Revisiting Two Influential Papers in the History of Computational Physics

Andrew Siegel
Argonne National Laboratory

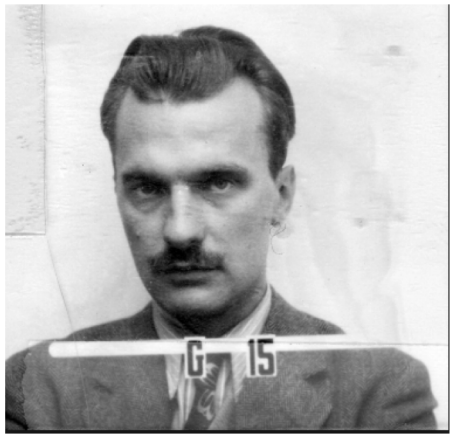
Choice of works

- These papers are not typically discussed together but have some key issues in common
  - required use of computers
  - were in some sense enabled by limitations in computer resources
  - were result of solving specific science problem
  - Their impact took time and wasn't initially obvious

Markov Chain Monte Carlo

- ***Equation of State Calculations on Fast Computing Machines,*** 1953, J. Chem Phys.

  Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller



Metropolis

Rosenbluth

Rosenbluth

Teller

Teller

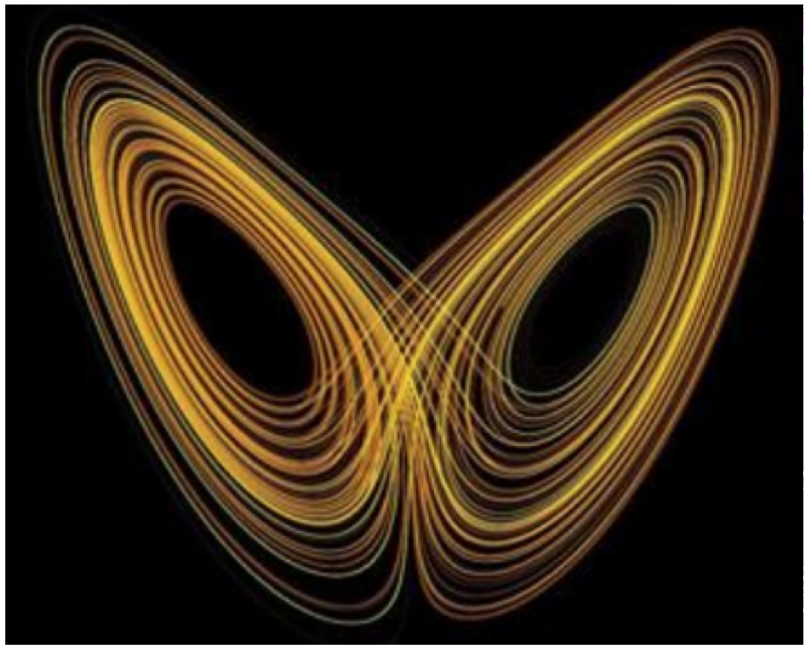
Before there was Bogdan …

Edward Teller

Deterministic Chaos

- ***Deterministic Nonperiodic Flow***, 1963, Journal of the Atmospheric Sciences

  Edward N. Lorenz



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.

I. INTRODUCTION

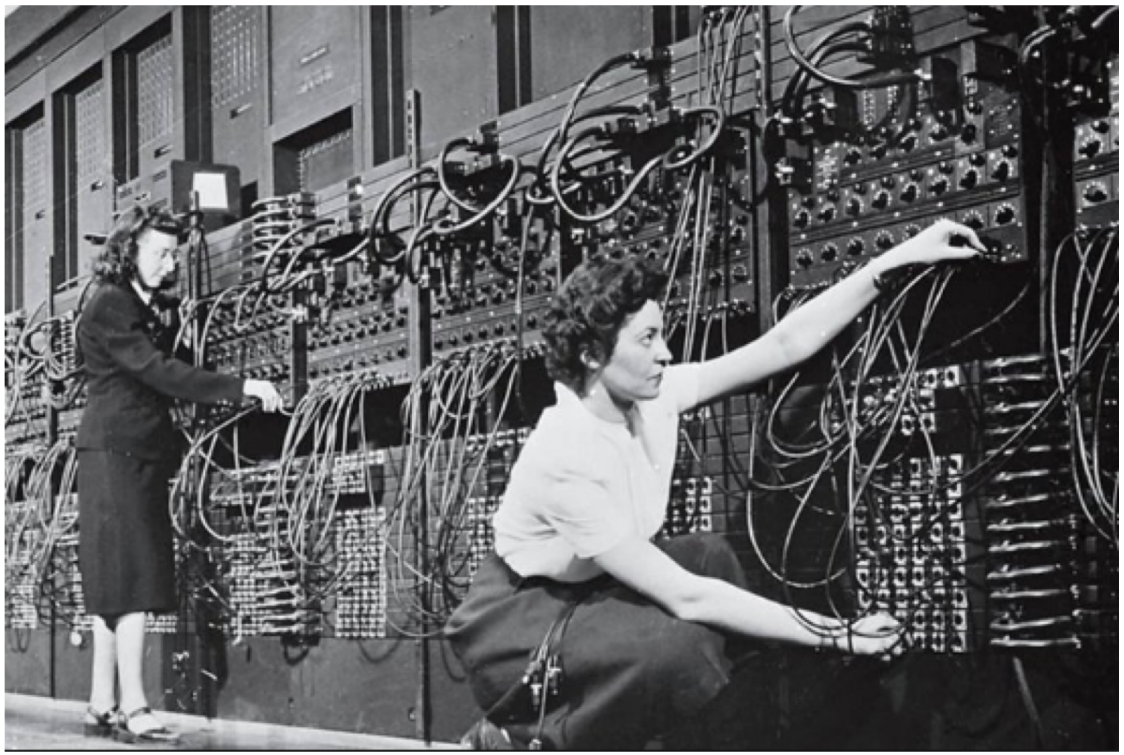
THE purpose of this paper is to describe a general method, suitable for fast electronic computing machines, of calculating the properties of any substance which may be considered as composed of interacting individual molecules. Classical statistics is assumed, only two-body forces are considered, and the potential field of a molecule is assumed spherically symmetric. These are the usual assumptions made in theories of liquids. Subject to the above assumptions, the method is not restricted to any range of temperature or density. This paper will also present results of a preliminary two-dimensional calculation for the rigid-sphere system.

II. THE GENERAL METHOD FOR AN ARBITRARY POTENTIAL BETWEEN THE PARTICLES

In order to reduce the problem to a feasible size for numerical work, we can, of course, consider only a finite number of particles. This number N may be as high as several hundred. Our system consists of a square† containing N particles. In order to minimize the surface effects we suppose the complete substance to be periodic, consisting of many such squares, each square containing N particles in the same configuration. Thus we define d_{AB}, the minimum distance between particles A and B, as the shortest distance between A and any of the particles B, of which there is one in each of the squares which comprise the complete substance. If we

ENIAC (1946)

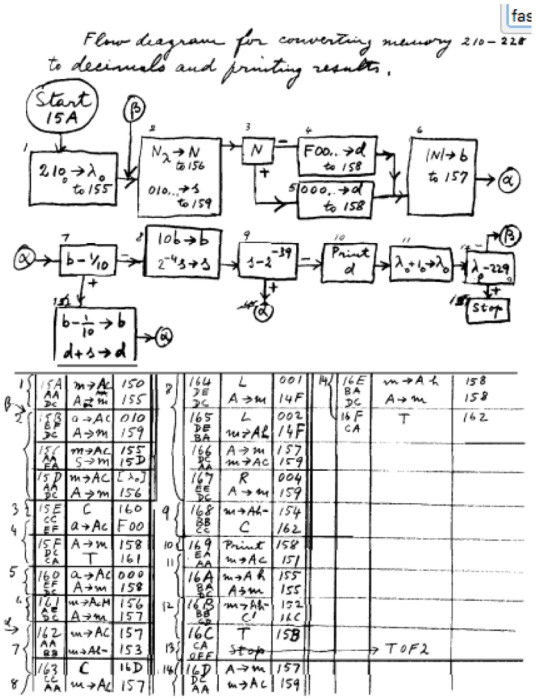
- Electronic Numerical Integrator and Computer
- Could not store programs
- 17,480 tubes operating at 100,000 pulses per second
- 50K instructions, 300 multiplications per second
- Intended to compute ballistic trajectories before war ended
- Used to study nuclear reactions

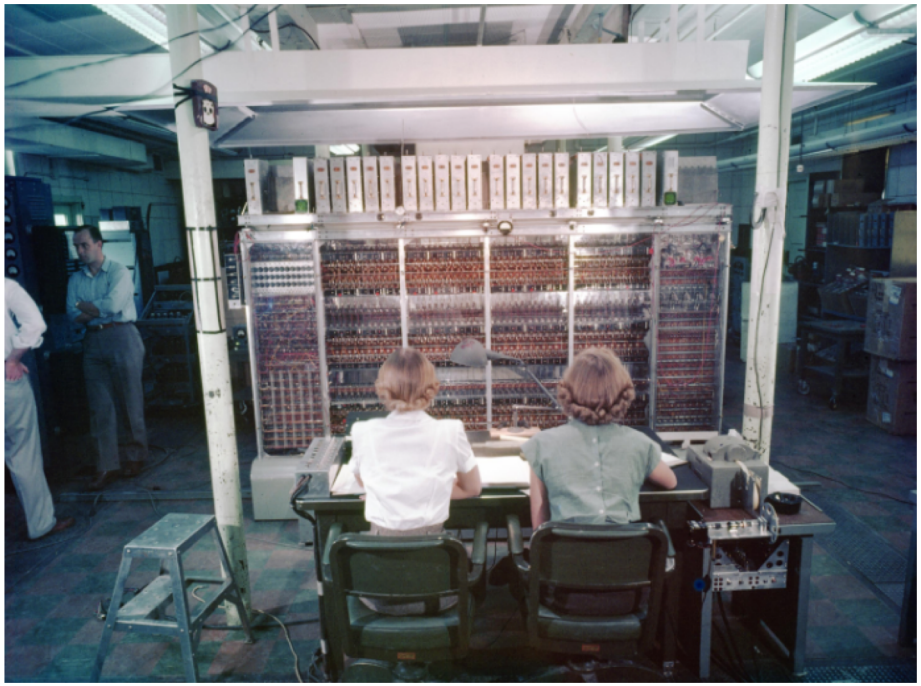



20th
CENTURY
FOX
Presents

MANIAC (1952)

- Based on Von Neumann Architecture
- Built at Los Alamos by team led by Metropolis
- Used to design hydrogen bomb
- Pioneering developing of Monte Carlo methods (Ulam)
- Intended to compute ballistic trajectories before war ended
- Used to study nuclear reactions





Metropolis et al driven by science goals – understanding phase transitions

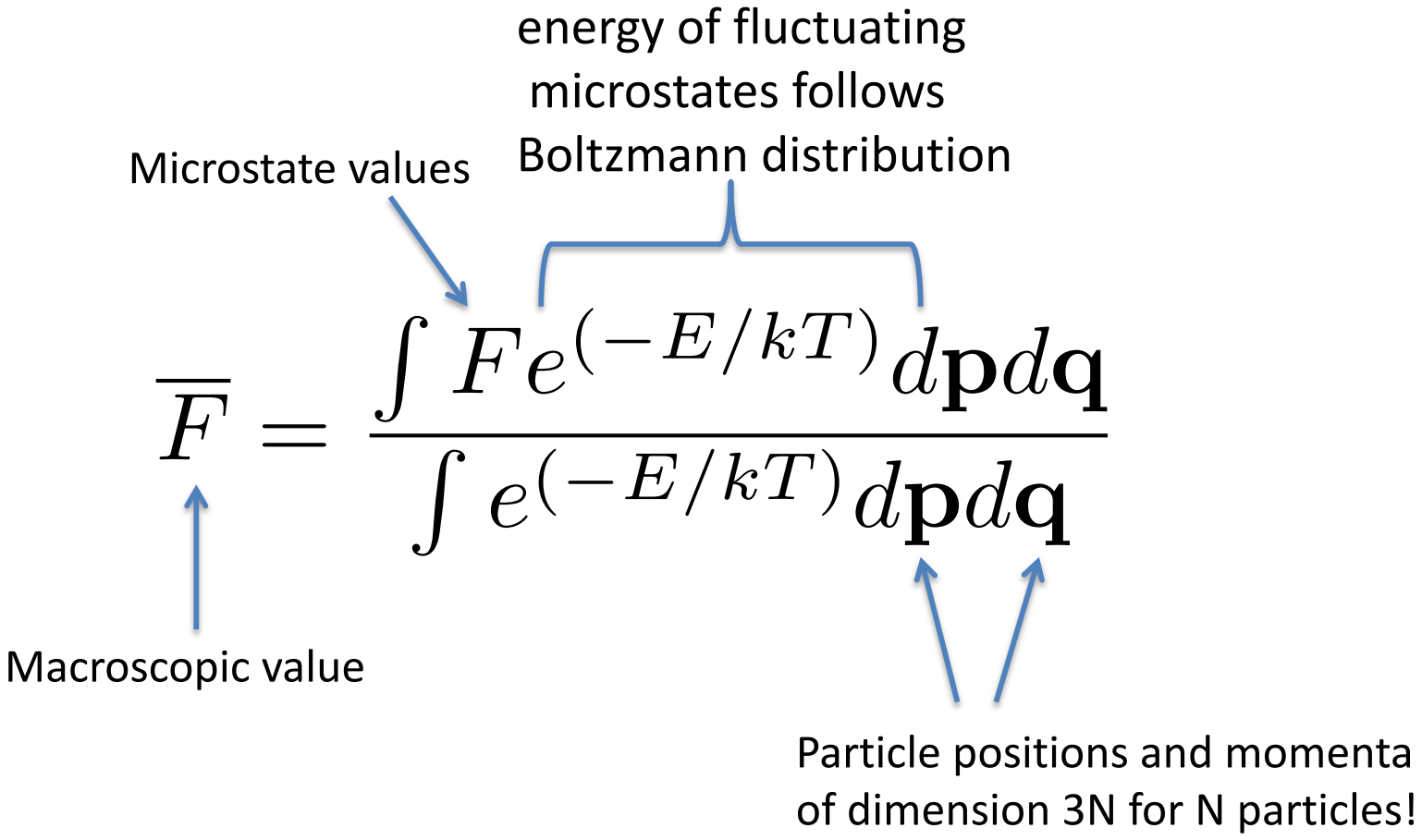
"The purpose of this paper is to describe a general method, suitable for fast electronic computing machines, of calculating the properties of any substance which may be considered as composed of interacting particles" , Metropolis et al. (1953*)*

Role of Teller

"Teller made the crucial suggestion, pointing out that statistical mechanical averages could be performed by ensemble averaging instead of time averaging. Time averaging required following the detailed kinematics of the interacting particles through the time integration of Newton's laws. Marshall comments that Metropolis's computers as the others of the timed was not yet big enough and fast enough to do this." Gubernatis (2005)

$$\overline{F} = \frac{\int F e^{(-E/kT)} d\mathbf{p} d\mathbf{q}}{\int e^{(-E/kT)} d\mathbf{p} d\mathbf{q}}$$

Computing Macroscopic Quantities



Estimating the integral

- For even small number of particles integral far too high dimensionality to discretize directly
- New approach then was Monte Carlo integration
- Pick a random configuration of particles, calculate energy of given configuration, and weight configuration by its probability

$$\overline{F} = \sum_i F_i \frac{e^{(E_i/kT)}}{\sum_j e^{(E_i/kT)}} = \sum_i F_i p_i$$

“Naïve” Monte Carlo Simulation

- Problem: each microstate was expensive on MANIAC, took several minutes to evaluate
- Most states sample were in very low probability regions, contributed little or nothing to integral.
- Couldn’t get good estimate of macroscopic quantities

"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." Metropolis et al (1953)

"... A simple way to do this sampling configurations with the Boltzmann weight , as emerged after discussions with Teller, would be to make a trial move: if it decreased the energy of the system, allow it; if it increased the energy, allow it with probability exp(− E/kT) as determined by a comparison with a random number. Each step, after an initial annealing period, is counted as a member of the ensemble, and the appropriate ensemble average of any quantity determined." Rosenbluth(2005)

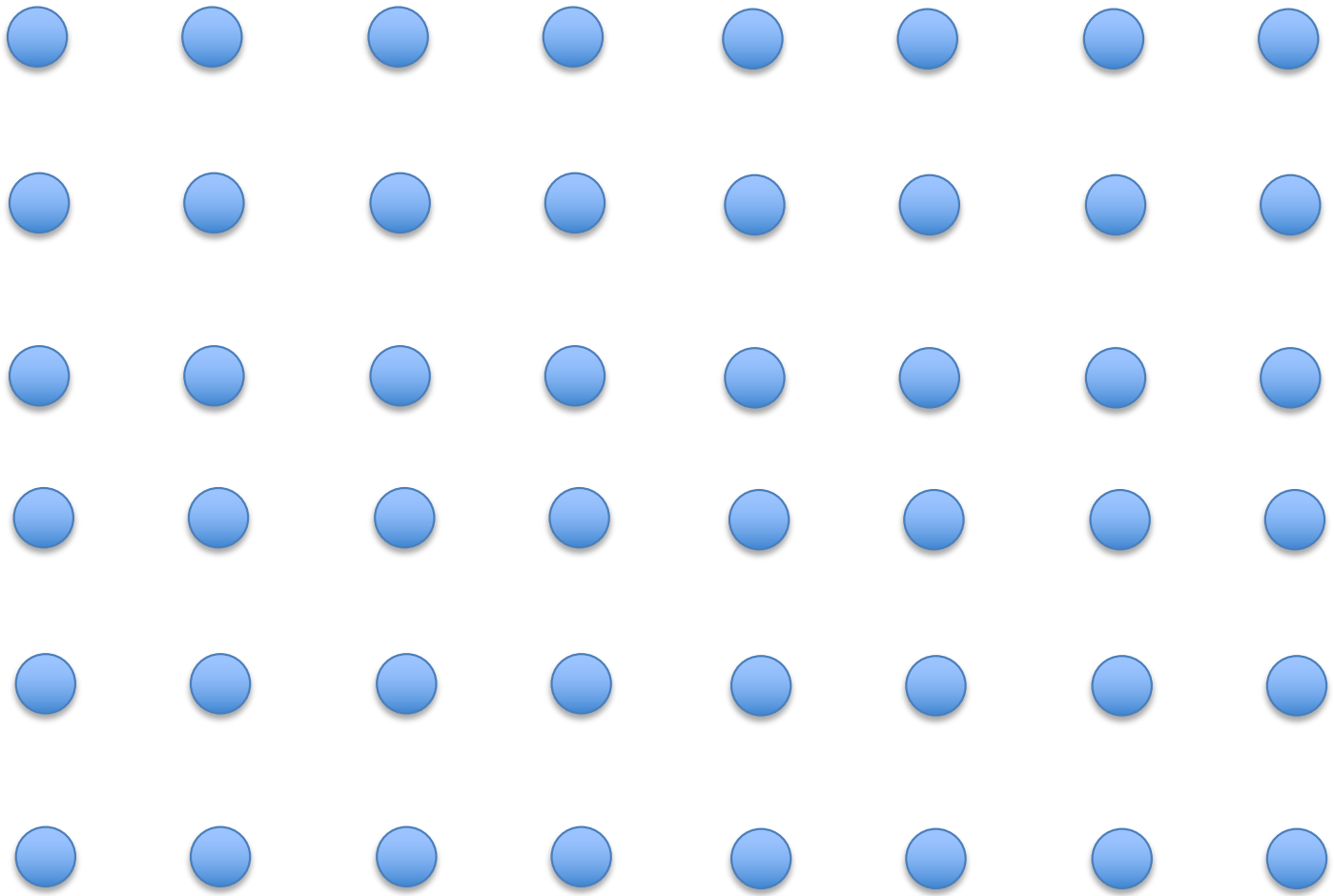
$$\overline{F} = \frac{1}{m} \sum F_m$$

Sampling from Boltzmann

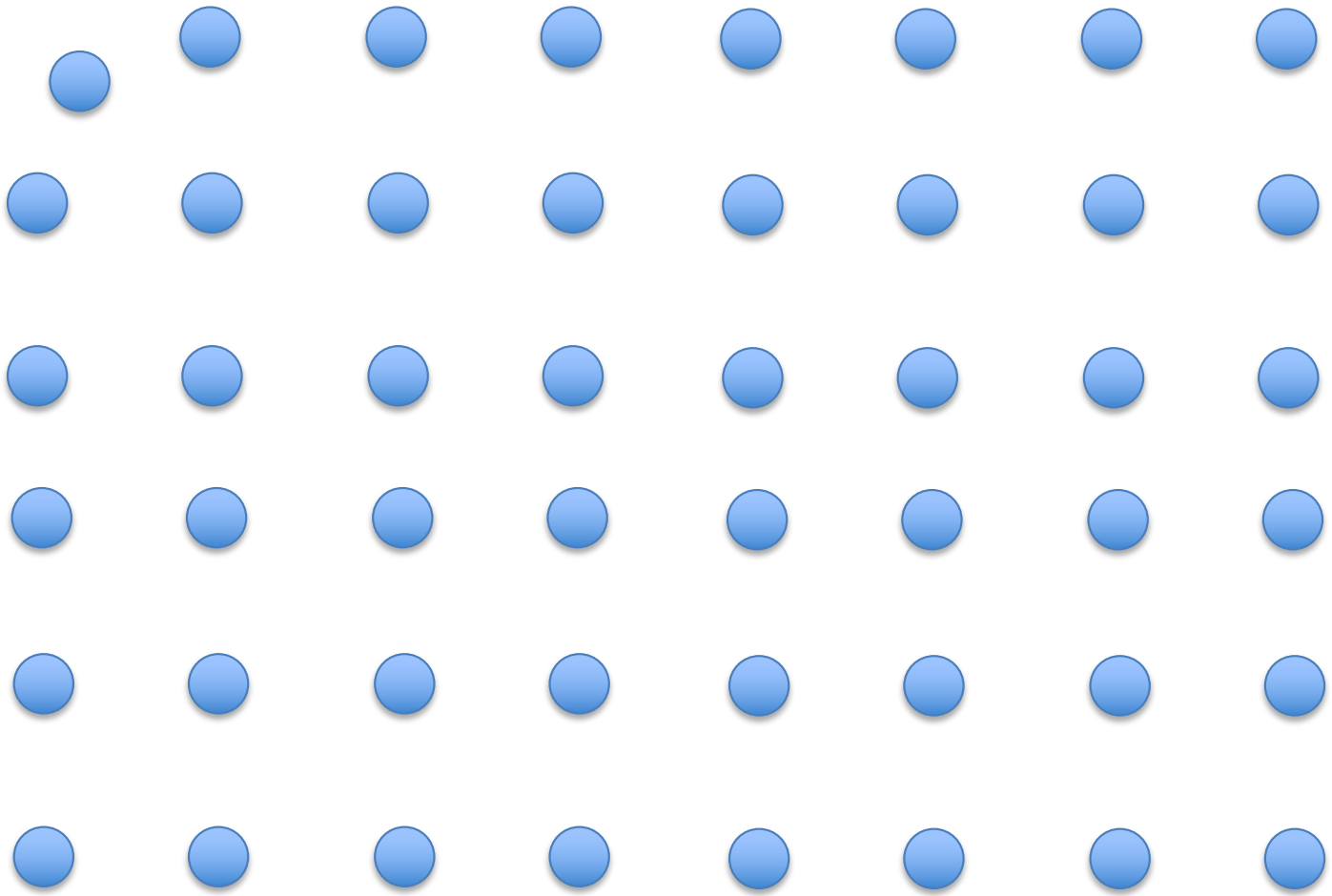
- Not as easy as it sounds
- Don’t know energy until configuration is selected
- Technique has to be simple/inexpensive to compute

Start with arbitrary configuration and calculate

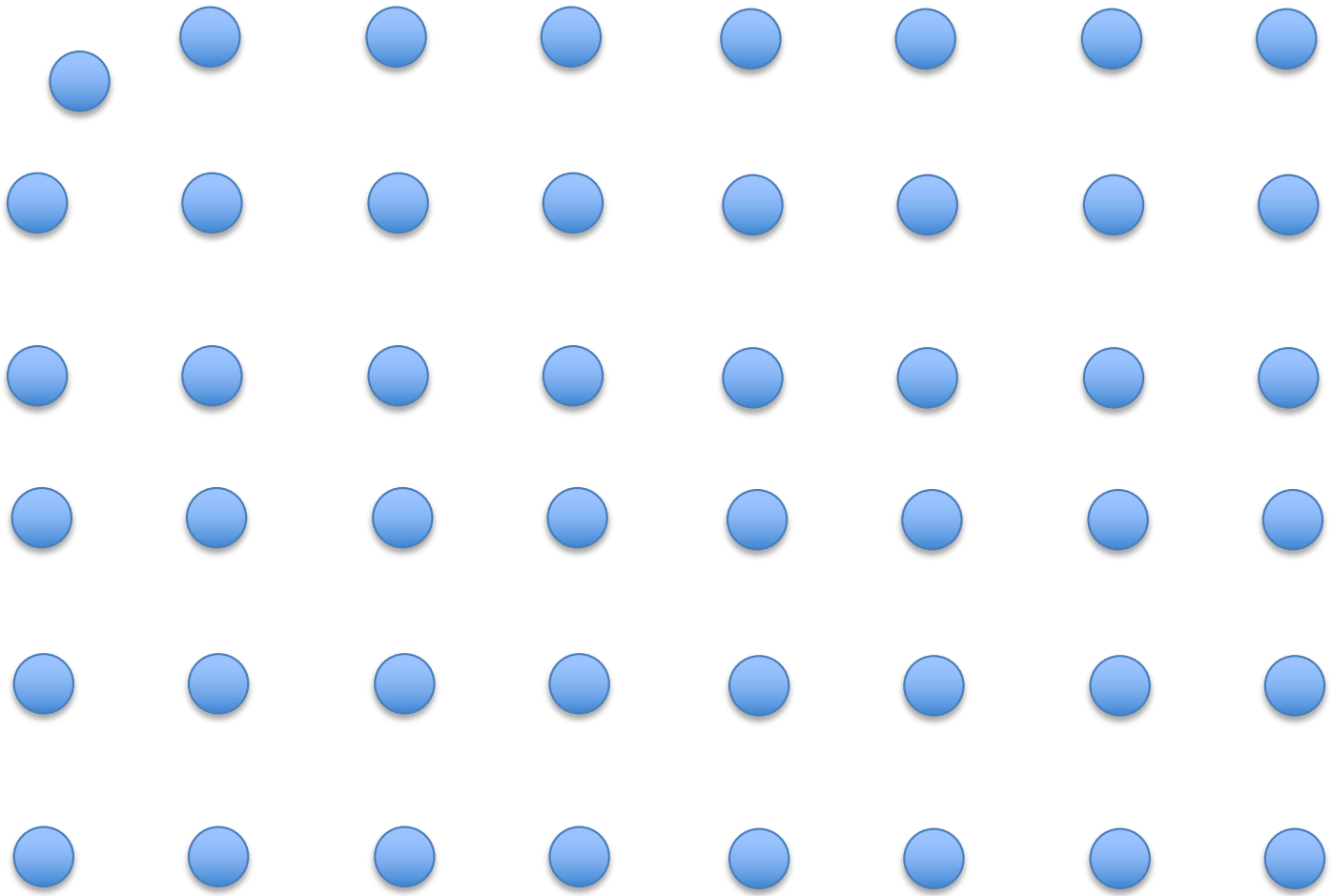
$$E_0 = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} V(d_{ij})$$



Make Trial Move of first Particle



Calculate new energy: $E_1 = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} V(d_{ij})$



- If $E_1 < E_0$ accept move, calculate F of new configuration, and add to ensemble.

- If $E_1 > E_0$ accept move with probability $e^{-(E_1 - E_0)/kT}$ **otherwise resample previous configuration**

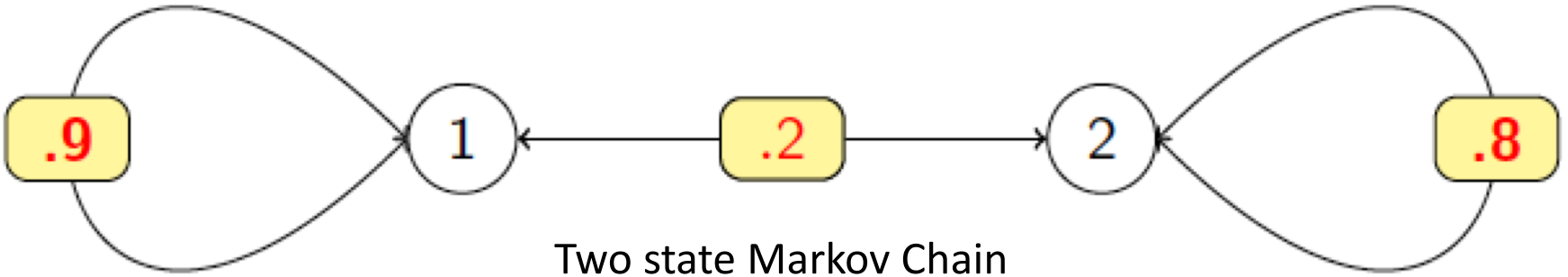
"It is, incidentally, clear from the above derivation that after a forbidden move we must count again the initial configuration. Not to do this would correspond in the above case to removing from the ensemble those systems which tried to move from s to r and were forbidden. This would unjustifiably reduce the number in state s relative to r."

Comments on proof

- The proof contained in text is not very mathematical but represents very clear and simple reasoning
- Kirkwood (LANL statistician) was skeptical
- Not clear authors understood result was not specific to the Boltzmann distribution

Linking to Modern Idea of MCMC

- Markov Chains are a simple idea:
  - Given a state space χ and transition probabilities between each pair of states, carry out random walk
  - Under suitable conditions after burn-in period node visitations represent samples from some underlying distribution p(x).



Two state Markov Chain

Markov Chain Monte Carlo

- Idea is to rig Markov chain to sample from specified distribution.
- What is known as the "Metropolis Algorithm" does this for general p(x) as:
  - Select an arbitrary stochastic trial matrix **Q** and starting state x_i. For Metropolis Q is symmetric.
  - For each step 1 … n

    Sample candidate next state j from i'th row of **Q**

$$P(i \to j) = min[1, p(i)/p(j)]$$

$$P(i \to i) = 1 - \sum_{i \neq j} p(i)/p(j)$$

Metropolis Algorithm

Note that for Boltzmann:

$$p(i)/p(j) = e^{(-\Delta E/kT)}$$

with

$$e^{-(\Delta E)/kT} > 1$$ for decreasing energy

so that for **decreasing energy** P(i,j) = 1

This exactly matches algorithm in paper

Rosenbluth Algorithm

- Fifty years later, just before his death, Rosenbluth asserted that he and Arianna did all of the work and that Metropolis only provided the computer time.

- When asked how he felt about changing the algorithm's name, Rosenbluth wasn't interested

"... life has been good to me. I feel rewarded in knowing that this algorithm will allow scientists to solve problems ranging from fluid flow to social dynamics to elucidating the nature of elementary particles."

Deterministic Nonperiodic Flow[1]

EDWARD N. LORENZ

Massachusetts Institute of Technology

(Manuscript received 18 November 1962, in revised form 7 January 1963)

ABSTRACT

Finite systems of deterministic ordinary nonlinear differential equations may be designed to represent forced dissipative hydrodynamic flow. Solutions of these equations can be identified with trajectories in phase space. For those systems with bounded solutions, it is found that nonperiodic solutions are ordinarily unstable with respect to small modifications, so that slightly differing initial states can evolve into considerably different states. Systems with bounded solutions are shown to possess bounded numerical solutions.

A simple system representing cellular convection is solved numerically. All of the solutions are found to be unstable, and almost all of them are nonperiodic.

The feasibility of very-long-range weather prediction is examined in the light of these results.

- 20,000 citations
- Only 3 in first 11 years!

Lorenz's driving questions

- *statistical* vs. *dynamic* weather forecasts
- Could weather analogs be trusted?
- How do errors in initial conditions propogate through scale and what are predictability limits?
- Study stability through numerical models

Lorenz 12-equation system

$$\frac{dX_i}{dt} = \sum_{m,n} a_{imn} X_m X_n + \sum_m b_{im} X_m + c_i$$

$$i, m, n = 1 : 12$$

- System still too large for numerical studies at the time
- Crazy idea at time
  - bought his own Royal Mcbee LGP-30, 60 Flop/s. 4Kb storage.

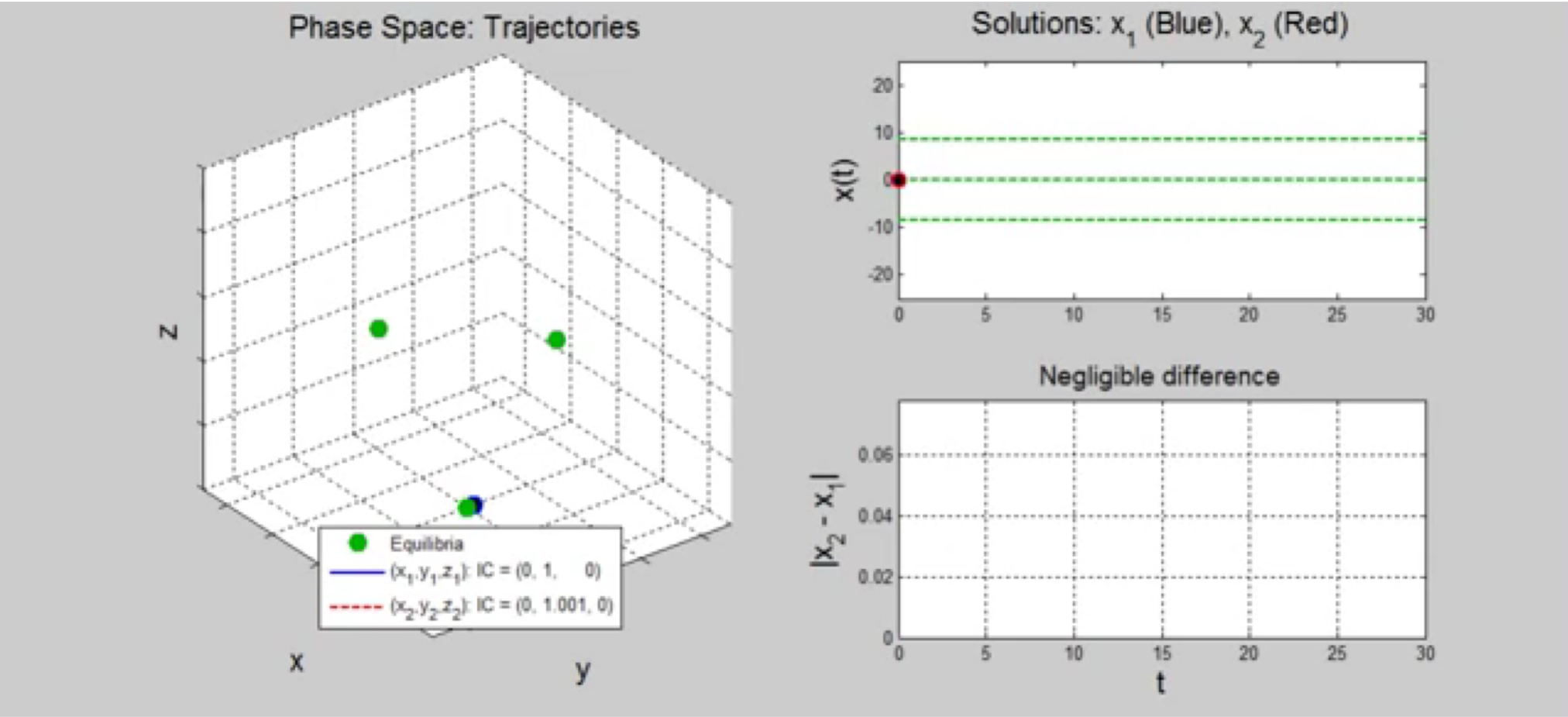
Lorenz 3-equation system

$$\frac{dx}{dt} = c_1(y - x)$$
$$\frac{dy}{dt} = x(c_2 - z) - y$$
$$\frac{dz}{dt} = xy - c_3 z$$

Restarted simulation to save time, with roundoff error. Massive change in solution was not expected. **Recounted in *Chaos* by James Gleick**

Developed in collaboration with Barry Saltzman

Lorenz Attractor



Key discovery from 1969 paper: Non-limiting process

- *"It is proposed that certain formally deterministic fluid systems which possess many scales of motion are observationally indistinguishable from indeterministic systems; specifically, that two states of the system differing initially by a small "observational error" will evolve into two states differing as greatly as randomly chosen states of the system within a finite time interval, which cannot be lengthened by reducing the amplitude of the initial error*." –Lorenz (1969)

Lorenz's Contributions

- In discovering "deterministic chaos," Dr. Lorenz established a principle that "profoundly influenced a wide range of basic sciences and brought about one of the most dramatic changes in mankind's view of nature since Sir Isaac Newton," said a committee that awarded him the 1991 Kyoto Prize for basic sciences.