Ising model (cont'd)

PHYS 250 (Autumn 2019) – Lecture 6

David Miller

Department of Physics and the Enrico Fermi Institute
University of Chicago

October 17, 2019

Outline

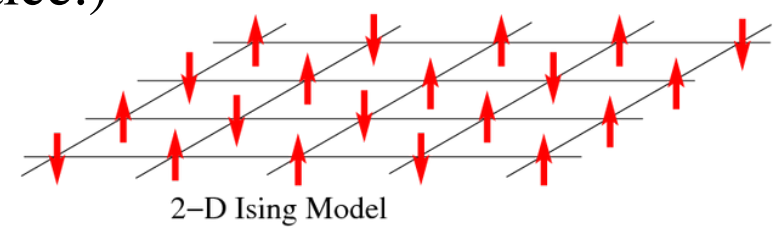
1. *Computational Ising model*
   - Reminders
   - Practical implications and relationships
   - Complications and numerical and computational obstacles
   - Monte Carlo Methods
   - The origins of Monte Carlo
   - Assumptions of the Metropolis algorithm
   - Metropolis-Hastings algorithm

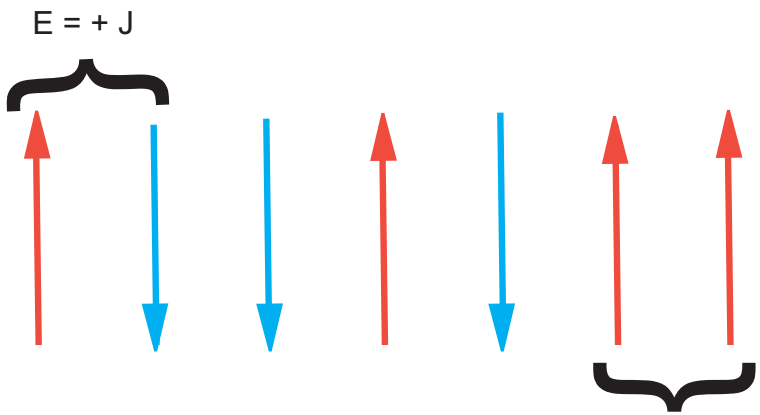
Quantum mechanics of the Ising model

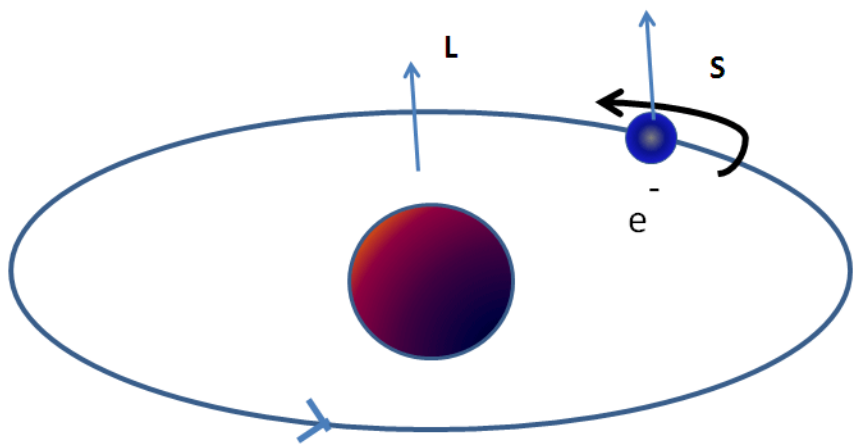
The Hamiltonian for a system of N spins s_i in free space, which may take on values of $s_i = \pm 1$ is given by the interaction of the spins

$$\mathcal{H} = -J \sum_{\langle ij \rangle}^{N} s_i s_j, \qquad (1)$$

where the sum $\langle ij \rangle$ is over all pairs of nearest-neighbor sites, J is the **coupling** between these neighboring sites. (For example, there are four neighbors per site on the square lattice.)







Magnetism and spins (I)

As we've discussed in E&M, magnetism is caused by charged particles **"spinning"** in closed orbits or about their axes

- For elementary particles, of course, we mean "spinning" in the quantum mechanical sense, not the rotational Newtonian sense

So atoms may have **both an L (orbit) and an S (spin) contribution** (capital letters for "operators" for those of you in the know) to the magnetic properties. For the Ising model, we're concerned with just a single contribution which we will refer to as s as on the previous slide.

In addition to the contribution to the Hamiltonian (and thus the energy) from the spin-spin interactions, there is the possibility that we apply an external magnetic field, H:

$$\mathcal{H} = -J \sum_{\langle ij \rangle}^{N} s_i s_j - H \sum_i s_i, \tag{2}$$

Note that the "magnetization" is given by $M = \sum_i s_i$.

Magnetism and spins (II)

Let's discuss our Hamiltonian

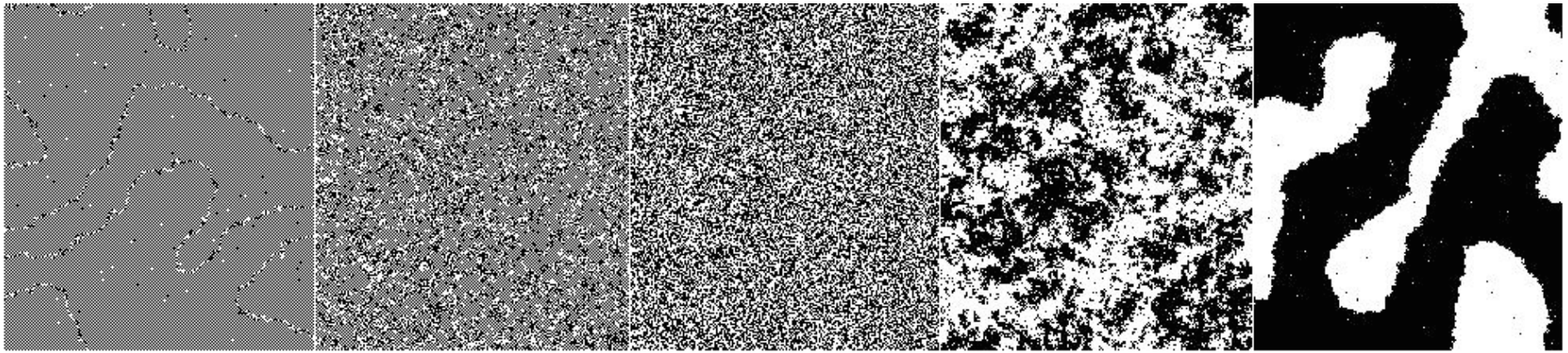
$$\mathcal{H} = -J \sum_{\langle ij \rangle}^{N} s_i s_j - H \sum_i s_i, \tag{3}$$

- **If $J > 0$: ferromagnetic**
  - Energy is minimized if the spins point in the same direction
- **If $J < 0$: antiferromagnetic**
  - Energy is minimized if spins locally point in the opposite direction

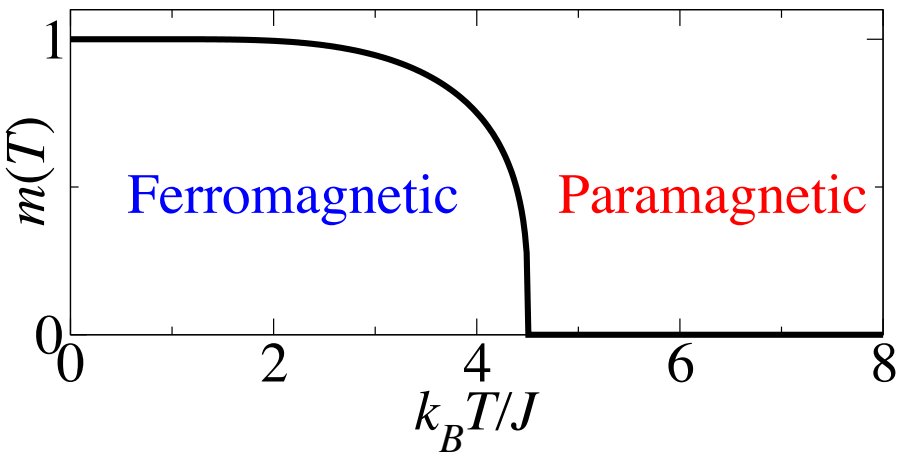
Important observations which we will study

- At low temperatures, the **spins will organize themselves** to either mostly point up or mostly point down, forming a ferromagnetic phase.
- The spins will tend to align (in a 2D square lattice) in a **checkerboard antiferromagnetic phase** at low temperatures.
- At high temperatures, independent of the sign of J, we expect entropy to dominate; the spins will fluctuate wildly in a **paramagnetic phase** and the magnetization per spin $m(T) = M(T)/N$ is zero

Emergent properties of the Ising model



Amazing features result from the **ensemble behavior** of the lattice sites of a 2D Ising model. Phase transitions, criticality, and other phenomena are all part of what we will study computationally with this model.



A few quantum mechanics notes

The Ising model parameters are rescaled from the microscopic ones. The Ising spin $s_i = \pm 1$ represents twice the z-component of a spin-1/2 atom in a crystal, $\sigma_i^z = s_i/2$.

The Ising interactions between spins, $Js_is_j = 4J\sigma_i^z\sigma_j^z$, is thus shifted by a factor of four from the $z-z$ coupling between spins.

The coupling of the spin to the external magnetic field is microscopically $g\mu_B H \cdot \sigma_i^z$, where g is the gyromagnetic ratio for the spin (close to two for the electron) and $\mu_B = \dfrac{e\hbar}{2m_e}$ is the Bohr magneton.

Hence, for those of you who've taken quantum mechanics, the Ising external field is rescaled from the physical one by $g\mu_B/2$.

Broader context for the Ising model

Lattice models are a big industry within statistical mechanics. Many systems and their computational evaluation and study take place on a lattice:

- **Critical phenomena and phase transitions**
- **lattice QCD and quantum field theories**
- **quantum magnetism and models for high-temperature superconductors**
- **phase diagrams for alloys**
- **the behavior of systems with dirt or disorder**
- **non-equilibrium systems exhibiting avalanches and crackling noise**

And many more!

(now do you see why I spent time showing you `np.mgrid`*?!)*

Computational complexity and simulation

The 1D model was solved analytically in 1924.

The 2D model was not solved analytically until Lars Onsager did it in 1944.

- Onsager solved is using the **transfer-matrix method**, where you essentially break up the partition function, Z into many pieces and decompose the interactions in a matrix formulation and then do an eigenvector decomposition and solve it that way.

However, this only really works for a small system. If there are many states (i.e. many sites) it can easily become analytically intractable even if solvable in principle.

Instead, we will approach this using our Monte Carlo based approach!

Reminder of the Ising model

In Lecture 5 we began discussing the Ising model and the general principles that we seek to understand about this deceptively **"simple"** model for a set of spin states $\mathbb{S} = \{s_i\}$, with a Hamiltonian, $\mathcal{H}$, for a given interaction coupling, J, and magnetic field, H, given by:

$$\mathcal{H} = -J \sum_{\langle ij \rangle}^{N} s_i s_j - H \sum_i s_i, \tag{4}$$

Given a randomized set of N spins, the Ising model has 2^N states $\mathbb{S} = \{s_i\}$ in $d = 2$ dimensions. What happens to the system when we let it evolve according to the laws of statistical mechanics? Does the average spin remain random? How does the application of an external field affect its evolution? Can we calculate a specific heat for this system (the temperature change required to raise the system's energy by a given amount)?

The statistical mechanics of the system

The energy E_μ of a particular microstate μ of this (discrete) system is given by the operator $\mathcal{H}$ acting on that microstate, or specific $\mathbb{S} = \mathbb{S}_\mu$. We assume that although the system is in an **equilibrium state** (i.e. the energy of a particular element is proportional to the temperature, T), it is a dynamic one in which each element's energy fluctuates as it exchanges energy with its environment. The probability for the ensemble $\mathbb{S}$ to have energy E_μ is

$$P(\mathbb{S}_\mu) = \frac{e^{-\beta E_\mu}}{\sum_\mu e^{-\beta E_\mu}}, \tag{5}$$

And the mean and variance of the energy $\langle E \rangle$ (or *any* observable) is given by

$$\langle E \rangle = \sum_\mu E_\mu P(\mathbb{S}_\mu) = \frac{1}{Z} \sum_\mu E_\mu e^{-\beta E_\mu} \tag{6}$$

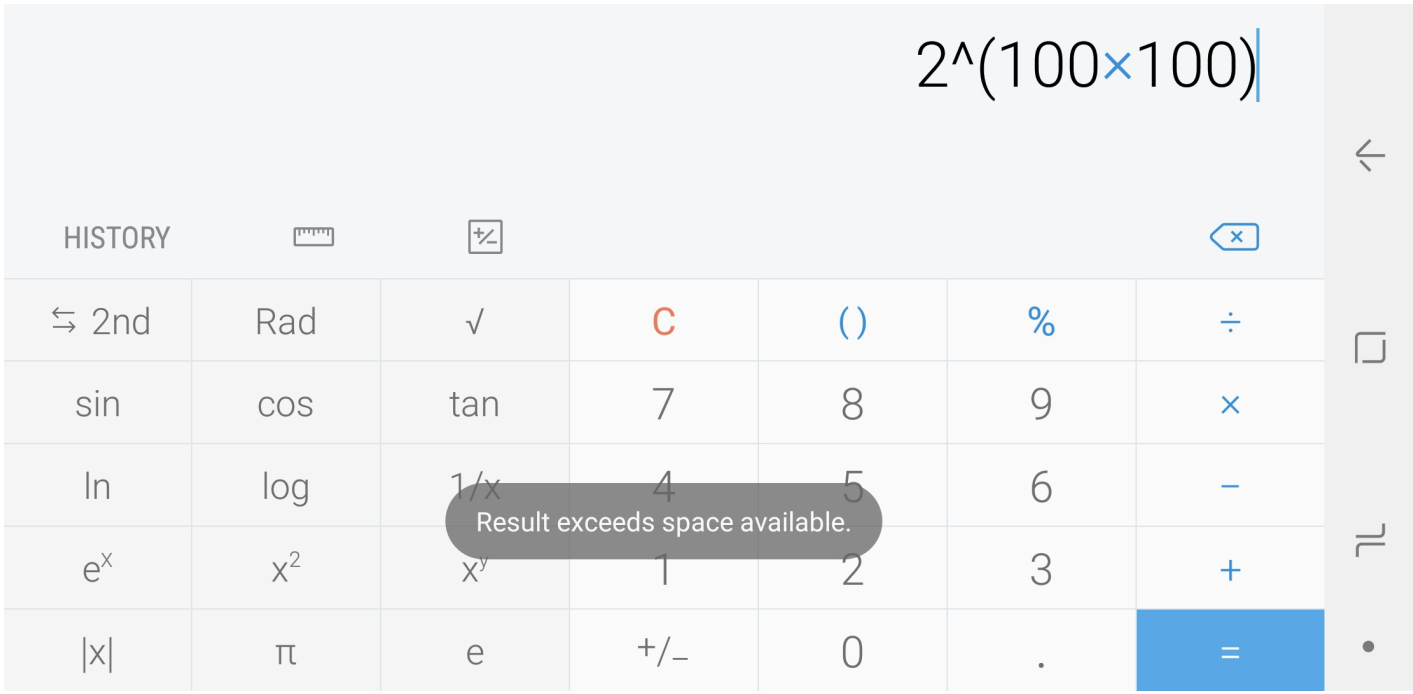
$$Var(E) = \langle (E - \langle E \rangle^2)^2 \rangle = \sum_\mu E_\mu^2 P(\mathbb{S}_\mu) - \left(\sum_\mu E_\mu P(\mathbb{S}_\mu) \right)^2 \tag{7}$$

Computational obstacles (big ones)

If we were to brute-force calculate the sum in the previous slide, we would have to perform $4 \times N \times 2^N$ calculations, which quickly becomes $\sim \infty$. For example, for just a 5×5 lattice, we have $2^{25} = 33,554,432$ states. For a 100×100 lattice, well...

Computational obstacles (big ones)

If we were to brute-force calculate the sum in the previous slide, we would have to perform $4 \times N \times 2^N$ calculations, which quickly becomes $\sim \infty$. For example, for just a 5×5 lattice, we have $2^{25} = 33,554,432$ states. For a 100×100 lattice, well...



Computational obstacles (big ones)

If we were to brute-force calculate the sum in the previous slide, we would have to perform $4 \times N \times 2^N$ calculations, which quickly becomes $\sim \infty$. For example, for just a 5×5 lattice, we have $2^{25} = 33,554,432$ states. For a 100×100 lattice, well...

We can save summing over half of these due to up/down symmetry, which means that for every state there is another one in which every spin is simply flipped upside down, with exactly the same energy (for $H = 0$). So, we can simplify the calculation by just taking one out of every pair of such states, for a total of $16,777,216$ states, and summing up the corresponding terms in the partition function.

We could save further time by using other symmetries too. For example, a square lattice also has a reflection symmetry and a four-fold rotational symmetry (symmetry group C_4), meaning that the states actually group into sets of 16 states (including the up/down symmetry pairs), all of which have the same energy. But $\infty/16$ still ain't great...

Monte Carlo methods

There is essentially only **one known numerical method** for calculating the partition function of a model such as the Ising model on a large lattice, and that method is **Monte Carlo simulation**.

- If we are clever enough, we can obtain a **relatively good estimate** by only performing a **subset** of the calculations, $\{\mu\}$ instead of all μ
- One way to be clever is to only sample the distribution that we are attempting to model, $P(\mathbb{S}_\mu)$ in regions where it is important.
- To put it another way, we want to perform a **weighted sampling**

$$\langle E \rangle = \frac{\sum_{\{\mu\}} E_\mu e^{-\beta E_\mu} W_\mu^{-1}}{\sum_{\{\mu\}} e^{-\beta E_\mu} W_\mu^{-1}} \tag{8}$$

$$\approx \frac{\sum_{\{\mu\}} E_\mu}{\sum_{\{\mu\}} 1} = \frac{1}{N'} \sum_{\{\mu\}} E_\mu \tag{9}$$

where N' is the number of terms in the subset $\{\mu\}$ and $W_\mu = e^{-\beta E_\mu}/Z$.

Recap of the Monte Carlo methods

We discussed (briefly) two approaches to the Monte Carlo implementation

- **Heat bath method:** always flip a spin according to the Boltzmann factor based on the energy difference (i.e. an energy difference of $\Delta E = 0$ has a 50% probability of flipping)
- **Metropolis method:** always flip a spin if the energy difference is negative (i.e. the new spin configuration results in a lower energy), but only flip the spin according to the Boltzmann factor if the result is an energy increase.

In both cases, there are two underlying assumptions which were first proposed in 1949 by Nicolas Metropolis (hence the eponymous algorithm).

The beginning of the Monte Carlo

The idea of Monte Carlo calculation is a lot older than the computer. Under the older name of "statistical sampling" the method has a history stretching back well into the last century, when numerical calculations were performed by hand using pencil and paper and perhaps a slide-rule.

The name "Monte Carlo" is relatively recent: it was coined by Nicolas Metropolis in 1949 in a paper with Stanislav Ulam

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 247 SEPTEMBER 1949 *Volume 44*

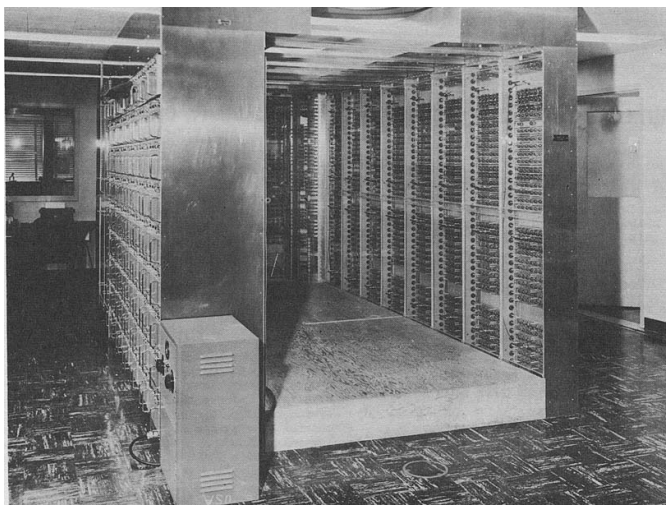
THE MONTE CARLO METHOD

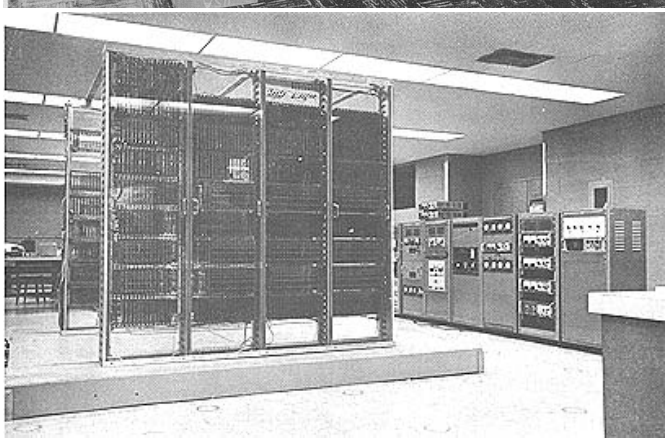
NICHOLAS METROPOLIS AND S. ULAM
Los Alamos Laboratory

We shall present here the motivation and a general description of a method dealing with a class of problems in mathematical physics. The method is, essentially, a statistical approach to the study of differential equations, or more generally, of integro-differential equations that occur in various branches of the natural sciences.

Nicolas Metropolis

Nicholas Constantine Metropolis was born in 1915, in Chicago (where all the coolest people were born). In 1936 he received his bachelor's degree, and in 1941, his doctorate, both from UChicago, and both in experimental physics. While here, Metropolis worked at the Met Lab as an assistant to Fermi. The MANIAC was Metropolis' computer at Los Alamos: **"Mathematical Analyzer, Numerical Integrator, and Computer"**. In 1957, Metropolis returned to UChicago and founded the Institute for Computer Research. Here, he oversaw the construction MANIACIII machine.





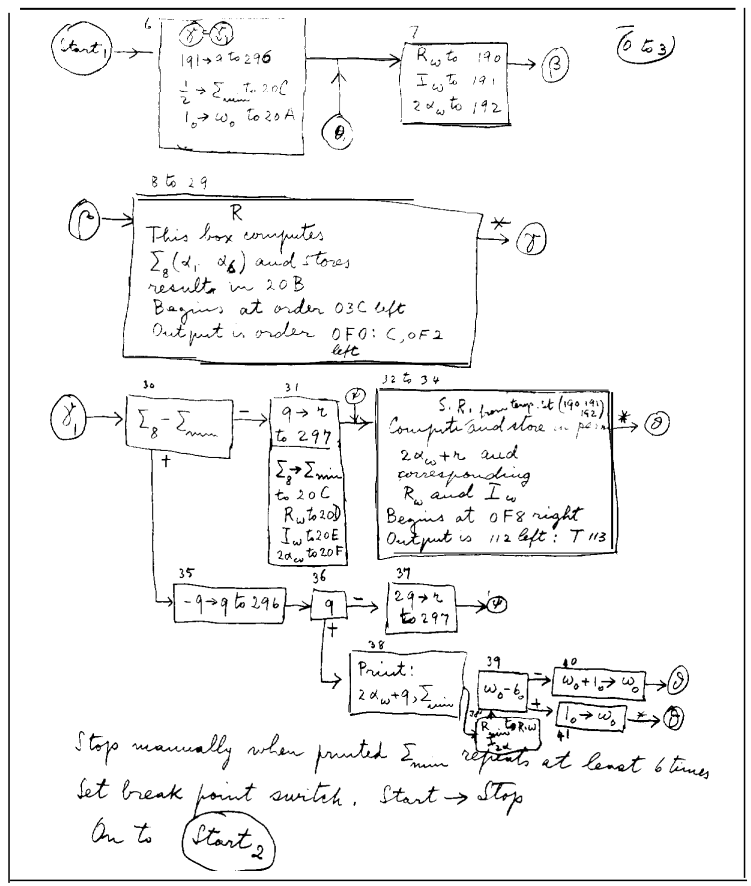


Fig. 4. A subprogram written by Fermi for calculating phase shifts by finding a minimum chi-squared in a fit to the data.

The concepts behind Monte Carlo

The concept that Metropolis and Ulam sketched out is exactly what we have been discussing:

> Here again one might try to perform a finite number of "experiments" and obtain a class or sample of possible genealogies. These experiments will of course be performed not with any physical apparatus, but theoretically. If we assume that the probability of each possible event is given, we can then play a great number of games of chance, with chances corresponding to the assumed probability distributions. In this fashion one can study empirically the asymptotic properties of powers of matrices with positive coefficients, interpreted as transition probabilities.

but is also has a very important embedded concept from deep within statistical physics (we will come back to this in a moment):

> It follows from simple application of ergodic theorems that the estimate should be, *with great probability*, valid within a few per cent.

Computational physics was born

In the same paper, Metropolis and Ulam made a compelling suggestion: with a combination of good ideas (from Fermi) and computers, this could be very powerful!

> We want now to point out that modern computing machines are extremely well suited to perform the procedures described. In practice,
>
> :
>
> given equation into an equivalent one, possessing the form of a diffusion equation with possible multiplication of the particles involved. For example as suggested by Fermi, the time-independent Schrödinger equation
>
> $$\Delta\psi(x, y, z) = (E - V)\psi(x, y, z)$$
>
> could be studied as follows. Re-introduce time dependence by considering
>
> $$u(x, y, z, t) = \psi(x, y, z)e^{-Et}$$
>
> u will obey the equation
>
> $$\frac{\partial u}{\partial t} = \Delta u - Vu.$$

Deploying the tools

In a follow-up paper, now with some of the Manhattan project's heaviest hitters (from Chicago, of course) they put those ideas into action

THE JOURNAL OF CHEMICAL PHYSICS VOLUME 21, NUMBER 6 JUNE, 1953

Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

AND

EDWARD TELLER,* *Department of Physics, University of Chicago, Chicago, Illinois*

(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.

The Metropolis algorithm

It is in this paper that the proposal to use a Boltzmann factor for attempting the **trials** a given change in configuration

a configuration of very low weight. So the method we employ is actually a modified Monte Carlo scheme, where, instead of choosing configurations randomly, then weighting them with $\exp(-E/kT)$, we choose configurations with a probability $\exp(-E/kT)$ and weight them evenly.

VI. CONCLUSION

The method of Monte Carlo integrations over configuration space seems to be a feasible approach to statistical mechanical problems which are as yet not analytically soluble. At least for a single-phase system a sample of several hundred particles seems sufficient.

The Metropolis algorithm's assumptions: ergodicity

Over a sufficiently long period of time, **all accessible microstates will be sampled**. A corollary is that the time spent by a system in some region of the phase space of microstates with the same energy is proportional to the volume of this region.

The implication for our computational approach is that in principle, it should be **possible for our Markov process to reach any state of the system from any other state**, if we run it for long enough.

The condition of ergodicity tells us that we are allowed to make some of the transition probabilities of our Markov process zero, but that there must be at least one path of non-zero transition probabilities between any two states that we pick.

In practice, most Monte Carlo algorithms set almost all of the transition probabilities to zero, and we **must be careful that in so doing we do not create an algorithm which violates ergodicity**.

The Metropolis algorithm's assumptions: detailed balance

The detailed balance equation then tells us that the transition probabilities should satisfy

$$\frac{P(\mu \to \nu)}{P(\nu \to \mu)} = \frac{P(\mathbb{S}_\nu)}{P(\mathbb{S}_\mu)} = e^{-\beta(E_\nu - E_\mu)} = e^{-\beta \Delta E} \tag{10}$$

This equation and the requirement that the sum of all probabilities is unity are the basic constraints on our choice of transition probabilities. If we satisfy these, as well as the condition of ergodicity, then the equilibrium distribution of states in our Markov process will be the Boltzmann distribution. It is also useful to break this down a bit further:

$$P(\mu \to \nu) = g(\mu \to \nu) A(\mu \to \nu) \tag{11}$$

$$\frac{P(\mu \to \nu)}{P(\nu \to \mu)} = \frac{g(\mu \to \nu) A(\mu \to \nu)}{g(\nu \to \mu) A(\nu \to \mu)} \tag{12}$$

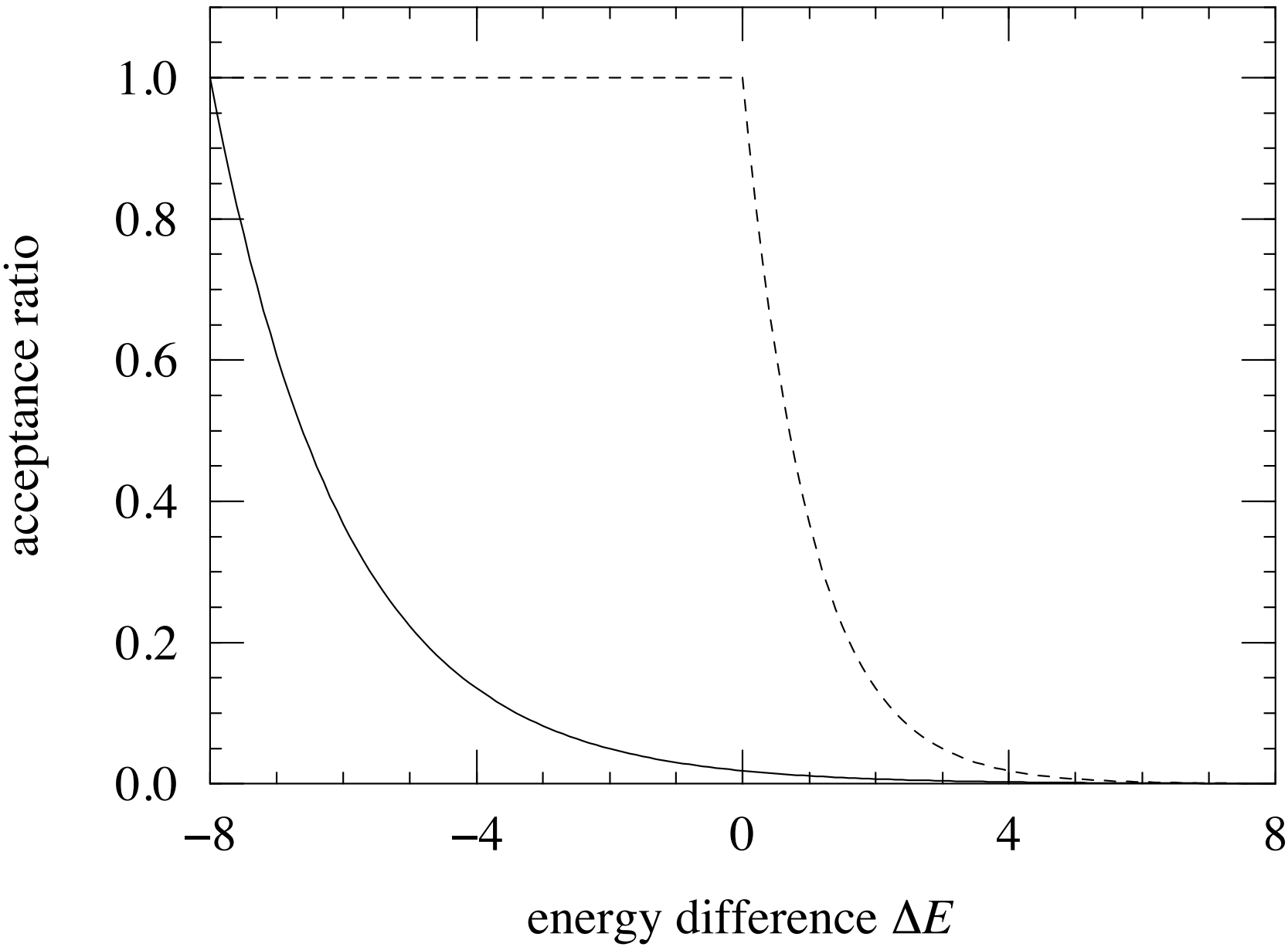
where g is a **selection probability** and A is the **acceptance ratio**

Selection probability

- $g(\mu \to \nu)$ **is the selection probability**
  - this is the probability, given an initial state μ, that our algorithm will generate a new target state ν.
- $A(\mu \to \nu)$ **is the acceptance ratio (sometimes also called the "acceptance probability")**
  - The acceptance ratio says that if we start off in a state μ and our algorithm generates a new state ν from it, we should accept that state and change our system to the new state ν a fraction of the time $A(\mu \to \nu)$.
  - The rest of the time we should just stay in the state μ.
  - If the acceptance ratios for our moves are low, then the algorithm will on most time steps simply stay in the state.

In the Metropolis algorithm the selection probabilities $g(\mu \to \nu)$ for each of the possible states ν are all chosen to be equal.

Acceptance probability



Metropolis Monte Carlo

"Pseudocode" for the Metropolis algorithm:

1. Choose initial configuration
2. For all spins:
   1. Trial flip ($+1 \rightarrow -1$ or vice versa)
   2. Compute change in energy
   3. If $\delta = e^{-\Delta E/k_B T} > r$ (where r is a uniform deviate) flip the spin