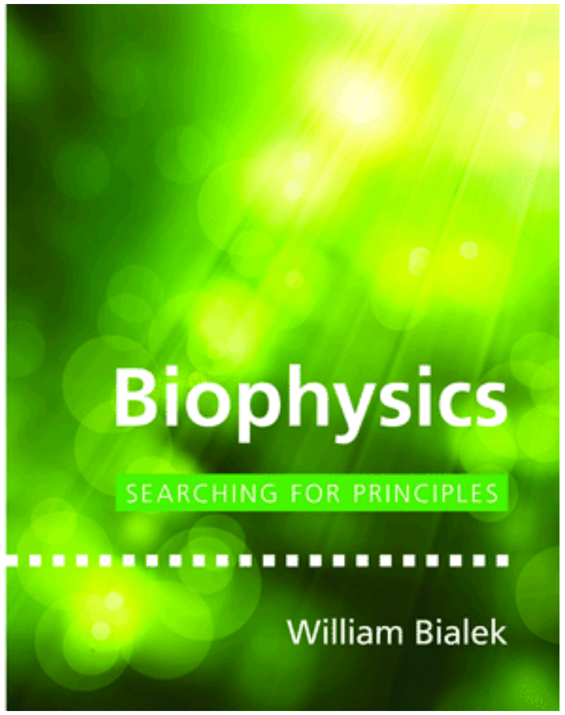
Statistical Physics of The Symmetric Group

……

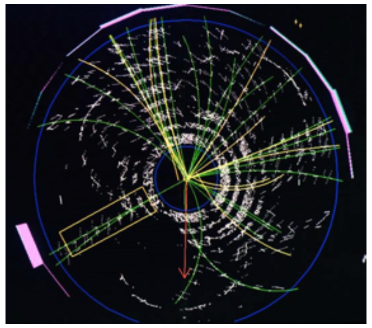
Mobolaji Williams — Harvard Physics Oral Qualifying Exam— Dec. 12, 2016

Theoretical Physics of Living Systems

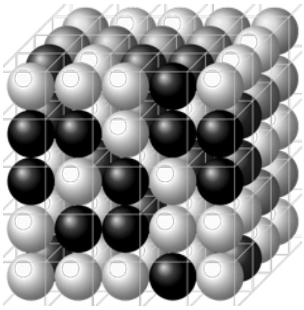


Biophysics: A Search for Principles
William Bialek

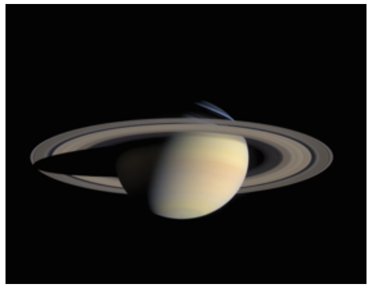
Physics



Particle Physics



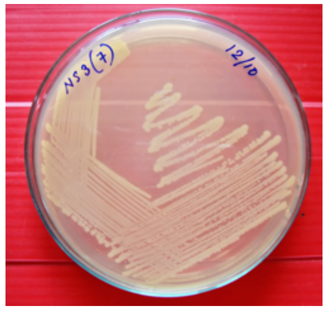
Condensed Matter



Astrophysics

"All of (theoretical) physics shares a method of approach"

Biology



Microbiology

Plant Biology

Marine Biology

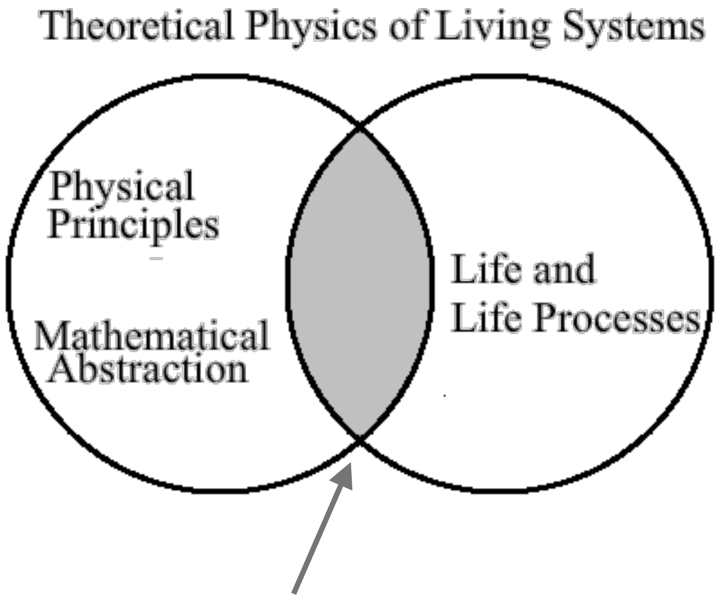
"All of biology shares an object of study"

What then is biological physics?

or, more specifically,

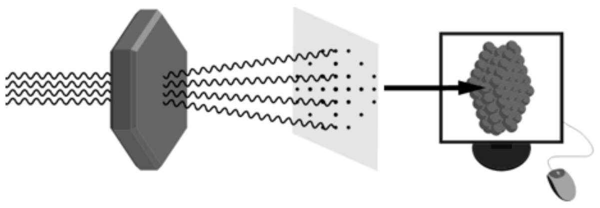
What constitutes the theoretical physics of living systems?

Theoretical Physics of Living Systems

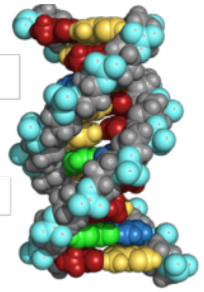


Questions/methods which are of interest/helpful to physicists and biologists

Physics can help biology



→



X-Ray Crystallography

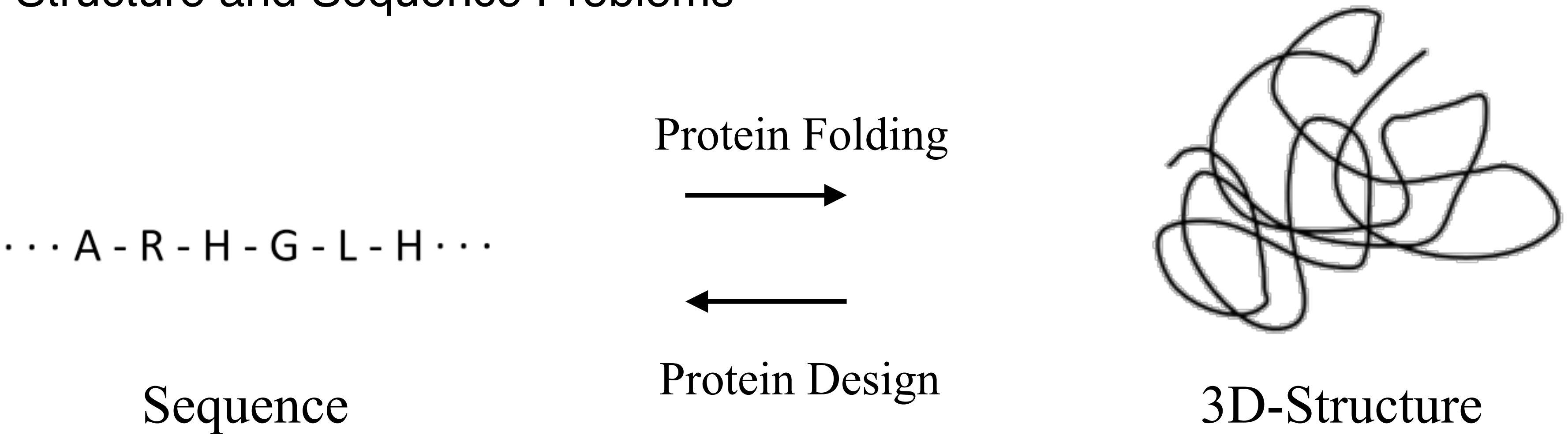
DNA Structure

But can biology help physics?

Some biological problem/idea → Something interesting to physicists

Protein Folding and Design

Two Protein Structure and Sequence Problems



$\cdots$ A - R - H - G - L - H $\cdots$

Protein Folding

Protein Design

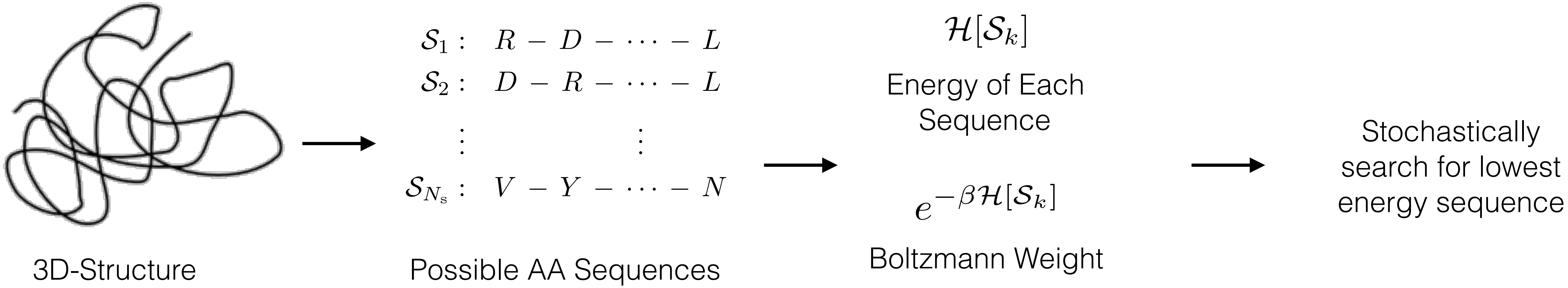
Sequence

3D-Structure

This work is concerned with **Protein Design**

Common sense statistical mechanics approach to problem

(Shakhnovich and Gutin. "A new approach to the design of stable proteins."(1993))



$$\begin{aligned} \mathcal{S}_1 &: \; R - D - \cdots - L \\ \mathcal{S}_2 &: \; D - R - \cdots - L \\ &\;\;\vdots \qquad\qquad \vdots \\ \mathcal{S}_{N_\mathrm{s}} &: \; V - Y - \cdots - N \end{aligned}$$

$\mathcal{H}[\mathcal{S}_k]$

Energy of Each Sequence

$e^{-\beta \mathcal{H}[\mathcal{S}_k]}$

Boltzmann Weight

Stochastically search for lowest energy sequence

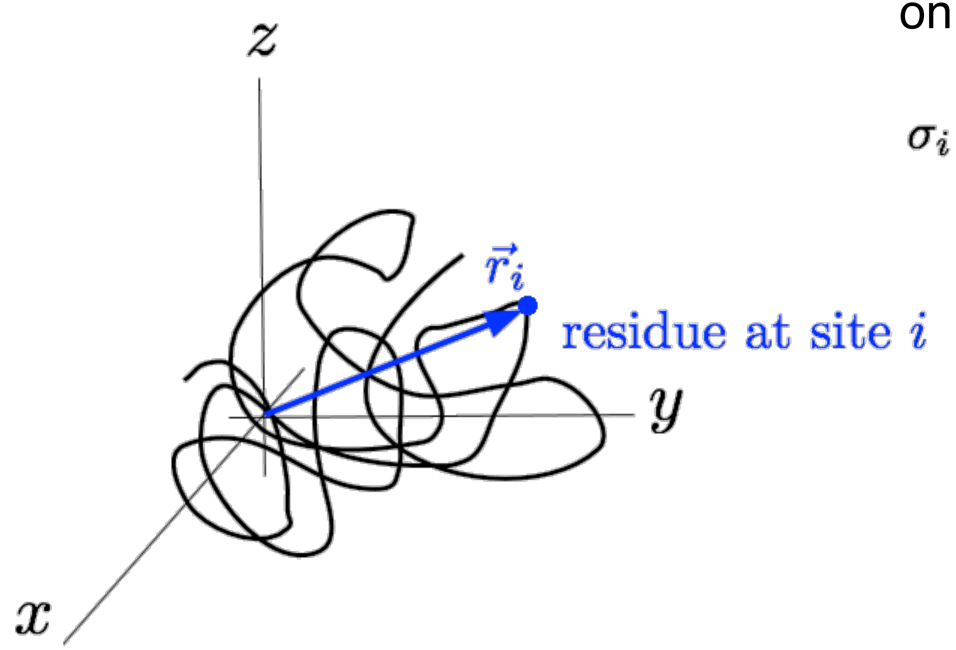
3D-Structure

Possible AA Sequences

Protein Folding and Design

Concrete Example:

Begin with a simple 3D structure with *N* amino acids



State Space

For simplicity, split amino acids into only polar and nonpolar residues

$$\sigma_i = \begin{cases} 1 & \text{nonpolar residue} \\ 0 & \text{polar residue} \end{cases}$$

[Solution to design problem consists of the sequence $\{\sigma_1, \ldots, \sigma_N\}$]

Hamiltonian

non-polar residues interact through a hydrophobic interaction for a given distance; all other contacts non-interacting

$$\mathcal{H}(\{\sigma_k\}) = \frac{1}{2} \sum_{i,j} U_0 \sigma_i \sigma_j \Delta(|\vec{r}_i - \vec{r}_j|)$$

$$\Delta(r) = \begin{cases} 1 & r_{\text{low}} < r < r_{\text{high}} \\ 0 & \text{otherwise} \end{cases}$$

$$(U_0 < 0)$$

The stochastic solution is easy to predict

[Lowest energy sequence is $\{\sigma_i\} = \{1, \ldots, 1\}$]

All residues are nonpolar

The polypeptide is a homopolymer?

Protein Folding and Design

Problem with formalism:

> *"For virtually any target structure, the lowest energy sequence will be a homopolymer, consisting of the amino acid with the largest self-attraction"**

*(Morrissey and Shakhnovich. "Design of proteins with selected thermal properties." (1996))

3D- Structure

→

$$I - \cdots - I - \cdots I$$

Homopolymer?

<u>This cannot be correct…</u>

Suggested way forward:

Constrain amino acid composition

New Design Question:
Given **a fixed number of each type of amino-acid**
What sequence yields the lowest energy?

Implication of New Design Question

For a chain of length *N*, what is the size of our state space?

✗ 20^N (or 2^N under a polar/nonpolar framing)

✓ $\dfrac{N!}{n_1!n_2!\cdots n_{20}!}$ (or $\dfrac{N!}{n_{\text{polar}}!n_{\text{nonpolar}}!}$ under a polar/nonpolar framing)

We should search over the space of **permutations** of components

Model Building on the Polypeptide Chain

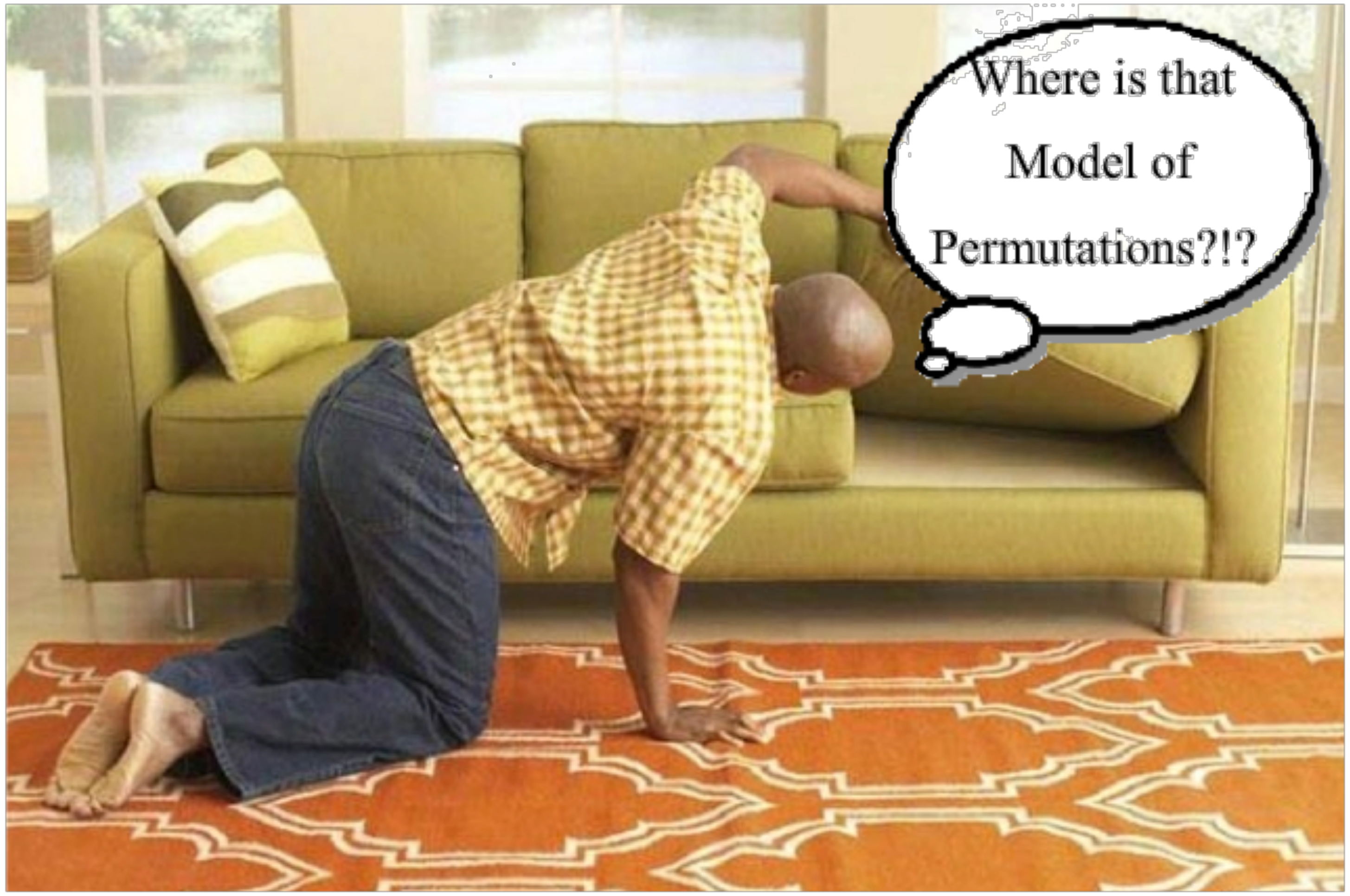
For a statistical mechanics model of protein design, we can construct a spin-like model

e.g., $\mathcal{H}(\{\sigma_k\}) = \frac{1}{2}\sum_{i,j} U_0 \sigma_i \sigma_j \Delta(|\vec{r}_i - \vec{r}_j|)$

But it is apparent that if we assume a fixed number of amino acids of each type, our space of states consists of various **permutations** of the amino acid list

i.e., $|\mathcal{S}_N| \neq 20^N$ but , $|\mathcal{S}_N| = \frac{N!}{n_0! n_1! \ldots n_{20}!}$

Where are statistical mechanics models where the space of states consists of permutations?



No such models seem to exist…

…So we must build one for ourselves!

Model of Permutations

Goal: Develop a statistical physics model of permutations

Simplifications

1. Consider equilibrium statistical mechanics $\lim_{t \to \infty} P(\mathcal{S}, t) = P^{\text{eq}}(\mathcal{S})$
2. Assume we only have one type of each unit i.e., $n_k = 1$ so $|\mathcal{S}_N| = N!$

Our System:
a chain (considered only in a single dimension) with *N* subunits (e.g., amino acids) which can be placed in various orderings.

Our Space of States:
By definition, our space of states is isomorphic to the **Symmetric Group.**

(Dixon and Mortimer, *Permutation Group,* 1987)

> " Let Ω be an arbitrary nonempty set... The set of all permutations of Ω forms a group, under composition of mappings, called the **symmetric group on Ω**."

Indicates that the mathematics for this would involve combinatorics, symmetric polynomials, derangements, etc.

e.g.,

$$\binom{N}{j} = \frac{\Gamma(N+1)}{\Gamma(j+1)\Gamma(N-j+1)}$$

$$\prod_{k=1}^{N} (1 + z\, x_i) = \sum_{\ell=0}^{N} z^{\ell}\, \Pi_{\ell}(x_1, x_2, \ldots, x_N)$$

$$d_N \equiv N! \sum_{k=0}^{N} \frac{(-1)^k}{k!}$$

Model of Permutations

Let’s define the system more precisely and introduce a Hamiltonian.

System Definition:

- subunits are labeled as ω_i with $i = 1, \ldots, N$
- the ordering of subunits with the zero energy is $\vec{\omega} \equiv (\omega_1, \omega_2, \ldots, \omega_N)$
- an arbitrary state is $\vec{\theta}$ where $\vec{\theta} \in \{\text{perm}(\omega_1, \omega_2, \ldots, \omega_N)\} \equiv Sym(\omega)$

Energy Definition:

- The state $\vec{\theta} = \vec{\omega}$ has zero energy and has subunits in the *correct* order. For all other states, there is an energy cost of λ_i for $\theta_i \neq \omega_i$.

$$\mathcal{H}_N(\{\theta_i\}) = \sum_{i=1}^{N} \lambda_i I_{\theta_i \neq \omega_i}$$

$$I_A = \begin{cases} 1 & A \text{ is true} \\ 0 & A \text{ is false} \end{cases}$$

Example: Three components

State	Energy
$(\omega_1, \omega_2, \omega_3)$	0
$(\omega_2, \omega_1, \omega_3)$	$\lambda_1 + \lambda_2$
$(\omega_3, \omega_2, \omega_1)$	$\lambda_1 + \lambda_3$
$(\omega_1, \omega_3, \omega_2)$	$\lambda_2 + \lambda_3$
$(\omega_2, \omega_3, \omega_1)$	$\lambda_1 + \lambda_2 + \lambda_3$
$(\omega_3, \omega_1, \omega_2)$	$\lambda_1 + \lambda_2 + \lambda_3$

Note: There is a a **2-fold** degeneracy for the most disordered state.

In general, for *N* subunits, the degeneracy of the most disordered state is. . .

$$d_N \equiv N! \sum_{k=0}^{N} \frac{(-1)^k}{k!}$$

“number of derangements”:
\# of ways
to *completely*
reorder list

Model of Permutations

System Definition:

- subunits are labeled as ω_i with $i = 1, \dots, N$
- the ordering of subunits with the zero energy is $\vec{\omega} \equiv (\omega_1, \omega_2, \dots, \omega_N)$
- an arbitrary state is $\vec{\theta}$ where $\vec{\theta} \in \{\mathrm{perm}(\omega_1, \omega_2, \dots, \omega_N)\} \equiv Sym(\omega)$

Energy Definition:

- The state $\vec{\theta} = \vec{\omega}$ has zero energy and has subunits in the *correct* order. For all other states, there is an energy cost of λ_i for $\theta_i \neq \omega_i$.

$$\mathcal{H}_N(\{\theta_i\}) = \sum_{i=1}^{N} \lambda_i I_{\theta_i \neq \omega_i}$$

$$I_A = \begin{cases} 1 & A \text{ is true} \\ 0 & A \text{ is false} \end{cases}$$

Now let’s compute the partition function

$$d_N \equiv N! \sum_{k=0}^{N} \frac{(-1)^k}{k!}$$

$$Z_N(\{\beta\lambda_i\}) = \sum_{\vec{\theta} \in Sym(\omega)} \exp\left(-\beta \sum_{i=1}^{N} \lambda_i \, I_{\theta_i \neq \omega_i}\right)$$

e.g.,

$$\Pi_2(x_1, x_2, x_3) = x_1 x_2 + x_2 x_3 + x_3 x_1$$

$= $ [Application of derangement formula, symmetric polynomials, etc.]

$$\boxed{Z_N(\{\beta\lambda_i\}) = \int_0^\infty ds\, e^{-s} \prod_{\ell=1}^{N} \left[1 + (s-1)e^{-\beta\lambda_\ell}\right]}$$

. . . so we can obtain a closed form expression . . .

. . .but what does it mean?

Model of Permutations

Energy Definition:

- The state $\vec{\theta} = \vec{\omega}$ has zero energy and has subunits in the *correct* order. For all other states, there is an energy cost of λ_i for $\theta_i \neq \omega_i$.

$$\mathcal{H}_N(\{\theta_i\}) = \sum_{i=1}^{N} \lambda_i I_{\theta_i \neq \omega_i}$$

Partition Function:

- $Z_N(\{\beta\lambda_i\}) = \int_0^\infty ds\, e^{-s} \prod_{\ell=1}^{N} \left[1 + (s-1)e^{-\beta\lambda_\ell}\right]$

What physics is contained in this partition function?

Progress comes from a simplification: $\lambda_i = \lambda_0 \quad \text{for all } i$

"The same energy penalty for each subunit"

The partition function then simplifies to

$$Z_N(\beta\lambda_0) = \int_0^\infty ds\, e^{-s} \left(1 + (s-1)e^{-\beta\lambda_0}\right)^N$$

And by steepest descent we find the free energy

$$\begin{aligned}\beta F &= -\ln Z_N(\beta\lambda_0) \\ &\simeq N\beta\lambda_0 - \left(e^{\beta\lambda_0} - N - 1\right) + F_0(N).\end{aligned}$$

And the **average number of incorrect components** is

$$\langle j \rangle \equiv \left\langle \sum_{i=1}^{N} I_{\theta_i \neq \omega_i} \right\rangle = \frac{\partial(\beta F)}{\partial(\beta\lambda_0)} \longrightarrow \boxed{\langle j \rangle \simeq N - e^{\beta\lambda_0}}$$

Model of Permutations

We are trying to understand the physics of

$$Z_N(\beta\lambda_0) = \int_0^\infty ds\, e^{-s}\left(1 + (s-1)e^{-\beta\lambda_0}\right)^N$$

By considering the average number of incorrect components…

$$\langle j \rangle \equiv \left\langle \sum_{i=1}^{N} I_{\theta_i \neq \omega_i} \right\rangle \simeq N - e^{\beta\lambda_0}$$

The function on the right has no constraints, but <u>the average number of incorrect components must be greater than 0</u>.

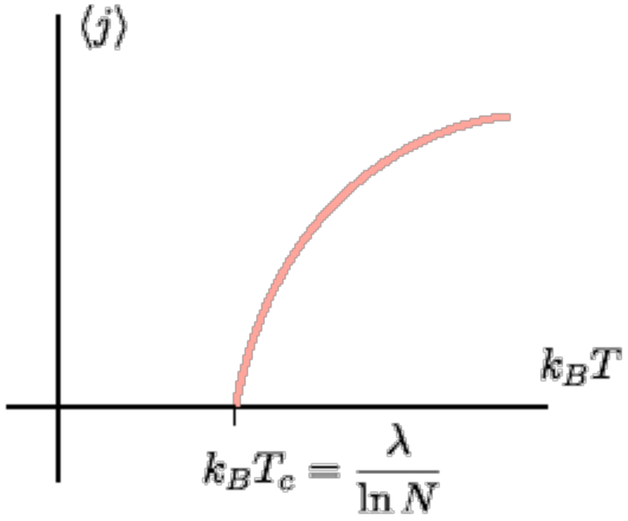
(Result is only valid so long as $\langle j \rangle \geq 0$.)

Implies the existence of a **transition temperature:**

$$k_B T_c = \frac{\lambda}{\ln N}$$

The correct solution is then

$$\boxed{\langle j \rangle / N = \begin{cases} 0 & \text{for } T < T_c \\ 1 - e^{\beta\lambda_0}/N & \text{for } T > T_c \end{cases}}$$



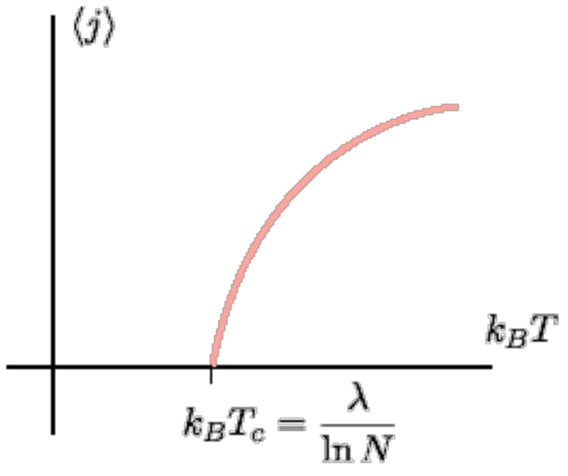
→ Above a certain transition temperature, **there is a spectrum of states** (different from "correct state") **which defines the free energy minimum**

Model of Permutations

This system appears to exhibit phase transitions. . .

“Average number of incorrect components”

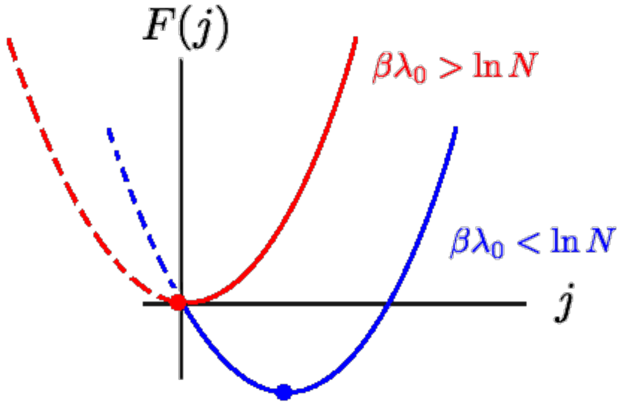
$$\langle j \rangle / N = \begin{cases} 0 & \text{for } T < T_c \\ 1 - e^{\beta\lambda_0}/N & \text{for } T > T_c \end{cases}$$



. . . But we do NOT actually have a phase transition here

The **functional form** of the free energy **remains the same** above and below transition temperature…

$$F(j) = \lambda_0 j - \frac{1}{\beta} \ln \left[\binom{N}{j} d_j \right]$$



It is only the excluded region of the free energy that changes.

Qualitative argument for why there is no phase transition:

<u>Modified Landau-Van Hove:</u>

as *N* becomes large, entropy (i.e., disorder) always beats energy (i.e., order)

Model of Permutations

For the model of permutations with Hamiltonian

$$\mathcal{H}_N(\{\theta_i\}) = \lambda_0 \sum_{i=1}^{N} I_{\theta_i \neq \omega_i}$$

$$\langle j \rangle / N = \begin{cases} 0 & \text{for } T < T_c \\ 1 - e^{\beta\lambda_0}/N & \text{for } T > T_c \end{cases}$$

We found that above a certain amount of thermal fluctuations, the "correct ordering" (*j=0*) is no longer thermodynamically favored.

Interesting result, but this model is relatively simple.

How would we extend it to the case where sites directly interact with one another?

Full Generalization:
Introduce site-dependent interaction terms and higher-order contributions to Hamiltonian:

$$\mathcal{H}_N(\{\theta_i\}) = \sum_i \lambda_i I_{\theta_i \neq \omega_i} + \frac{1}{2} \sum_{i,j} \mu_{ij} I_{\theta_i \neq \omega_i} I_{\theta_j \neq \omega_j} + \frac{1}{3!} \sum_{i,j,k} \tau_{ijk} I_{\theta_i \neq \omega_i} I_{\theta_j \neq \omega_j} I_{\theta_k \neq \omega_k}$$

(energy cost of incorrectness depends on **one other or two other** components.)

But this is too complicated!

Let's begin more simply, with two terms and no site-dependent interactions:

$$\mathcal{H}_N(\{\theta\}) = \lambda_1 \sum_{i=1}^{N} I_{\theta_i \neq \omega_i} + \frac{\lambda_2}{2N} \sum_{i,j=1}^{N} I_{\theta_i \neq \omega_i} I_{\theta_j \neq \omega_j}$$

What is the phenomenology of such a system?

Interacting Model

Goal:
Understand the basic physics
of the "all-site-interacting" model:

$$\mathcal{H}_N(\{\theta_i\}) = \lambda_1 \sum_{i=1}^{N} I_{\theta_i \neq \omega_i} + \frac{\lambda_2}{2N} \sum_{i,j=1}^{N} I_{\theta_i \neq \omega_i} I_{\theta_j \neq \omega_j}$$

We took three approaches to this end…

- **"Landau" Free Energy**

$$F_N(j,\beta) = \lambda_1 j + \frac{\lambda_2}{2N} j^2 + \frac{1}{\beta} \ln \Gamma(N - j + 1) + f_0$$

- **Gibbs-Feynman Inequality**

$$F[\mathcal{H}] \leq F[\mathcal{H}_0] + \langle \mathcal{H} - \mathcal{H}_0 \rangle_0 .$$

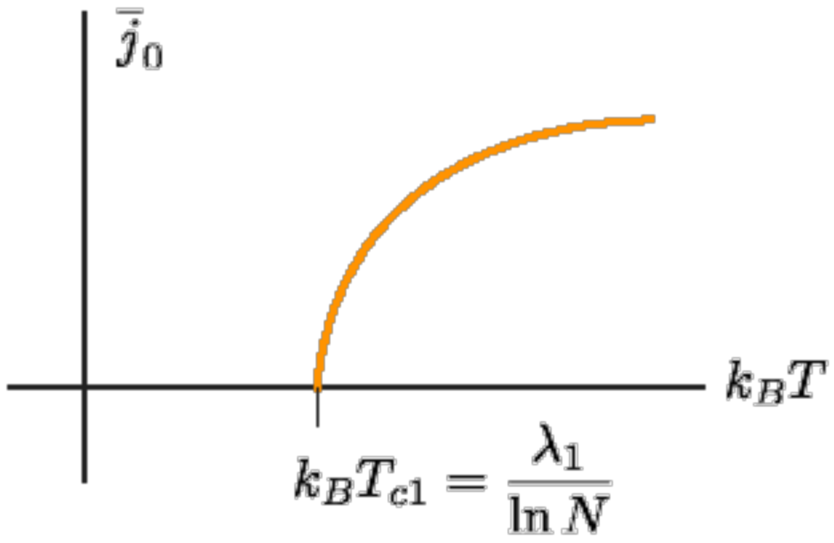
- **Method of Steepest Descent**

$$Z_N(\beta; \lambda_1, \lambda_2) = \sqrt{\frac{N}{2\pi\beta|\lambda_2|}} \int_0^{\infty} ds \int_{-\infty}^{\infty} dx \, \left(1 + (s-1)e^{-\beta\lambda_1 - x}\right)^N e^{-s - Nx^2/2\beta|\lambda_2|} .$$

…and all three yield the same order-parameter.

$$\boxed{\frac{\bar{j}_0}{N} = 1 - \frac{1}{\beta\lambda_2} W_0\left(\frac{\beta\lambda_2}{N} e^{\beta\lambda_1 + \beta\lambda_2}\right) + \mathcal{O}\left(\frac{1}{N}\right)}$$

where $W_0(xe^x) = x$ (for $W_0 > -1$)

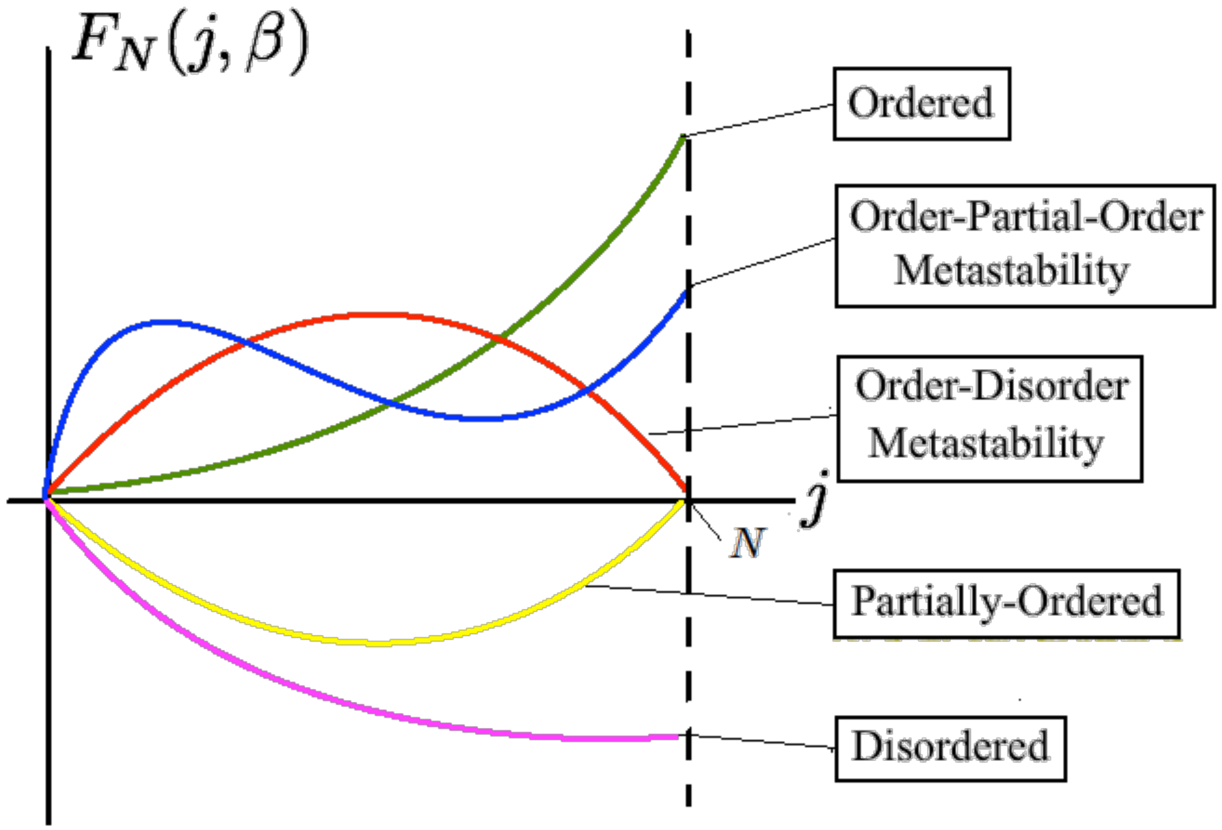


Interacting Model

The utility of the Landau Free Energy is that we get a qualitative picture the equilibria of the system.

$$F_N(j,\beta) = \lambda_1 j + \frac{\lambda_2}{2N} j^2 + \frac{1}{\beta}\ln\Gamma(N-j+1) + f_0$$

For various parameter choices, we find different types of free energy functions



Parameter Space Regions

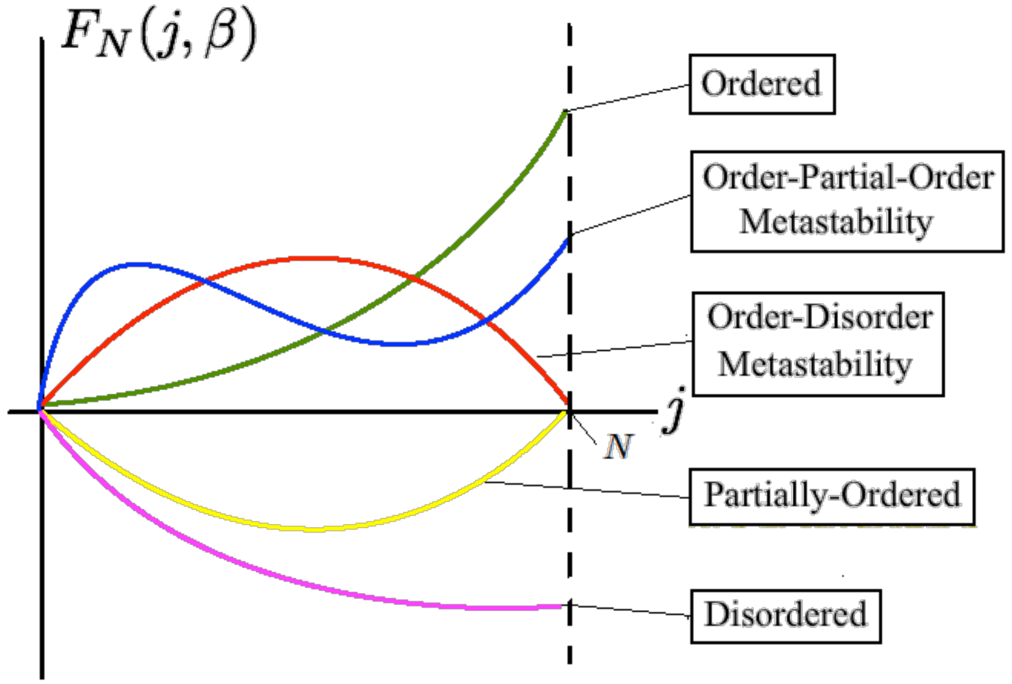
- **Ordered Region:** $j = 0$ stable
- **Order and Partial-Order Metastable Region:** $j = 0$ and $j = \bar{j}_0$ stable
- **Order and Disorder Metastable Region:** $j = 0$ and $j = N$ stable
- **Partially-Ordered Region:** $j = \bar{j}_0$ stable
- **Disordered Region:** $j = N$ stable

Interacting Model

Each type of free energy function is associated with a specific region in the parameter space of the Hamiltonian

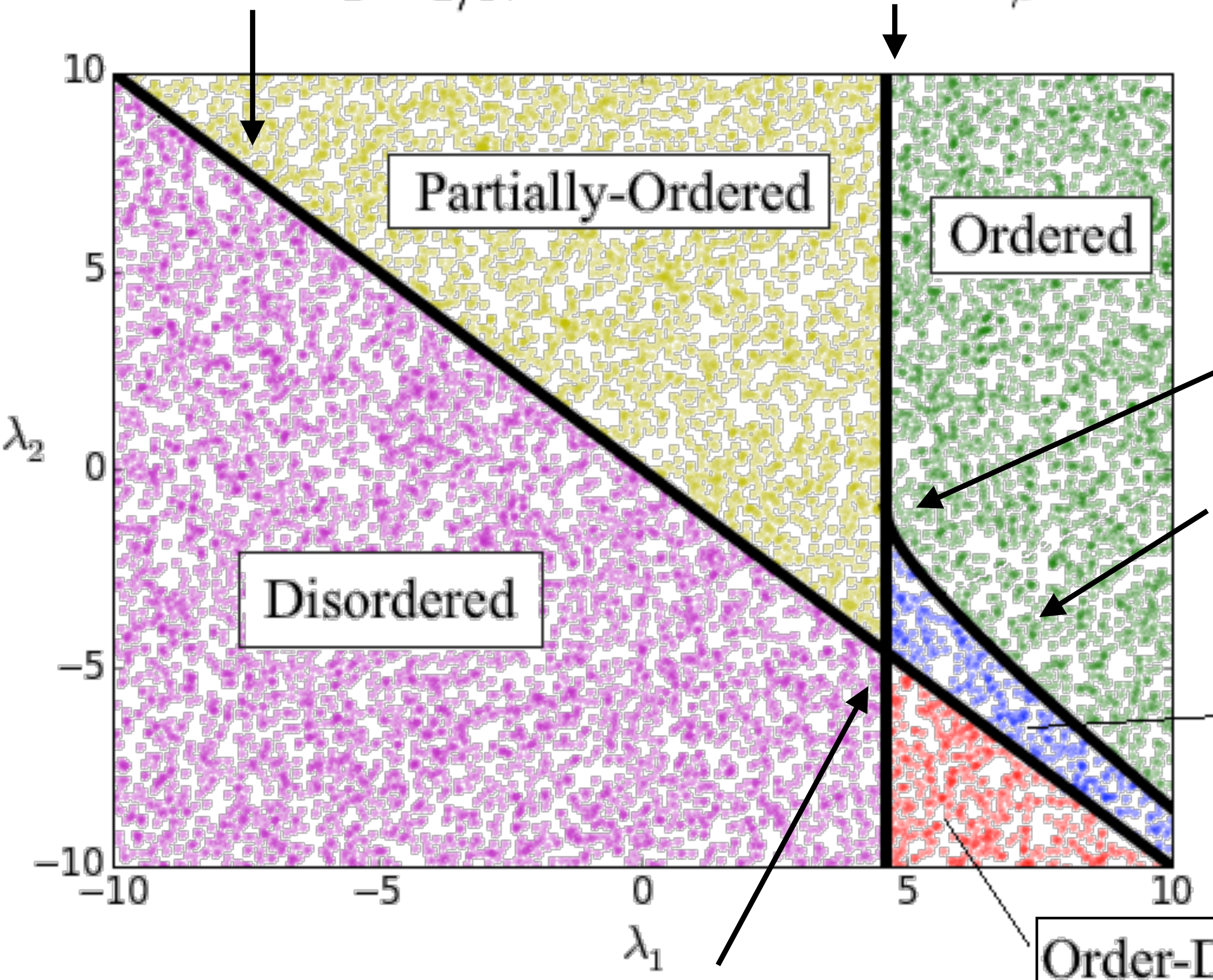
$$W_0(xe^x) = x \ (\text{for } W_0 > -1)$$

$$W_{-1}(xe^x) = x \ (\text{for } W_{-1} \le -1)$$



$$\lambda_2 = -\frac{\lambda_1}{1 - 1/N}$$

$$\lambda_1 = \frac{\ln N}{\beta}$$



$$(\ln N/\beta, -1/\beta)$$

$$\lambda_2 = \frac{1}{\beta} W_{-1}\left(-\frac{N}{e} e^{-\beta\lambda_1}\right)$$

Order-Partial-Order Metastability

Order-Disorder Metastability

$$(\ln N/\beta, -\ln N/\beta)$$

<u>Features</u>

- Three regime boundaries
- Two transition temperatures
  - $k_B T_{c1} = \dfrac{\lambda_1}{\ln N}$
  - $\dfrac{1}{k_B T_{c2}} = \dfrac{\lambda_1 + \lambda_2}{W_0\left(-\frac{N}{e\lambda_2}(\lambda_1 + \lambda_2)\right)}$
- A quadruple and a triple point

1D Chain of Partition Functions

Equilibrium properties can be framed as **a chain of partition functions**

For Example

If we have a Hamiltonian written in terms of the number of correct components as . . .

$$\mathcal{H}_N(\{\theta_i\}) : \mathcal{E}_N(j) = \lambda j$$

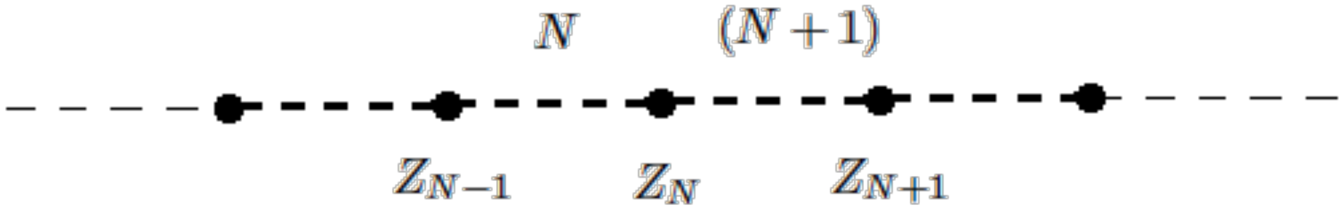
Then we know that the associated partition function is . . .

$$Z_N(\beta\lambda) = \int_0^\infty ds\, e^{-s} \left(1 + (s-1)e^{-\beta\lambda}\right)^N$$

This partition function satisfies the identity

$$\frac{\partial}{\partial\beta} Z_N(\beta\lambda) = -\lambda N \Big(Z_N(\beta\lambda) - Z_{N-1}(\beta\lambda) \Big)$$

. . . which is reminiscent of a 1D chain of oscillators . . .



Taking the “partition function chain” to the continuum limit we find

$$\left(\frac{\partial}{\partial\beta} + \lambda \frac{\partial}{\partial \ln \eta} \right) \mathcal{Z}(\eta, \beta\lambda) = 0$$

$$\text{(with boundary condition } \mathcal{Z}(\eta, 0) = \Gamma(\eta + 1))$$

1D Chain of Partition Functions

Continuum limit of Partition Function Chain

$$\left(\frac{\partial}{\partial\beta}+\lambda\frac{\partial}{\partial\ln\eta}\right)\mathcal{Z}(\eta,\beta\lambda)=0 \qquad \text{(with boundary condition } \mathcal{Z}(\eta,0)=\Gamma(\eta+1)\text{)}$$

Solution to PDE:

$$\mathcal{Z}(\eta,\beta\lambda)=\Gamma(\eta e^{-\beta\lambda}+1)=\int_0^\infty dt\,\exp\left[\eta e^{-\beta\lambda}\ln t-t\right]$$

Order Parameter:

$$\bar{j}=-\frac{\partial}{\partial(\beta\lambda)}\ln\mathcal{Z}(\eta,\beta\lambda)=\eta e^{-\beta\lambda}\psi_0(\eta e^{-\beta\lambda}+1) \qquad *\ \psi_0(x)\equiv\frac{\Gamma'(x)}{\Gamma(x)}$$

By the condition $\bar{j}\geq 0$,
we find a transition temperature $\quad k_BT_c=\dfrac{\lambda}{\ln(\eta/\epsilon)}$

(where $\psi_0(1+\epsilon)=0$)

Which is reminiscent of the discrete result

$$Z_N(\beta\lambda)=\int_0^\infty ds\,e^{-s}\left(1+(s-1)e^{-\beta\lambda}\right)^N \quad\rightarrow\quad k_BT_c=\frac{\lambda}{\ln N}$$

1D Chain of Partition Functions

Continuum limit of Partition Function Chain

We find a similar correspondence between

$$\mathcal{E}_N(j) = \lambda_1 j + \lambda_2 j^2$$

(original permutations model)

and

$$\left(\frac{\partial}{\partial\beta} + \lambda_1 \frac{\partial}{\partial \ln\eta} + \lambda_2 \frac{\partial^2}{\partial \ln\eta^2}\right) \mathcal{Z}(\eta, \beta\lambda_1, \beta\lambda_2) = 0$$

(continuum model)

We can generalize these results…

Physics of
original partition function
of permutations

$\simeq$

Physics of
continuum limit of
partition function chain

$$Z_N[\beta\mathcal{E}_N] = \sum_{j=0}^{N} g_N(j) e^{-\beta\mathcal{E}_N(j)}$$

$\simeq$

Solution to

$$\left[\frac{\partial}{\partial\beta} + \mathcal{E}\left(\eta, \frac{\partial}{\partial \ln\eta}\right)\right] \mathcal{Z}(\eta; \beta\mathcal{E}) = 0$$

$$* \ \mathcal{E}_N(j) \equiv \sum_{k=0}^{M} \lambda_k(N) j^k$$

Continuum Model

Our model of permutations is connected to a PDE,…

$$Z_N[\beta\mathcal{E}_N] = \sum_{j=0}^{N} g_N(j)\, e^{-\beta\mathcal{E}_N(j)} \quad \rightarrow \quad \left[\frac{\partial}{\partial\beta} + \mathcal{E}\left(\eta, \frac{\partial}{\partial \ln\eta}\right)\right]\mathcal{Z}(\eta;\beta\mathcal{E}) = 0$$

…and the PDE seems to yield consistent results…

Transition Temperature

Discrete Model	Continuum Model
$\lambda/\ln N$	$\lambda/\ln(\eta/\epsilon)$

…but what *other* questions could we be asking?

Numerical Studies
(Finite Difference Methods)

Return to discrete system

$$\frac{\partial^2\mathcal{Z}(\beta)}{\partial\eta^2} \simeq \frac{\mathcal{Z}^{\ell}_{i+1} - 2\mathcal{Z}^{\ell}_{i} + \mathcal{Z}^{\ell}_{i-1}}{\Delta n}$$

Possible Question:
Is analyzing system with
finite difference methods more
efficient than standard MC methods?

Analytical Studies
(Implications of the Path Integral)

Schrödinger-Like PDE, suggest existence of…

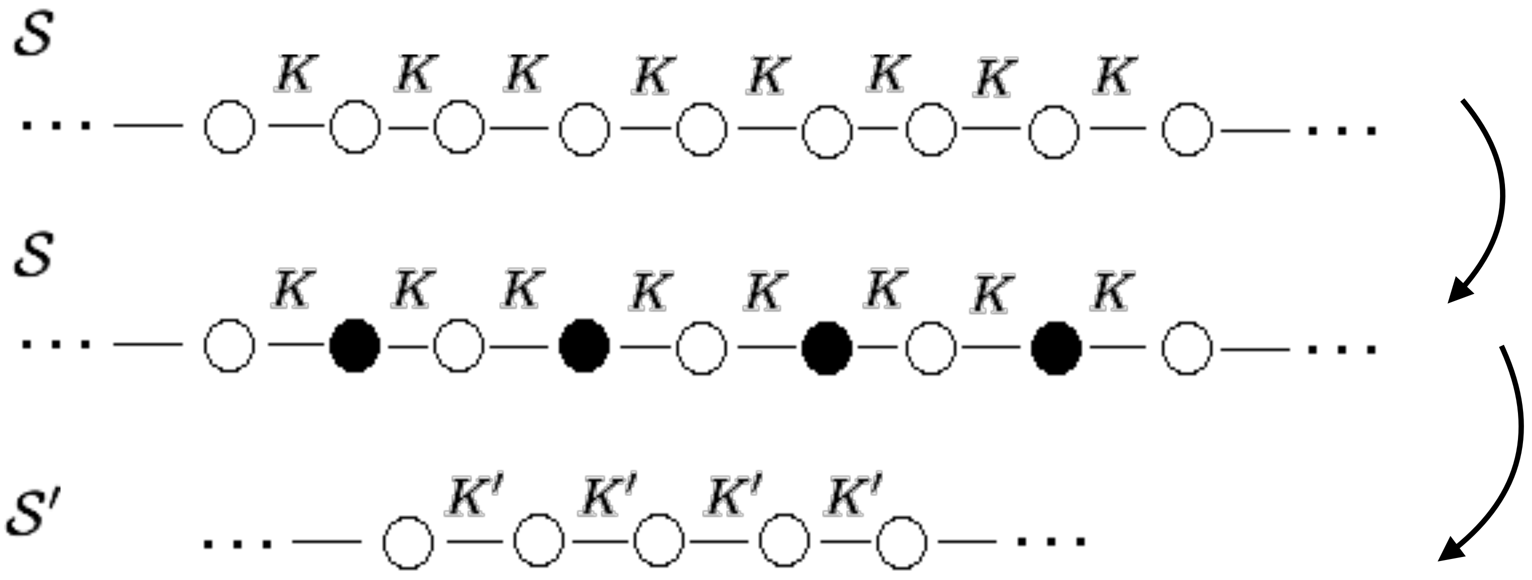
$$K(\ln\eta, \ln\eta_1; \beta\mathcal{E}) = \int \mathcal{D}[\ln\eta]\, \exp\left[-\int_0^{\beta} d\beta_1\, A(\eta, \ln\eta'; \beta)\right]$$

with A an action-like function.

Possible Question:
Does this yield new way
to study system perturbatively?

Extension: Renormalization Group

Steps of a Renormalization Group Transformation: Ising Model



1. Label half of the sites/d.o.f.s

2. Sum over all states the labeled sites can occupy

Resulting system is the "renormalized/effective theory"!

* Procedure allows us establish relations between original and renormalized theory (e.g. for the coupling constant)

$$K' = \frac{1}{2} \ln \cosh(2K)$$

Key to Procedure:
What makes this possible is that the state space of the Ising Model can be factorized.

$$\mathcal{S}_{\text{ising}} = \mathcal{S}_1 \otimes \mathcal{S}_2 \otimes \cdots \otimes \mathcal{S}_N$$

We need a new way to perform the renormalization group transformation for this system of permutations.

Problem:
How do we sum over the states of half the indices when the state space cannot be factorized?

$$\mathcal{S}_{\text{perm}} \neq \mathcal{S}_1 \otimes \mathcal{S}_2 \otimes \cdots \otimes \mathcal{S}_N$$

Extension: Permutation Glass

Similar to the Sherrington-Kirkpatrick Model, we can imagine a system of **permutations** with **quenched disorder**.

Mathematically more tractable, and still potentially interesting given non-trivial nature of state space

There are two choices in this direction:

1. $$\mathcal{H}_N(\{\theta_i\}) = \sum_{i=1}^{N} \lambda_i I_{\theta_i \neq \omega_j} \qquad \text{where } \lambda_i \sim \mathcal{N}(\lambda_0, \tilde{\lambda}^2)$$

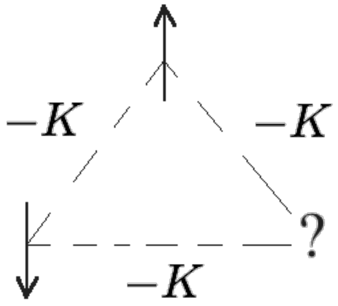
More Analogous to SK Spin glass Model

2. $$\mathcal{H}_N(\{\theta_i\}) = \frac{1}{2} \sum_{i,j=1}^{N} \mu_{ij} I_{\theta_i \neq \omega_i} I_{\theta_j \neq \omega_j} \qquad \text{where } \mu_{ij} \sim \mathcal{N}(\mu_0, \tilde{\mu}^2)$$

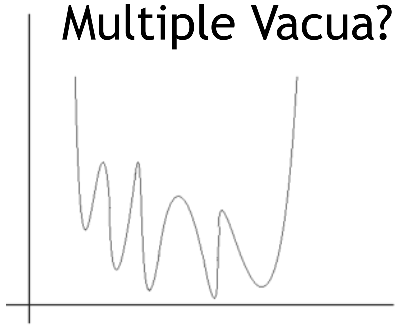
In either case the question is
“What is the basic phenomenology of such models?”

Qualitative Features of Models of Protein Design?

Frustration

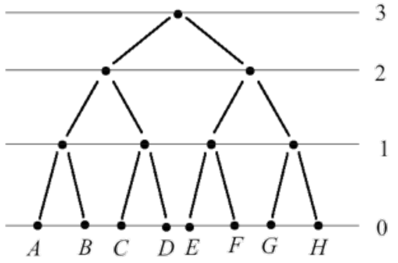


Multiple Vacua?



and

Ultrametricity?



Replicas and RSB?

Ergodicity Breaking?

Continuous Sequence of Phase Transitions?

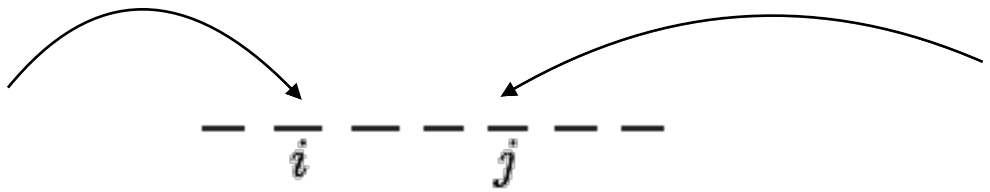
Extension: Non-Trivial Interactions

Mean Field Theory: In the first pass at studying an interacting model, we looked at an **“all-indices-coupled-to-all-indices”** model (Analogous to the mean field Ising model)

$$\mathcal{H}_N(\{\theta_i\}) = \frac{\lambda_2}{2N}\sum_{i,j=1}^{N} I_{\theta_i\neq\omega_i} I_{\theta_j\neq\omega_j} + \cdots$$

Matrix Model: To make progress beyond this point, we can consider **site-dependent interactions** between indices.

$$\mathcal{H}_N(\{\theta_i\}) = \frac{1}{2}\sum_{i,j=1}^{N} \mu_{ij} I_{\theta_i\neq\omega_i} I_{\theta_j\neq\omega_j} + \cdots$$



Result is proportional to matrix element μ_{ij}

The energy penalty for incorrectly occupying site *i*

Depends on whether *j* is correctly occupied

Possibly Soluble Model: **Nearest Neighbor Interaction** $\mu_{ij} = 2\mu\,\delta_{i,j+1}$ (similar to 1D Ising Model)

Possible Ways to Compute The Partition Function?

1. Transfer Matrices? $Z_N \overset{?}{=} \mathrm{Tr}\left[P_1 P_2 \cdots P_N\right]$

2. Write in terms of a basis of energy values? $Z_N \overset{?}{=} \sum_{\ell=0}^{N} h_N(\ell) e^{-\beta\mu\ell}$ where $\ell \equiv \sum_{i=1}^{N-1} I_{i+1} I_i$

Extension: Kinetic Models

$Z_N(\{\beta\lambda_i\})$
Partition Function

$k_B T_c$
Transition Temperature

$\bar{j}_0 = 1 - W_0(\cdots)$
“Order Parameter”

So far we’ve only been concerned with equilibrium physics.

How would we formulate the non-equilibrium physics of this system?

If we begin with an arbitrary ordering, how does that ordering evolve in time?

$$\underline{\omega_2}\quad \underline{\omega_N}\quad \cdots\quad \underline{\omega_1}\quad \overset{\text{a time } t \text{ passes}}{\Longrightarrow}\quad \underline{?}\quad \underline{?}\quad \cdots\quad \underline{?}$$

Naive Approach: Use the Master Equation (In analogy to the Kinetic Ising Model)

$$\frac{\partial}{\partial t} P_N(j,t) = \sum_{j'=0}^{N} P_N(j',t)\,w(j' \to j) - \sum_{j'=0}^{N} P_N(j,t)\,w(j \to j').$$

$(j = \text{number of incorrect components})$

<u>Example Dynamics</u>
In each time step, we exchange two components.

$$\theta_i(t_1 + \Delta t) = \theta_k(t_1) \text{ and } \theta_k(t_1 + \Delta t) = \theta_i(t_1)$$

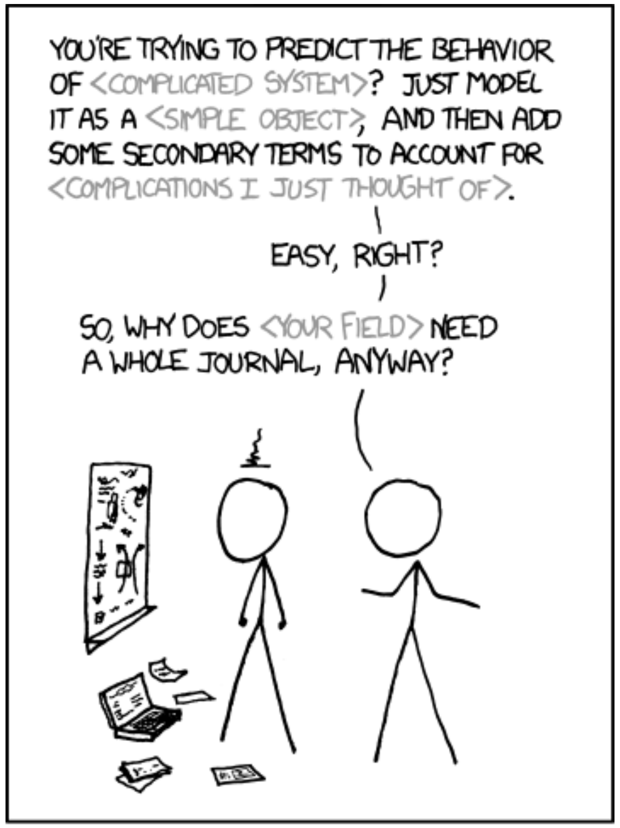
— Resulting *j* will increase/decrease or stay the same

$$j \to j\,,\, j \pm 1\,,\, j \pm 2,$$

$\longrightarrow$ What is $\langle j(t) \rangle$?

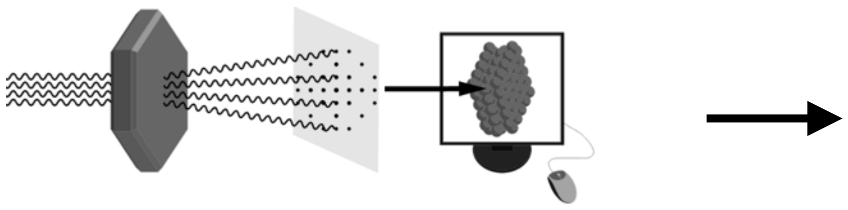
Bucking the Trend

https://imgs.xkcd.com/comics/physicists.png

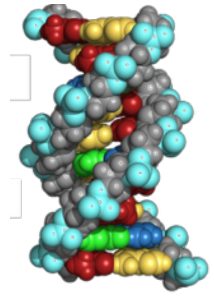


LIBERAL-ARTS MAJORS MAY BE ANNOYING SOMETIMES, BUT THERE'S *NOTHING* MORE OBNOXIOUS THAN A PHYSICIST FIRST ENCOUNTERING A NEW SUBJECT.

Physics can help biology



X-Ray Crystallography



DNA Structure

But can biology help physics?

Some biological problem or idea		Something interesting to physicists
Protein Design (with fixed a.a. composition)	→	Statistical Physics of the Symmetric Group

END