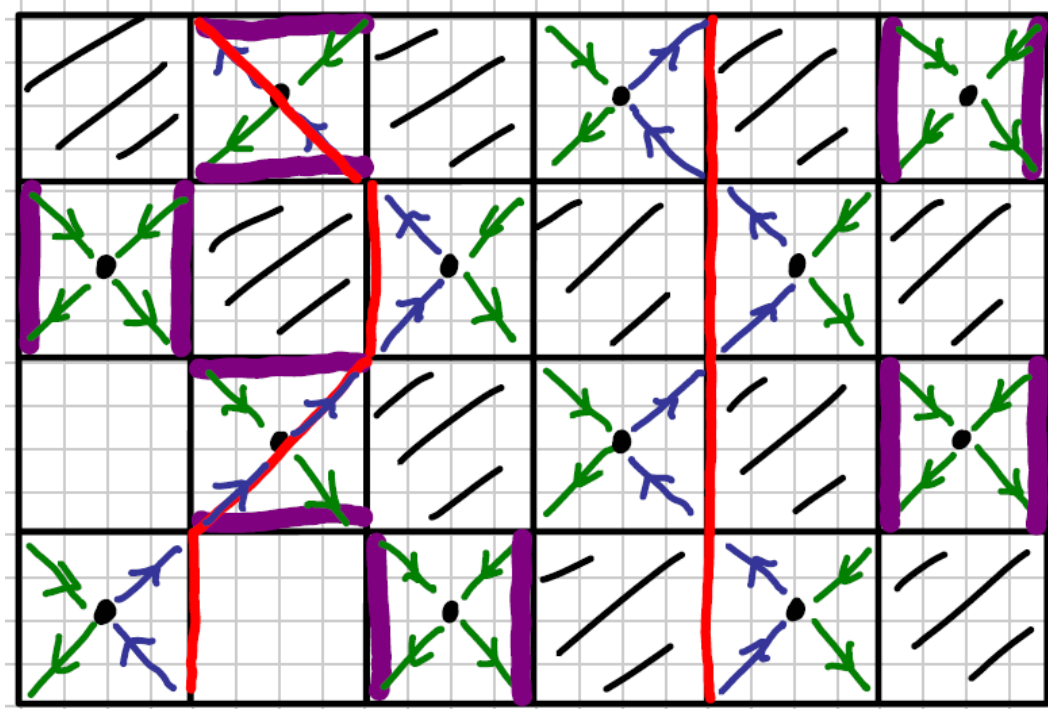
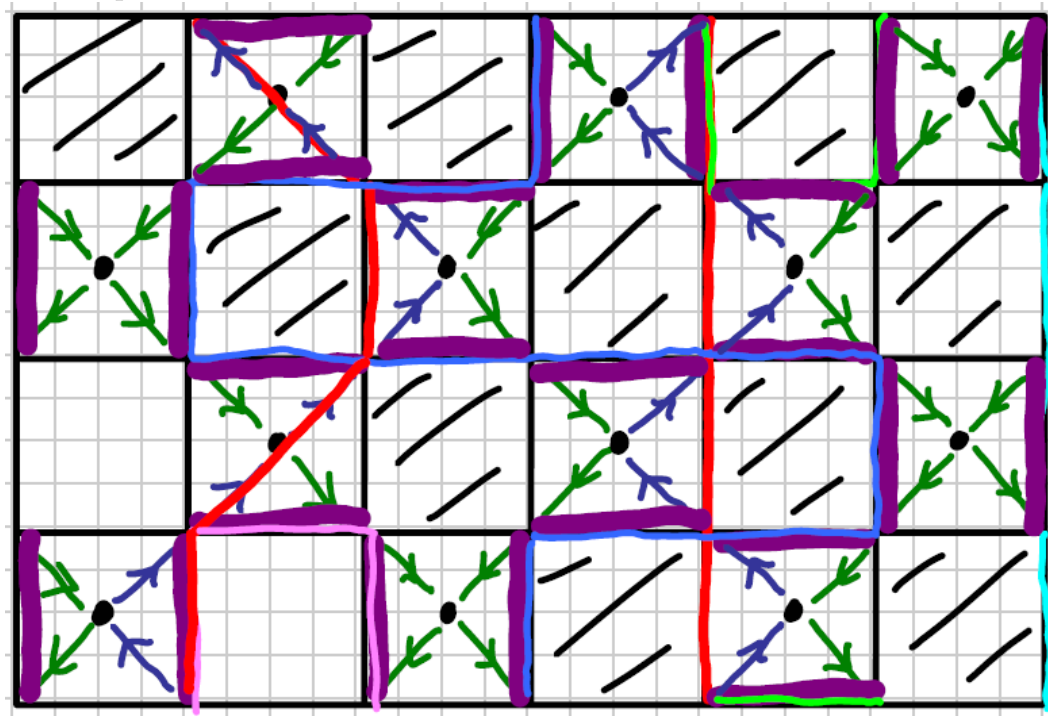
$$\begin{aligned} W_{1,1} &\stackrel{!}{=} W_{3,1} = W_3 \\ W_{1,2} &\stackrel{!}{=} W_{2,2} = W_2 \end{aligned} \quad (3.54)$$

Dies ist konsistent, da

$$\begin{aligned} W_{1,1} + W_{1,2} &= W_3 + W_2 \\ &= e^{\Delta\tau J/4} (e^{-\Delta\tau J/2} + \sinh(\Delta\tau J/2)) \\ &= e^{\Delta\tau J/4} \frac{2e^{-\Delta\tau J/2} + e^{\Delta\tau J/2} - e^{-\Delta\tau J/2}}{2} \\ &= e^{\Delta\tau J/4} \cosh\left(\frac{\Delta\tau J}{2}\right) = W_1 \end{aligned} \quad (3.55)$$

Abschließend noch ein Beispiel:

1. erzwungene Graphen

2. alle Graphen mit $P[\square \rightarrow \blacksquare] = \tanh(4\tau_z^2)$ 

3. Update hier: größter Cluser umgeklappt

