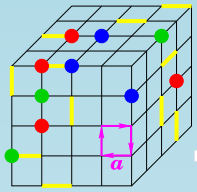
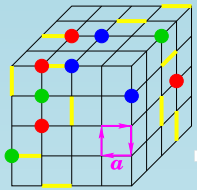


Numerical Simulations

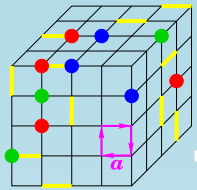


Is it theory or experiment?

Wikipedia: *Simulation is the imitation of some real thing, state of affairs, or process. The act of simulating something generally entails representing certain key characteristics or behaviours of a selected physical or abstract system.*

Numerical Simulation

- instead of the real experiments we cannot/do not want to perform (airplanes, nuclear war, evolution)

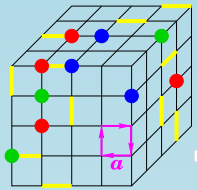


Is it theory or experiment?

Wikipedia: *Simulation is the imitation of some real thing, state of affairs, or process. The act of simulating something generally entails representing certain key characteristics or behaviours of a selected physical or abstract system.*

Numerical Simulation

- instead of the real experiments we cannot/do not want to perform (airplanes, nuclear war, evolution)
- problems with no analytic solution; complex systems, non-linearity, critical phenomena (independence of details: universality)

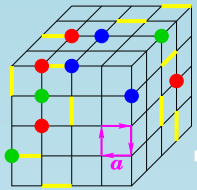


Is it theory or experiment?

Wikipedia: *Simulation is the imitation of some real thing, state of affairs, or process. The act of simulating something generally entails representing certain key characteristics or behaviours of a selected physical or abstract system.*

Numerical Simulation

- instead of the real experiments we cannot/do not want to perform (airplanes, nuclear war, evolution)
- problems with no analytic solution; complex systems, non-linearity, critical phenomena (independence of details: universality)
- e.g. cellular automaton: local evolution rules $\Rightarrow$ parallel processing. Note: dynamics may also be introduced (not *physical*)

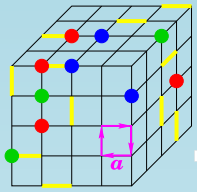


Is it theory or experiment?

Wikipedia: *Simulation is the imitation of some real thing, state of affairs, or process. The act of simulating something generally entails representing certain key characteristics or behaviours of a selected physical or abstract system.*

Numerical Simulation

- instead of the real experiments we cannot/do not want to perform (airplanes, nuclear war, evolution)
- problems with no analytic solution; complex systems, non-linearity, critical phenomena (independence of details: universality)
- e.g. cellular automaton: local evolution rules $\Rightarrow$ parallel processing. Note: dynamics may also be introduced (not *physical*)
- theoretical treatment, with experimental aspects (data, errors, measurements in time)

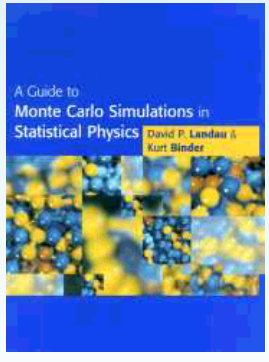


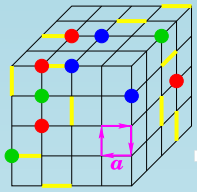
Monte Carlo Method

Stochastic systems may be simulated on a computer using pseudo-random numbers provided by some random-number generator, which employs an algebraic prescription to produce a (deterministic) sequence of numbers with the desired distribution, starting from a given seed.

> *Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin.*
>
> John Von Neumann (1951)

Reference: A Guide to Monte Carlo Simulations in Statistical Physics, Landau & Binder (Cambridge, 2000)





Monte Carlo Method: Crash Course

In Statistical Mechanics, the probability of a configuration ψ for a physical system in equilibrium at temperature T is given (in the canonical ensemble) in terms of the system's Hamiltonian $\mathcal{H}(\psi)$ by the Boltzmann distribution

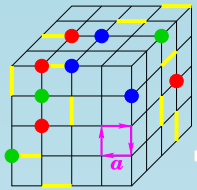
$$P(\psi) = \frac{e^{-\beta\mathcal{H}(\psi)}}{Z}; \qquad Z = \int d\psi\, e^{-\beta\mathcal{H}(\psi)}; \qquad \beta = 1/KT$$

Thermodynamic averages, e.g. $E = <\mathcal{H}(\psi)>$, given by

$$<A> = \int d\psi\, A(\psi)\, P(\psi)$$

(analogous to expectation values in Euclidean field theory)

This is a very complicated (high-dimensional) integral to compute!



Monte Carlo Method: Crash Course

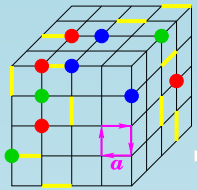
Think of an integral as a sum of (equally distributed) random variables

$$I = \int_0^1 f(x)\,dx \quad \rightarrow \quad \frac{\sum_i f(x_i)}{N}$$

with x_i uniformly distributed in [0,1]. Actually, for finite N $\overline{I} \equiv \sum_i f(x_i)/N$ is also a random variable, converging to its mean value I with an error proportional to $1/\sqrt{N}$ (central limit theorem)

$$\sigma_{\overline{I}}^2 = \frac{\sigma_f^2}{N} = \frac{< f^2 > - < f >^2}{N}$$

Exercise: derive the above expression



Monte Carlo Method: Crash Course

Think of an integral as a sum of (equally distributed) random variables

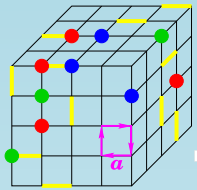
$$I = \int_0^1 f(x)\,dx \quad \rightarrow \quad \frac{\sum_i f(x_i)}{N}$$

with x_i uniformly distributed in [0,1]. Actually, for finite N $\overline{I} \equiv \sum_i f(x_i)/N$ is also a random variable, converging to its mean value I with an error proportional to $1/\sqrt{N}$ (central limit theorem)

$$\sigma_{\overline{I}}^2 = \frac{\sigma_f^2}{N} = \frac{< f^2 > - < f >^2}{N}$$

Exercise: derive the above expression

Typical deterministic methods for integrals have errors $\mathcal{O}(N^{-2})$ (trapezoidal rule) or $\sim \mathcal{O}(N^{-4})$ (Simpson's rule) for the 1d case



Monte Carlo Method: Crash Course

Think of an integral as a sum of (equally distributed) random variables

$$I = \int_0^1 f(x)\,dx \quad \rightarrow \quad \frac{\sum_i f(x_i)}{N}$$

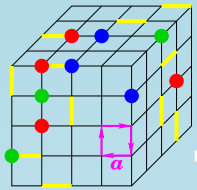
with x_i uniformly distributed in [0,1]. Actually, for finite N
$\overline{I} \equiv \sum_i f(x_i)/N$ is also a random variable, converging to its mean value I with an error proportional to $1/\sqrt{N}$ (central limit theorem)

$$\sigma^2_{\overline{I}} = \frac{\sigma^2_f}{N} = \frac{< f^2 > - < f >^2}{N}$$

Exercise: derive the above expression

Typical deterministic methods for integrals have errors $\mathcal{O}(N^{-2})$ (trapezoidal rule) or $\sim \mathcal{O}(N^{-4})$ (Simpson's rule) for the 1d case

$\Longrightarrow$ Monte Carlo method $\mathcal{O}(N^{-1/2})$ is not so great... (!?)



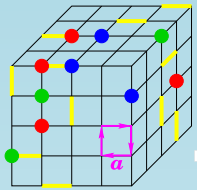
Monte Carlo Method: Crash Course

For d-dimensional integral: error $N^{-2/d}$ (trapezoidal rule) or $N^{-4/d}$ (Simpson) $\Rightarrow$ Monte Carlo attractive for d ≥ 8...

Typically $d \sim 10^3$ (simple statistical mechanical models, e.g. 3d Ising model with 10 sites per direction)

$\Rightarrow$ time to sum up 2^{1000} terms on a 1 Tflops machine:

$$t \; = \; 10^{288} s \; = \; 10^{270} \times \text{age of the universe}$$



Monte Carlo Method: Crash Course

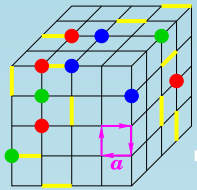
For d-dimensional integral: error $N^{-2/d}$ (trapezoidal rule) or $N^{-4/d}$ (Simpson) $\Rightarrow$ Monte Carlo attractive for d ≥ 8...

Typically $d \sim 10^3$ (simple statistical mechanical models, e.g. 3d Ising model with 10 sites per direction)

$\Rightarrow$ time to sum up 2^{1000} terms on a 1 Tflops machine:

$$t \; = \; 10^{288} s \; = \; 10^{270} \times \text{age of the universe}$$

$\Rightarrow$ Monte Carlo is not the best choice, it is the only one!



Monte Carlo Method: Crash Course

For d-dimensional integral: error $N^{-2/d}$ (trapezoidal rule) or $N^{-4/d}$ (Simpson) $\Rightarrow$ Monte Carlo attractive for d ≥ 8...

Typically $d \sim 10^3$ (simple statistical mechanical models, e.g. 3d Ising model with 10 sites per direction)
$\Rightarrow$ time to sum up 2^{1000} terms on a 1 Tflops machine:

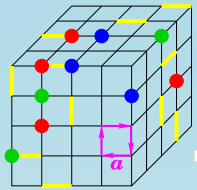
$$t \;=\; 10^{288} s \;=\; 10^{270} \times \text{age of the universe}$$

$\Rightarrow$ Monte Carlo is not the best choice, it is the only one!

Note: computing

$$I \;=\; \int_0^1 f(x)\, w(x)\, dx \;=\; \frac{\sum_i f(x_i)\, w(x_i)}{N}$$

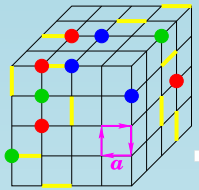
[where $\int_0^1 w(x)dx = 1$] with x_i uniformly distributed in [0,1] is very inefficient if $w(x)$ is concentrated. Take $\sum_i f(x_i)/N$ with x_i distributed according to $w(x_i) \Rightarrow$ importance sampling



Monte Carlo Method: Crash Course

There are good methods for sampling a random variable with a given distribution, starting from a uniform distribution in [0,1]. However, there is no hope of direct (i.e. independent) sampling of a joint distribution of many d.o.f. such as the Boltzmann distribution

$$< A >= \int A(x)\, w(x)\, dx\,, \qquad w(x) \;=\; \frac{e^{-\beta \mathcal{H}(x)}}{Z}$$

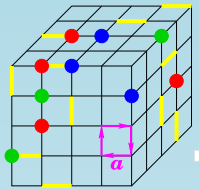


Monte Carlo Method: Crash Course

There are good methods for sampling a random variable with a given distribution, starting from a uniform distribution in [0,1]. However, there is no hope of direct (i.e. independent) sampling of a joint distribution of many d.o.f. such as the Boltzmann distribution

$$< A >= \int A(x)\, w(x)\, dx\,, \qquad w(x) \;=\; \frac{e^{-\beta \mathcal{H}(x)}}{Z}$$

Solution: dynamic Monte Carlo. Invent a time evolution such that the generated configurations will be distributed according to $w(x)$. This can be done for a Markov chain, i.e. a stochastic process X_0, X_1, ..., X_t such that $P(X_{t+1}|x_0 \ldots x_t) \;=\; P(X_{t+1}|x_t) \;\Rightarrow$ chain's history determined by $P(X_0)$ and the transition matrix p_{xy}; $\sum_y p_{xy} = 1$, all x



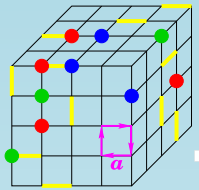
Monte Carlo Method: Crash Course

There are good methods for sampling a random variable with a given distribution, starting from a uniform distribution in [0,1]. However, there is no hope of direct (i.e. independent) sampling of a joint distribution of many d.o.f. such as the Boltzmann distribution

$$< A >= \int A(x)\, w(x)\, dx\,, \qquad w(x) = \frac{e^{-\beta \mathcal{H}(x)}}{Z}$$

Solution: dynamic Monte Carlo. Invent a time evolution such that the generated configurations will be distributed according to $w(x)$. This can be done for a Markov chain, i.e. a stochastic process X_0, X_1, ..., X_t such that $P(X_{t+1}|x_0 \ldots x_t) = P(X_{t+1}|x_t) \Rightarrow$ chain's history determined by $P(X_0)$ and the transition matrix p_{xy}; $\sum_y p_{xy} = 1$, all x

Note: if there is $w(x)$ s.t. $\sum_x w(x)\, p_{xy} = w(y)$ for all y, then the process converges to the stationary distribution $w(x)$ independently of $P(X_0)$



Monte Carlo Method: Crash Course

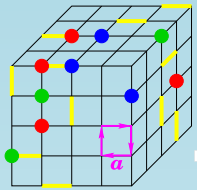
There are good methods for sampling a random variable with a given distribution, starting from a uniform distribution in [0,1]. However, there is no hope of direct (i.e. independent) sampling of a joint distribution of many d.o.f. such as the Boltzmann distribution

$$< A >= \int A(x)\, w(x)\, dx\,, \qquad w(x) = \frac{e^{-\beta \mathcal{H}(x)}}{Z}$$

Solution: dynamic Monte Carlo. Invent a time evolution such that the generated configurations will be distributed according to $w(x)$. This can be done for a Markov chain, i.e. a stochastic process X_0, X_1, ..., X_t such that $P(X_{t+1}|x_0 \ldots x_t) = P(X_{t+1}|x_t) \Rightarrow$ chain's history determined by $P(X_0)$ and the transition matrix p_{xy}; $\sum_y p_{xy} = 1$, all x

Note: if there is $w(x)$ s.t. $\sum_x w(x)\, p_{xy} = w(y)$ for all y, then the process converges to the stationary distribution $w(x)$ independently of $P(X_0)$

Our (inverse) problem: find p_{xy} having $w(x)$ as its stationary distribution



Monte Carlo Method: Crash Course

Conditions on the dynamics p_{xy} for the Markov chain to converge to the distribution $w(x)$ at large times

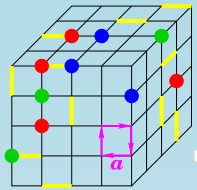
(A) Irreducibility (ergodicity): for all $x,\ y$ there is n s.t. $p_{xy}^{(n)} \neq 0$

(B) $w(x)$ is stationary: $\sum_x w(x)\, p_{xy} = w(y)$

Note: may also impose the sufficient condition

(B') Detailed balance: $w(x)\, p_{xy} = w(y)\, p_{yx}$

Exercise: show that (B') $\rightarrow$ (B)



Monte Carlo Method: Crash Course

Conditions on the dynamics p_{xy} for the Markov chain to converge to the distribution $w(x)$ at large times

(A) Irreducibility (ergodicity): for all $x,\ y$ there is n s.t. $p_{xy}^{(n)} \neq 0$

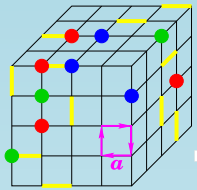
(B) $w(x)$ is stationary: $\sum_x w(x)\, p_{xy} = w(y)$

Note: may also impose the sufficient condition

(B') Detailed balance: $w(x)\, p_{xy} = w(y)\, p_{yx}$

Exercise: show that (B') $\rightarrow$ (B)

Generally the evolution of configurations (e.g. x or ψ) for the system is done by freezing the field variables at all points but one, which is then sampled by a local method. An iteration of the algorithm, i.e. one step of the Markov chain, is obtained by sweeping in this way over all sites



Monte Carlo Method: Crash Course

Conditions on the dynamics p_{xy} for the Markov chain to converge to the distribution $w(x)$ at large times

(A) Irreducibility (ergodicity): for all $x,\ y$ there is n s.t. $p_{xy}^{(n)} \neq 0$

(B) $w(x)$ is stationary: $\sum_x w(x)\, p_{xy} = w(y)$

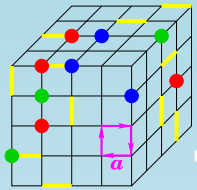
Note: may also impose the sufficient condition

(B') Detailed balance: $w(x)\, p_{xy} = w(y)\, p_{yx}$

Exercise: show that (B') $\rightarrow$ (B)

Generally the evolution of configurations (e.g. x or ψ) for the system is done by freezing the field variables at all points but one, which is then sampled by a local method. An iteration of the algorithm, i.e. one step of the Markov chain, is obtained by sweeping in this way over all sites

The heat-bath algorithm: exact sampling of the local (conditional) distribution; clearly a valid way to sample the joint distribution



Monte Carlo Method: Crash Course

The Metropolis algorithm: based on proposing and accepting/rejecting a step $x \to y$

- accept if $w(y)/w(x) \geq 1$
- otherwise accept with probability $w(y)/w(x)$

the probability of acceptance is $A_{xy} = \min\{1, w(y)/w(x\}$. Then consider the transition matrix $p_{xy} = T_{xy}\, A_{xy}$ (with general $T_{xy} = T_{yx}$)

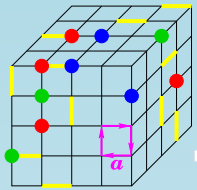
Exercise: show that the above choice satisfies detailed balance

For the Boltzmann distribution this means

$$w(x) = \frac{e^{-\beta E(x)}}{Z} \quad \Rightarrow \quad \frac{w(y)}{w(x)} = e^{-\beta \Delta E}\,; \quad \Delta E \equiv E(y) - E(x)$$

$\Rightarrow$ accept if $\Delta E \leq 0$; otherwise accept with probability $e^{-\beta \Delta E}$

Note: if proposed step is rejected, keep old value and move to a new site; when possible, choose T_{xy} such that acceptance is 50%



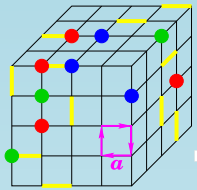
Monte Carlo Method: Crash Course

Program: follow the dynamics $X(t) = x_i$ and compute time averages

$$< A > = \int A(x)\, w(x)\, dx \; = \; \frac{\sum_i A(x_i)}{N}$$

which are expectation values in the desired distribution, i.e. the Boltzmann distribution. Note that the initial transient must be discarded. The resulting averages + errors are the output of our Monte Carlo simulation

Exercise: application to the Ising Model



Monte Carlo Method: Crash Course

Program: follow the dynamics $X(t) = x_i$ and compute time averages

$$< A > = \int A(x)\, w(x)\, dx \; = \; \frac{\sum_i A(x_i)}{N}$$

which are expectation values in the desired distribution, i.e. the Boltzmann distribution. Note that the initial transient must be discarded. The resulting averages + errors are the output of our Monte Carlo simulation

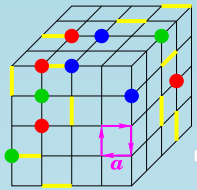
Exercise: application to the Ising Model

But... we have a problem: samples are not independent.
The program above is subject to systematic effects.

The time correlation between different steps of the Markov chain is

$$C(k) \; = \; \frac{< A_i\, A_{i+k} > - < A_i >^2}{< A_i^2 > - < A_i >^2}$$

$\Rightarrow$ independent samples only after $C(k) \approx 0$; k = decorrelation time



Monte Carlo Method: Crash Course

(Monte Carlo) average of A: $$\overline{A} = \frac{1}{N}\sum_{i=1}^{N} A_i$$

Variance: $$\sigma^2_{\overline{A}} = \frac{\sigma^2_A}{N}\left[1 + 2\sum_{k=1}^{N-1} C(k)\right] = \frac{\sigma^2_A}{N}(2\,\tau)$$

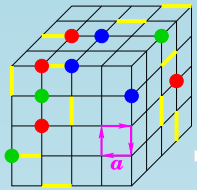
where the temporal correlation $C(k)$ was given above and τ is the auto-correlation time for observable A.

Consider $C(k) = e^{-k/\tau}$, τ large (but $\tau \ll N$)

$$1 + 2\sum_{k=1}^{N-1} C(k) \approx 2\sum_{k=0}^{\infty} e^{-k/\tau} - 1$$

$$\approx 2\,\tau \int_0^{\infty} e^{-u} du - 1 \approx 2\tau$$

We therefore define $\tau \equiv \frac{1}{2} + \sum_{k=1}^{N-1} C(k)$



Monte Carlo Method: Summary

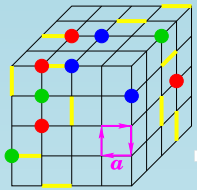
Integral becomes sum of random variables

$$\int f(x)\, d\mu\ , \quad d\mu = \frac{e^{-\beta \mathcal{H}(x)}}{Z}\, dx \quad \Rightarrow \quad \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

where x_i have statistical distribution μ

- Static Monte Carlo: independent sampling (error $\sim 1/\sqrt{N}$)
- Dynamic Monte Carlo: Simulation of a Markov chain with equilibrium distribution μ (error $\sim \sqrt{\tau/N}$). Autocorrelation time τ related to critical slowing-down. Note: similar to experimental methods, but temporal dynamics was artificially introduced

Errors: either consider only effectively independent samples (via temporal correlation analysis) and error is given by standard deviation, jack-knife, bootstrap or consider all samples and error is estimated taking correlations into account: binning method, self-consistent windowing method

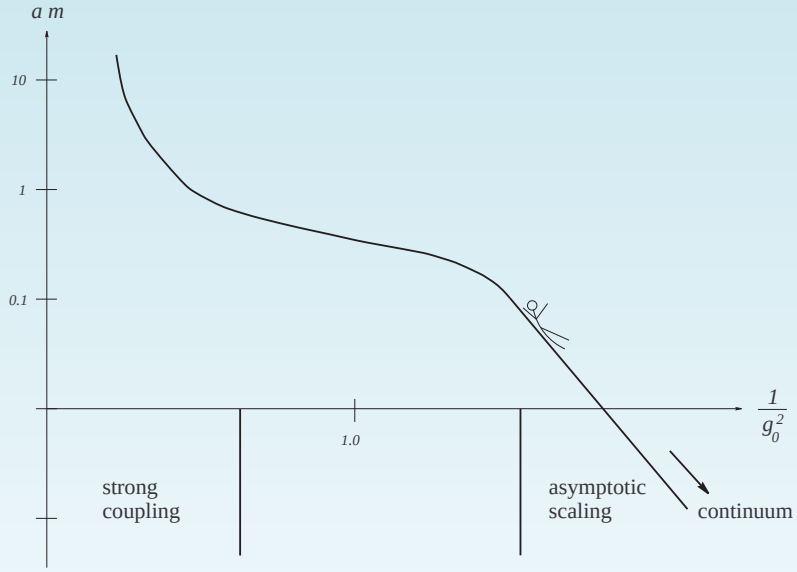


The Continuum Limit: how do I get there!?

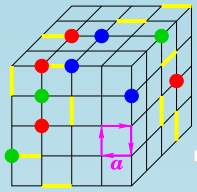
Physics is obtained after 3 limits:

1) The Thermodynamic Limit ($V = N^d \to \infty$): need $N \to \infty$ to keep physical lengths $L = aN$ fixed. Need $N > \xi_{latt}$, while $\xi_{latt}(a)$ diverges!

2) The Continuum Limit ($a \to 0$): correlation length $\leftrightarrow$ mass^{-1} from renormalization group:
$\log(\xi_{latt}) = \log(1/ma) \sim 1/g_0^2 \sim \beta$
thus continuum limit given by $g_0 \to 0$, $\beta \to \infty$ and $\xi_{latt} \sim e^\beta$ (asymptotic scaling), i.e. $\xi = 1/m \sim a\, e^\beta \Rightarrow$ eliminate e^β computing mass ratios (scaling law) or fix a using an experimental input (renormalization)



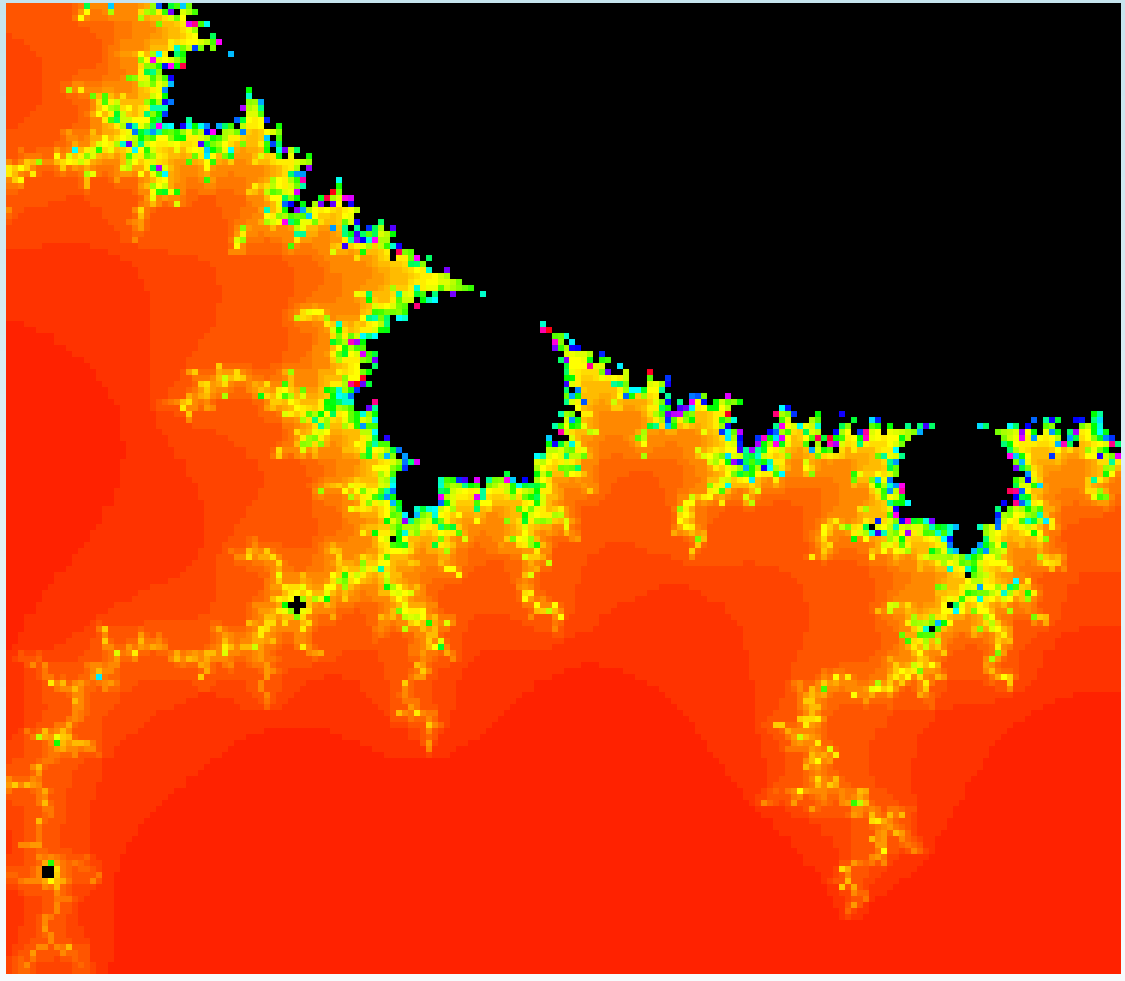
3) The Chiral Limit (small m_q): fit results to chiral perturbation theory predictions and extrapolate to physical masses

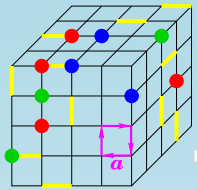


Critical point

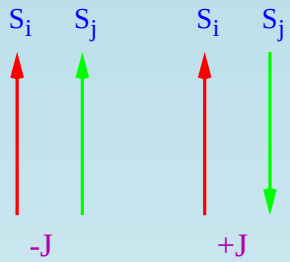
Lattices have to be huge because $\xi \to \infty$ at the critical point...

at the same time, there is scale invariance, universality, ...





Application to the Ising Model (I)

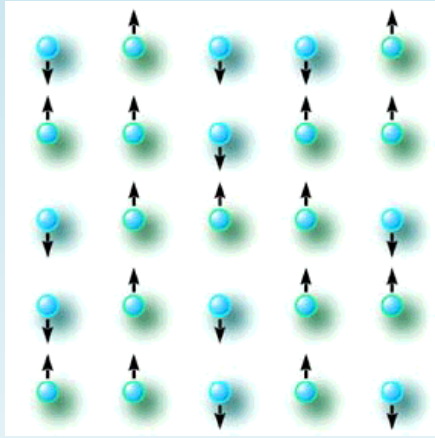


two-state “spins”, that prefer to be aligned

$$\mathcal{H}(S) = -J \sum_{<i,j>} S_i S_j - H \sum_i S_i$$

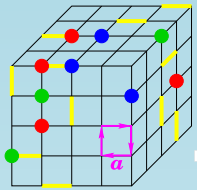
Observables of interest

- Energy: $E = < \mathcal{H}(S) >$
- Specific Heat: $C_V = \partial E / \partial T$
- Magnetization: $M = < \sum_i S_i >$
- Suscetibility: $\chi = \partial M / \partial H$



Exercise: write simple programs to simulate the 2d Ising model using the Metropolis and heat-bath methods. Compare the efficiency of both methods, plot the dependence on the magnetic field and temperature.

Note: partly ready programs, more details and examples are available at http://lattice.if.sc.usp.br/CADSC03/curso_MC.html



Application to the Ising Model (II)

Metropolis method for the Ising model: sweep over the lattice, at each site propose to flip the spin, i.e. $S_i \to -S_i$. Acceptance probability

$$\frac{e^{-\beta\, E(y)}}{e^{-\beta\, E(x)}} = \frac{e^{+\beta J\, S_i \sum_{j\ n.n.\ i} S_j}}{e^{-\beta J\, S_i \sum_{j\ n.n.\ i} S_j}} = e^{2\,\beta J\, S_i\, h_i}$$

an iteration consists of a complete "sweep" over the lattice. At the end of the iteration compute $A(S)$ for the generated configuration, and restart the process of generating configurations

Heat-bath method for the Ising model: exact sampling of the conditional probability at site i. Sweep over the lattice, at each site pick a new value for S_i independently of the old one, keeping all other spins fixed. Unnormalized probability $P(S_i) = e^{\beta J S_i \sum_{j\ n.n.\ i} S_j} \times const.$ Thus

- $P(S_i = +1) = e^{\beta J h_i}/(e^{\beta J h_i} + e^{-\beta J h_i}) \equiv p$
- $P(S_i = -1) = 1 - p$