# Part I: Some Ergodic Theory

I hope this is all the ergodic theory we'll need. Let $\nu_p$ be product measure on $\{0,1\}^{Z^d} \equiv X$ where the 1's have probability p, i.e.,

$$\nu_p = \prod_{x \in Z^d} ((1-p)\,\delta_0 + p\,\delta_1).$$

In other words, we have independent and identically distributed r.v. indexed by $Z^d$ where each r.v. is 1 with probability p and 0 with probability 1-p.

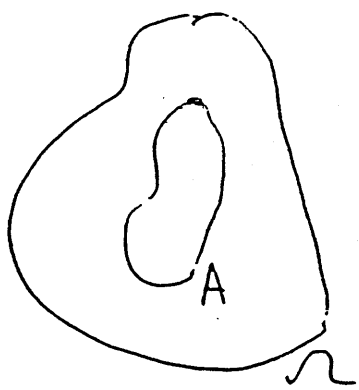
N.B. There are d canonical transformations $T_1 \dots T_d$ on X, the i:th such moving a configuration -1 unit in the i:th direction.

Theorem: Let $A \subseteq \{0,1\}^Z$ be such that $T_1 A = A$. Then A has measure 0 or 1 i.e. $\nu_p(A) = 0$ or 1.

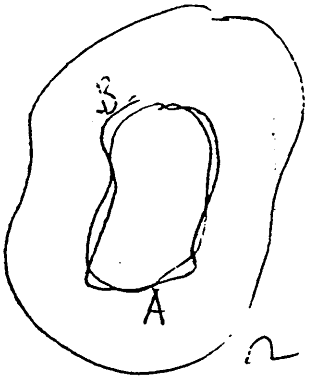
An example of such a A is the set of elements in $\{0,1\}^{Z^d}$ which have infinitely many 1's surrounded by 0's.

Proof of Theorem: We give a "picture proof" and note that such a picture proof can be used to prove Kolmogorov's 0-1 law and the Hewitt-Savage 0-1 law.

If A does not have measure 0 or 1, then it looks like this.

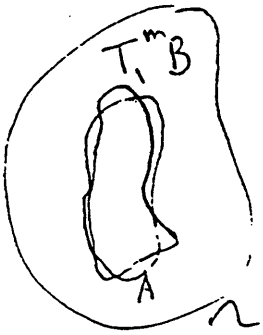



Basic facts from measure theory tell us that ∃ a cylinder set $B \subseteq X$ (i.e. B is of the form $\{\eta\colon \eta(x_i) = j_i,\ i = 1,\dots k\}$ where $x_1,\dots x_k \in Z^d$ $j_1,\dots, j_k \in \{0,1\}$) $\ni \nu_p(A \Delta B)$ is small.

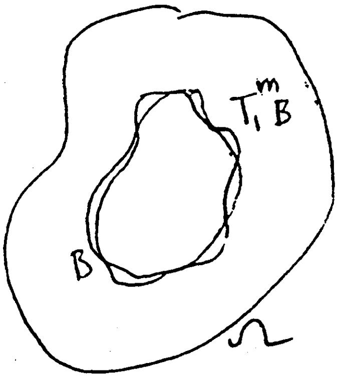



Now choose a high enough power of $T_1$, say $T_1^m$ so that the coordinates determining $T_1^m(B)$ are distinct from those of B. Then $T_1^mB$ and B are independent.

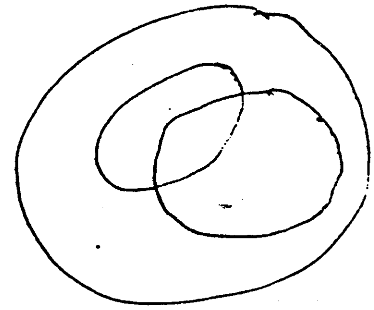
Since $T_1$ is measure preserving (i.e. $\nu_p(T_1C) = \nu_p(C)\ \forall\ C \subseteq X$), $\nu_p(T_1^m(A\Delta B))$ is small $\Rightarrow \nu_p(T_1^mA \,\Delta\, T_1^mB)$ is small $\Rightarrow \nu_p(A\Delta T_1^mB)$ is small as A is invariant. Hence




Since B and $T_1^mB$ are both close to A we also have




i.e. B and $T_1^mB$ are close. But independent sets do not look like this, they look like



That's the proof. You can fill in the $\epsilon$'s + $\delta$'s if you want. QED

We now state a special case of the "ergodic theorem", a very important and useful theorem.

Theorem: Consider $\nu_p$ as above. Let $A \subseteq X$. Then for $\nu_p$ a.e. $\omega \in X$,

$$\frac{1}{n^d} \left| (i_1, i_2, .. i_d) \in [0, n-1]^d: T_1^{i_1} T_2^{i_2} \ldots T_d^{i_d}(\omega) \in A \right| \to \nu_p(A).$$

In 1-dimension, we have only $T_1$ which we call $T$, which simply shifts the configuration over to the left one unit. Then $\frac{1}{n} \sum_{i=0}^{n-1} I_{\{T^i \omega \in A\}} \longrightarrow P(A)$ $\nu_p$ a.s.

Corollary: Strong law of large #'s.

Proof: Let $A = \{\eta: \eta(0) = 1\}$. Then

$$I_{\{T^i \omega \in A\}} = \begin{cases} 1 & \text{if } \omega_i = 1 \\ 0 & \text{if } \omega_i = 0 \end{cases}.$$

Hence $\frac{1}{n} \sum_0^{n-1} \omega_i = \frac{1}{n} \sum_{i=0}^{n-1} I_{\{T^i \omega \in A\}} \to P(A)$ QED

Some more ergodic theory background. (related to the existence of stationary distribution for Markov processes).

Theorem. Let $T$ be a continuous transformation on a compact metric space $X$. Then $\exists$ a T-invariant prob. measure $\mu$ on $X$, i.e. a measure $\mu \ni \mu(T^{-1}A) = \mu(A)$ $\forall A$.

Background on weak convergence of measures (generalizes convergence in distribution of r.v.)

Let $X$ be a compact metric space.

Definition. $\mu_n \longrightarrow \mu$ weakly if $\int f \, du_n \xrightarrow{n \to \infty} \int f d\mu$ $\forall$ continuous $f$

**Theorem.** If $\{\mu_n\}$ is a sequence of pm on X, then $\exists$ a subsequence $\mu_{n_k}$ that converges to some $\mu$.

"In words, the set of probability measures on a compact metric space is weakly compact." No proof but you should know this. (It's false for R).

**Proof.** Choose $x \in X$. Let $\mu_n = \frac{1}{n}\sum_{i=0}^{n-1} \delta_{T^i x}$. Let $\mu_{n_k}$ be a subsequence converging to some $\mu$.

We need to show

$$\int f \circ T d\mu = \int f d\mu \qquad \forall \text{ cont. f.}$$

$$|\int f \circ T d\mu - \int f d\mu| \leq |\int f \circ T d\mu_{n_k} - \int f d\mu_{n_k}| + \epsilon$$

$$= |\frac{1}{n_k}\sum_{i=0}^{n_k - 1} fT(T^i x) - \frac{1}{n_k}\sum_{i=0}^{n_k - 1} f(T^i x)| + \epsilon$$

$$\leq \frac{1}{n_k} |f(T^{n_k}(x) - f(x)| + \epsilon \to 0 \qquad \text{QED}$$

**Theorem.** Consider a discrete time Feller Markov process (MP) with compact metric state space X. Let $T\mu$ denote the distribution of the process at time 1 if the initial distiribution is $\mu$. Then any subsequential limit of the sequence

$$\frac{1}{n}\sum_{i=0}^{n-1} T^i \mu \text{ is a stationary distribution (S.D.)}$$

(In particular Stationary distributions exist).

NB: Consider a continuous time feller MP with compact State space X. Let $T^t\mu$ denote the distribution at time t when the initial distribution is $\mu$. Then any subsequential limit of $\frac{1}{T}\int_0^T T^t\mu \, dt$ is a stationary distribution.

# Part II: Percolation

1. Bond percolation.

Define 1) finite graphs 2) connectedness (components)

3)infinite graphs. 4) connectedness (components) 5) subgraph

Perc. is the study of random graphs.

Define the model. Start with $Z^2$. We will construct a random subgraph of $Z^2$ as follows: Fix $p \in (0,1)$. Let each edge be open with probability $p$ and be closed with probability $1-p$.

What is the probability $\theta(p)$ that $0$ can reach infinitely many vertices along the open bonds?

NB: If no edges are kept, this does not happen.

If all edges are kept, this does not happen.

So this might or might occur - it is random.

More rigorously. Let $E$ denote the edge set of $Z^2$. Let $\Omega = \prod_{e \in E} \{0,1\}$. $\mathcal{F} = \sigma$-algebra generated by the cylinder sets. Let $P_p = \prod_{e \in E} \mu_p$ where $\mu_p$ is the measure on $\{0,1\}$, $\mu_p(1) = p$, $\mu_p(0) = 1-p$.

Let $C(x)$ denote the cluster (component) containing $x$: (depends on $\omega$). $C = C(0)$.

What is $P_p(|C| = \infty)$? (Same as $\theta(p)$).

Easy Problem: For any sg of $Z^2$, $|C| = \infty$, iff $\exists\ 0 = x_0, x_1, x_2, \ldots$ all distinct $\ni \{x_i, x_{i+1}\} \in E \ \forall i$.

Exercise. $\theta(p)$ is nondecreasing.

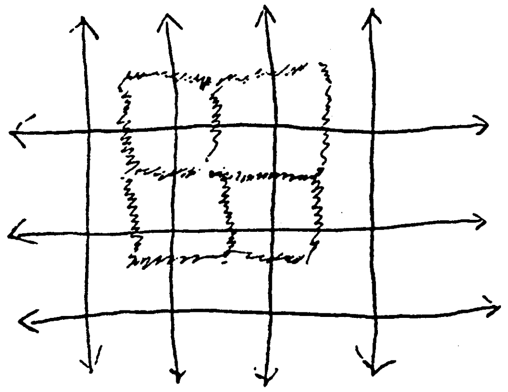
Theorem 1. If $p < 1/3$, $\theta(p) = 0$.

Let $F_n$ be the event there is a path of length $n$ starting at $0$. For any such path in $Z^2$, the prob that it is a path in the random graph is $p^n$. The number of such paths is $\leq 4 \cdot 3^{n-1} \Rightarrow \text{Prob}(F_n) \leq 4 \cdot 3^{n-1} p^n \to 0$ as $n \to \infty$. Since $\{|C| = \infty\} \subseteq F_n \ \forall_n$, Prob $\{|C| = \infty\} = 0$, i.e. $\theta(p) = 0$ and no percolation. QED

**Theorem 2.** If $\sum_{n=4}^{\infty} n 4 3^{n-1}(1-p)^n < 1$, then $\theta(p) > 0$ (so the model is interesting).

NB: $\theta(p)$ cannot be 1 for $p < 1$.

**Proof.** Introduce dual graph

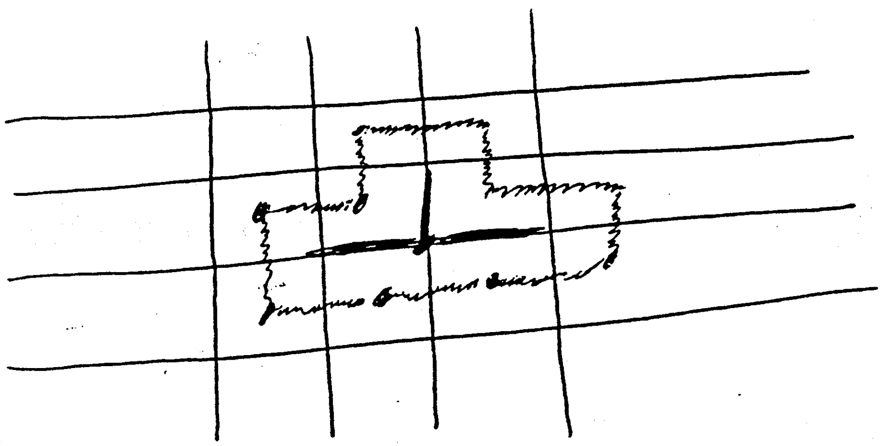


$V^* = Z^2 + (\frac{1}{2}, \frac{1}{2})$, and the edge set is $E + (\frac{1}{2}, \frac{1}{2})$. There is an obvious 1-1 correspondence between the edges of V and those of $V^*$. For any subgraph of $Z^2$, we get a subgraph of $V^*$ by calling an edge open if and only if the corresponding edge is open.

NB: If a subgraph of $Z^2$ is chosen according to $P_p$, then the corresponding subgraph of $V^*$ also has distribution $P_p$.

**Lemma 1.** Whitney (pure graph theory, no probability.)

$|C| < \infty$ iff $\exists$ a simple closed path in $V^*$ surrounding 0 consisting of all <u>closed</u> edges.



$P(|C| < \infty) = \text{Prob}$ ($\exists$ a simple closed path in $V^*$ surrounding 0 consisting of all closed edges) $\leq \sum_{n=4}^{\infty} n 4 3^{n-1}(1-p)^n$.

Since the sum is $< \infty$ if $p > \frac{2}{3}$, it can be made arbitrarily small if $p$ is close to $1$. In particular, the sum can be made $< 1 \Rightarrow P(|C| = \infty) > 0$ for $p$ close to $1$. QED

NB: We showed $\theta(p) \to 1$ as $p \to 1$.

NB: This argument is called a contour argument and variants of it are used often.

NB: The above proof can be modified to give $\theta(p) > 0$ for $p > \frac{2}{3}$.

**Proof.** Fix $p > \frac{2}{3}$ and choose $N \ni \sum_{n \geq N}^{\infty} n 4 3^{n-1} (1-p)^n < 1$.

Let $E_1$ be the event that all bonds are open in $[-N,N] \times [-N,N]$

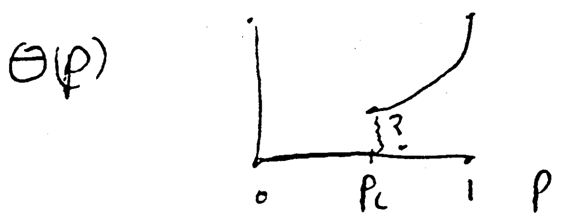
Let $E_2$ be the event that there are no simple closed loops in $V^*$ surrounding $[-N,N]$ consisting of closed edges.

1) $\{|C| = \infty\} \supseteq E_1 \cap E_2$ 2) $E_1, E_2$ independent

3) $P(E_1) > 0$ trivially while $P(E_2) > 0$ by choice of $N \Rightarrow P(|C| = \infty) > 0$. QED

It is natural to define the critical value $p_c$ by

$$p_c = \sup\{p: \theta(p) = 0\} = \inf\{p: \theta(p) > 0\}$$




What is $p_c$? Is $\theta(p)$ continuous?

**Theorem 3.** $p_c = 1/2$.

**Theorem 4.** $\theta(p)$ is cont on $[0,1]$.

**Theorem 5.** $\theta(p)$ is inf. diff. on $[\frac{1}{2}, 1]$.

Open Question 1: Is $\theta(p)$ analytic in $(\frac{1}{2}, 1]$. NB: obviously not at 1/2.

Everything trivially generalizes to $Z^d$.

NB: It is clear that $\theta_{d+1}(p) \geq \theta_d(p)$ $\forall p \ \forall d$

$$\Rightarrow p_c(d+1) \le p_c(d).$$

NB: If $\theta(p) = 0$, then $P_p$ (some $C(x)$ is infinite) $= 0$ since $P_p(\bigcup_x \{|C(x)| = \infty\}) \le \sum_x P_p(|C(x)| = \infty) = 0$.

**Theorem 6.** If $\theta(p) > 0$, then $P_p$ (some $C(x)$ is infinite) $= 1$.

**Proof.** Let $F$ be the event that some $C(x)$ is infinite. $F$ is $Z^d$ invariant $\Rightarrow P(F) = 0$ or 1. Since $P(|C(0)| = \infty) > 0$, $P(F) > 0 \Rightarrow P(F) = 1$. QED

**Theorem 7.** $\theta_d(p)$ is right continuous on $[0,1]$.

**Theorem 8.** $\theta_d(p)$ is continuous on $(p_c(d), 1]$. (Vanden Berg-Keane)

Motivates the Question.

Is there a jump at $p_c(d)$ or is $\theta_d(p_c(d)) = 0$.

Conjecture 2: $\theta_d(p)$ is continuous at $p_c$ (i.e. no percolation at the critical value).

**Theorem 9.** $\theta_d(p)$ is continuous at $p_c(d)$ $\forall d \ge 48$.

(so called mean field behaviour).

**Open Problem 3:** Find $p_c(d)$. NB: probably impossible.

**Proof of Theorem 7.** Let $g_n(p) = P_p$ (there is a path of length $n$ starting from the origin.) $g_n(p)$ is a polynomial in $p$ and $g_n(p) \downarrow \theta(p)$ as $n \to \infty$.

Now a decreasing limit of continuous functions is always upper semi-continuous and a nondecreasing upper semi-continuous function is right continuous. QED

Before proving Theorem 8, we need to enter the next area. Uniqueness of the infinite cluster.

Switch to site percolation. Methods are the same but conceptually easier to apply our developed ergodic theory. Now sites are on or off (all bonds are on)

$p_c^s(2) > p_c(2)$ but its exact value is unknown. ($s$ stands for site).

Theorem 9. Theorem 6 holds verbatum.

So if $p < p_c$, no infinite cluster.
if $p > p_c$, at least one infinite cluster.

Natural Question: How many infinite clusters.

NB: There are realizations with any number of clusters.

Theorem 10. If $\theta(p) > 0$, $P_p$ ($\exists$ a unique infinite cluster) = 1.

NB: First proofs were quite long and difficult. A number of papers have been written. Finally the canonical proof by Burton and Keane was obtained which we will do.

## Warm UP

The number of infinite clusters is nonrandom, i.e. it is constant a.s.

Let $E_k = \{\#$ of inf. cluster is $k\}$.

$E_k$ is $Z^d$-invariant $\Rightarrow$ Prob $(E_k) = 0$ or 1.

Since $\Omega = \bigcup_{k=0}^{\infty} E_k$, some $E_k$ has probability 1 $\Rightarrow$ # of infinite clusters is a.s. that value of k.

Here is a good exercise in applying the ergodic theorem.

Theorem 11.
$$\frac{|x \in [0,n-1]^d: \ x \text{ belongs to an infinite cluster}|}{n^d} \longrightarrow \theta(p) \quad \text{as } n \to \infty \quad P_p\text{- a.s.}$$

Warm up Theorem due to Newman and Schulman.

Theorem 12. The number of infinite clusters is either 0 a.s. 1 a.s. or $\infty$ a.s.

NB: We rule out that the number of cluster is > 1 but finite.

Proof. We show P(# of infinite clusters = 5) = 0, other cases similar.

Assume P(# inf cluster = 5) > 0 and so 1.

Let E = {there are 5 inf cluster and each intersects $[-M,M]^d$}.

$E_1 \subseteq E_2 \subseteq \ldots \subseteq \cup E_i =$ {there are 5 infinite clusters}.

So $P(E_i) \xrightarrow{i\to\infty} 1$. Choose $N \ni P(E_N) > 0$.

Let $\tilde{E}_N = \{\omega$: $\omega$ can be changed on $[-N,N]^d$ to be in $E_N)$

$E_N \subseteq \tilde{E}_N \Rightarrow P(\tilde{E}_N) > 0$ and $\tilde{E}_N$ is measurable with respect to $\{\omega_i\}_{i \notin [-N,N]^d}$.

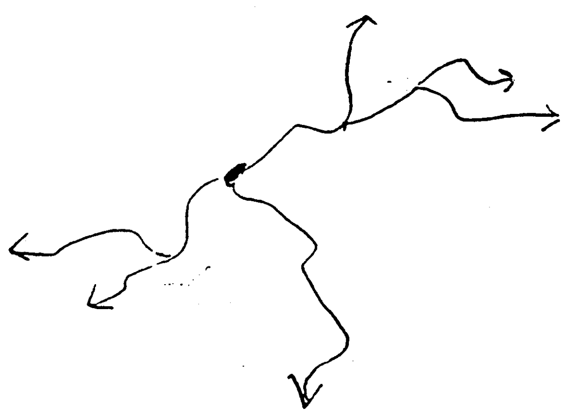
Let $F = \{\omega_i \equiv 1$ on $[-N,N]^d\} \Rightarrow P(F \cap \tilde{E}_N) > 0$ by independence.

$F \cap \tilde{E}_N \subseteq$ {there is 1 infinite cluster} $\Rightarrow$ P(1 infinite cluster) > 0 contradicting P(5 infinite clusters) = 1.

QED

It is much harder to rule out infinitely many infinite clusters, which we do now. This proof is due to Burton-Keane. Let N be the # of infinite clusters. Assume $P(N=\infty) = 1$ and we will get a contradiction. We call z an "encounter point" (for $\omega$) if

1. z belongs to an infinite cluster C.
2. $C \sim z$ has no finite components and exactly 3 infinite components.



Lemma. Prob 0 is an encounter point (ep) is > 0.

Proof. Let $E_M$ = {at least 3 infinite clusters intersect $[-M,M]^d$}

Since $E_1 \subsetneq E_2 \subsetneq \ldots$ and $\cup E_i = \{\exists \geq 3 \text{ infinite clusters}\}$.

$P(E_i) \xrightarrow{i \to \infty} 1$ and so we can choose $N \ni P(E_N) > 0$.

Let $\tilde{E}_N = \{\omega: \omega$ can be changed in $[-N,N]^d$ to be in $E_N\}$

$E_N \subsetneq \tilde{E}_N \Rightarrow P(\tilde{E}_N) > 0$ and $\tilde{E}_N$ is measurable with respect to $\{\omega_i\}_{i \notin [-N,N]^d}$.

Geometry and thinking tells us that if $\omega \in \tilde{E}_N$, then there is $\omega'$ which satisfies

1) $\omega' = \omega$ outside $[-N,N]^d$

2) $0$ is an encounter point for $\omega'$.

So, first we pick $\omega$ outside $[-N,N]^d$ and this is in $\tilde{E}_N$ with positive probability. Then using independence, with probability $> (\frac{1}{2N+1})^d$, we choose $\omega$ inside $[-N,N]^d$ so that we have an encounter point at 0. QED

Let $\delta = \text{Prob}$ (0 is an encounter point).

The ergodic theorem $\Rightarrow$

$$\lim_{N\to\infty} \frac{|x \in [-N,N]^d: \ x \text{ is an encounter point for } \omega|}{(2N+1)^d} = \delta \quad P_p \text{ a.s. } (\omega).$$

We now show that no configuration at all can have this property. This follows from the following lemma.

**Lemma**: For any configuration the number of encounter points in $[-N,N]^d$ is less than the number of boundary points of $[-N,N]^d$.

NB: At this point, there is no probability at all.

Proof. First note each "e.p" is associated to an infinite cluster. Let $C_1, C_2, \ldots, C_k$ be the infinite clusters that intersect $[-N,N]^d$. Let $Y_1, \ldots Y_k$ be defined by $Y_i = C_i \cap \partial[-N,N]^d$. Each "e.p" in $[-N,N]^d$ is associated to one of the $C_i$'s.

We show that the number of e.p. assoicated with $C_i$ is $\leq |Y_i|$ (actually we will show $\leq |Y_i| - 2$). Then total number of "e.p." in $[-N,N]^d \leq \sum_{i=1}^{k} |Y_i| \leq |\partial [-N,N]^d|$, as desired. Without loss of generality, we show that the number of e.p. for $C_1$ is $\leq |Y_1|$. From now on, all e.p. will be e.p's for $C_1$. Each e.p. by definition partitions $C_1$ into 3 pieces and so partitions $Y_1$ into 3 pieces, say $(P_1, P_2, P_3)$. If we take another e.p., it partitions $Y_1$ into $(Q_1, Q_2, Q_3)$. Note by drawing a picture, $Q_2 \cup Q_3 \subseteq P_1$ after relabeling. The lemma now follows from the following wierd combinational lemma. QED.

**Lemma.** Let $\rho_1, \rho_2, \ldots, \rho_n$ be a collection of partitions of a set $S$ where each $\rho_i$ partitions $S$ into 3 nonempty sets. Assume that they are compatible in the sense that if $(P_1, P_2, P_3)$ and $(Q_1, Q_2, Q_3)$ are 2 of the partitions, then after relabeling $Q_2 \cup Q_3 \subseteq P_1$. Then, $n \leq |S| - 2$.

**Proof.** This is not real hard but not real easy. It should be done by induction and is left to the reader. QED.

**Proof of Theorem 8.** (We are back in bond percolation.) Let $\tilde{p} > p_c$. We need to show $\lim_{\pi \nearrow \tilde{p}} \theta(\pi) = \theta(\tilde{p})$. The idea is to couple all percolations realizations (for various p) on the same probability space. This is easy. Let $\{X(e): e \in E^d\}$ be a collection of r.v. indexed by the bonds of $Z^d$ and having uniform distribution on $[0,1]$. We say $e \in E^d$ is p-open if $X(e) < p$.

NB1: Prob (e is p-open) $= p$ and these events are independent for different e's. Here the set of e's which are p-open is just a percolation realization with parameter p.

NB2: If $p_1 < p_2$, $\{e: e \text{ is } p_1 \text{ open}\} \subseteq \{e: e \text{ is } p_2 \text{ open}\}$. Now, let $C_p$ be the p-open cluster of the origin. Obvious $C_{p_1} \subseteq C_{p_2}$ if $p_1 < p_2$.

$$\theta(\tilde{p}) = P(|C_{\tilde{p}}| = \infty)$$

and

$$\lim_{\pi \nearrow \tilde{p}} \theta(\pi) = \lim_{\pi \nearrow \tilde{p}} P(|C_\pi| = \infty) = P(|C_\pi| = \infty \text{ for some } \pi < \tilde{p}).$$

Since $\{|C_\pi| = \infty \text{ for some } \pi < \tilde{p}\} \subsetneq \{|C_{\tilde{p}}| = \infty\}$, we need to show

$$|C_{\tilde{p}}| = \infty \Rightarrow |C_\pi| = \infty \text{ for some } \pi < \tilde{p} \text{ (a.s.)}.$$

Let $\alpha$ be such that $p_c < \alpha < \tilde{p}$. Then there is an infinite $\alpha$-open cluster $I_\alpha$ (not necessarily containing the origin).

Now, if $|C_{\tilde{p}}| = \infty$, then $I_\alpha \subsetneq C_{\tilde{p}}$ a.s. since there is a unique infinite cluster a.s. If $0 \in I_\alpha$, we are of course done with $\pi = \alpha$. Otherwise, there is a p-open path $\ell$ from the origin to $I_\alpha$. Let $\mu = \max\{X(e): e \in \ell\}$ which is $< p$.

Now, if $\mu, \alpha < \pi < p$, then there is a $\pi$ open path from 0 to $I_\alpha$ and $|C_\pi| = \infty$, as we wanted to show. QED

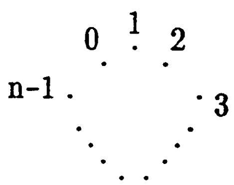
# Part III: Interacting Particle Systems

These are continuous time Markov processes on $\{0,1\}^{Z^d}$. For countable state Markov processes, irreducibility $\Rightarrow$ at most 1 s.d. This is false for $\{0,1\}^{Z^d}$ (nb: uncountable) and is what causes the richness of the theory.

## Contact Process

Start with finite state space.

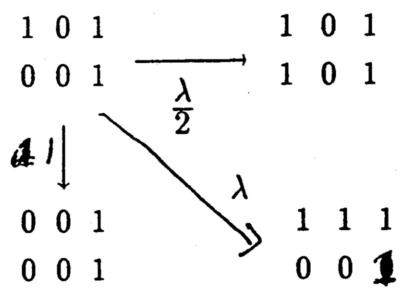
$$X = \{0,1\}^{Z/n}$$



Dynamics: Each 1 waits an exp (1) random time and becomes a 0.

Each 1 also waits an indep exp($\lambda$) amount of time and gives birth to a 1 and then places this 1 at random $(\frac{1}{2}, \frac{1}{2})$ at either of her 2 neighbors. If there is already a 1 there, the birth is suppressed (not counted).

The dynamics look like




if $\eta$ and $\delta$ differ at more than 1 lattice point, $q(\eta,\delta) = 0$.

If $\eta$ and $\delta$ differ only at x with $\eta(x) = 1$, $\delta(x) = 0$, then $q(\eta,\delta) = 1$.

If $\eta$ and $\delta$ differ at x with $\eta(x) = 0$, $\delta(x) = 1$,

$$q(\eta,\delta) = \lambda \left[\frac{\eta(x-1) + \eta(x+1)}{2}\right]$$

NB: all 0's is absorbing and everything can reach all 0's. Hence $\eta_t \to \underline{0}$ with probability 1 and the only S.D. is the trivial one, $\delta_{\underline{0}}$.

NB: true $\forall n$ and $\forall\lambda$.

Now we do this on $\{0,1\}^Z$. We think of $\lambda$ as a parameter just as p in the percolation model.

Question 1: Does $\exists\lambda$ sufficiently large $\ni$

$\text{Prob}_\lambda \{\eta_t \neq \underline{0}\ \forall t$ starting with only one 1 at 0) > 0.

Question 2: Does $\exists\lambda$ sufficiently large $\ni$

there exists a nontrivial s.d; i.e. $\neq \delta_{\underline{0}}$.

It is interesting to compare to a branching process. Each individual is dying at rate 1 and giving birth at rate $\lambda$. BUT sometimes the birth is suppressed. If the birth was not suppressed, then if $\lambda > 1$,

$$\text{Prob}(\eta_t \neq 0\ \forall t \text{ starting with only one 1 at 0}) > 0.$$

(This follows immediately from basic branching process theory).

So there is complicated geometrical structure involved. Since the branching process dies out if $\lambda \leq 1$, the above proves (using an elementary coupling of a Branching Process with the total # of 1's in the contact process) that if $\lambda \leq 1$,

$$P(\eta_t \neq 0\ \ \forall t \text{ starting with only one 1 at 0}) = 0.$$

(We call this event "survival" some times)

Also, this probability is nondecreasing in $\lambda$ and this is easy to show.

Theorem 1. If we let $\lambda_c = \sup\{\lambda\text{: } P(\text{survival}) = 0\}$, then $\lambda_c < \infty$, i.e. for sufficiently large $\lambda$, $P(\text{survival}) > 0$.

NB: $P(\text{survival}) < 1$ always.

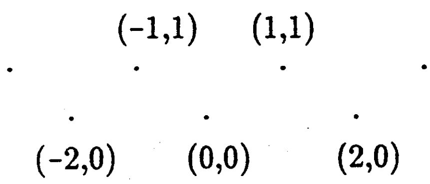
Theorem 2. $P_\lambda\,(\text{survival}) > 0 \Leftrightarrow \exists$ a nontrivial S.D. at $\lambda$.

NB: Theorem 2 is not true for general particle systems but holds for the contact process due to a self-duality property of the contact process.

Method of Proof: "Show the contact process for large $\lambda$ dominates ORIENTED Percolation".

## ORIENTED PERC

Let $G = \{(x,y) \in Z^2\text{: } x+y \text{ is even}\}$




As in percolation, each point is designated to be open with prob $p$ and closed with prob $(1-p)$, all independently.

**Definition:** $x \to y$ ($y$ can be reached from $x$), if there is an open path from $x$ to $y$, that is, there is a sequence $x_0 = x, x_1, x_2, \ldots, x_j = y$ in $G \ni$
each $x_i$ is open and

$$x_i = x_{i-1} + (1,1) \text{ or } x_{i-1} + (-1,1) \quad \text{for } i=1,\ldots j.$$

Let $C_{(0,0)} = \{x\text{: } (0,0) \to x\}$.

**Theorem:** If $p > 1 - 3^{-36}$, then $P(|C_{(0,0)}| = \infty) > 0$.
We will not give the proof. It is similar to the contour argument given for percolation but the geometry a little messier. See Durrett's book. We now show that for large $\lambda$, the contact process survives (with positive prob.)

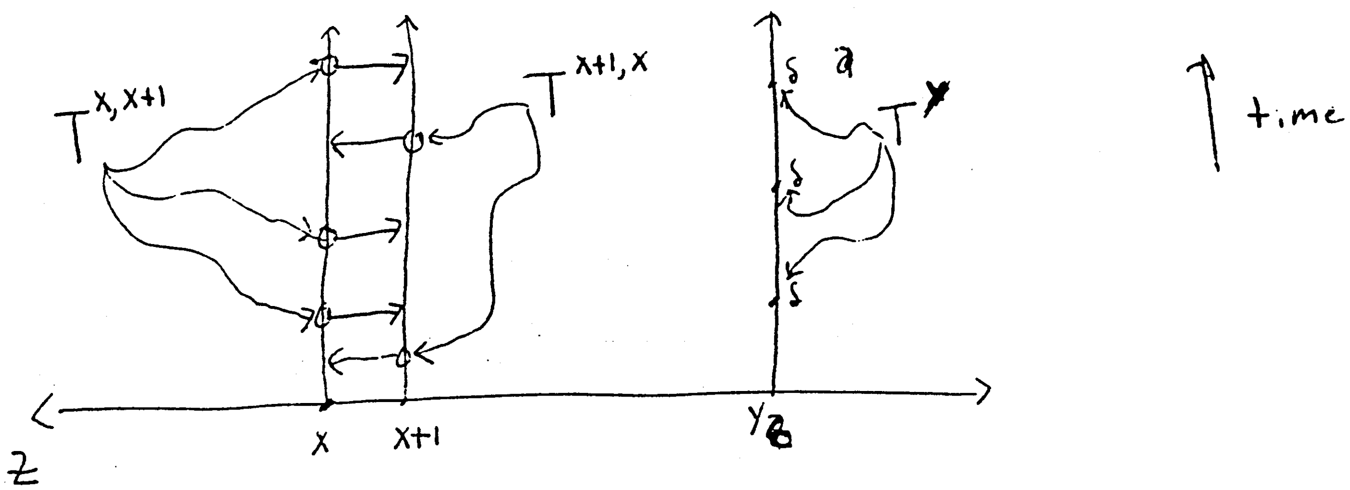
Before doing this, we introduce the "Graphical Representation". This is an extremely important tool for analyzing particle systems.

$$\forall x \in Z, \quad \forall y \in Z \text{ with } |y-x| = 1$$

let $T^{x,y}$ be a Poisson processes with rate $\frac{\lambda}{2}$ and let all these Poisson processes for different (x,y) be independent.

"The points of the Poisson process $T^{x,y}$ will be the "random times" at which x gives birth onto y".

We place around from x to y at these random times.




We also have $\forall x \in Z$, independent Poisson Processes $T^x$ with rate 1.

"These correspond to the times at which x dies i.e. a 1 switches to 0". We place $\delta$'s at these points.

**Definition:** There is a path from (x,0) to (y,t) if there is a sequence of times $s_0 = 0 < s_1 < s_2 < \ldots < s_n < s_{n+1} = t$ and spacial locations $x_0 = x$, $x_1, x_2 \ldots x_n = y$ so that

1) for $i = 1,\ldots n$ there is an arrow from $x_{i-1}$ to $x_i$ at time $s_i$.

2) the vertical segments $\{x_i\} \times (s_i, s_{i+1})$ $i = 0,\ldots,n$ contains no $\delta$'s.

i.e. if there is a path from (x,0) to (y,t) moving up vertical lines and across arrows never going through a $\delta$.

We identify subsets of Z with elements of $\{0,1\}^Z$ by

if $\eta \in \{0,1\}^Z$, let $A = \{x \in Z\colon \eta(x) = 1\}$.

The "formal" definition of our process is as follows. Let $A \subseteq Z$ be our initial configuration. Let

$$\eta_t^A = \{y: \text{ for some } x \in A, \text{ there is a path from } (x,0) \text{ to } (y,t)\}.$$

NB: One must think for a few minutes to see that $\{\eta_t^A\}$ is what we want it to be.

NB: Since X is uncountable, there is no Q-matrix anymore. One can generalize Q-matrices to this setting but this involves fairly complicated functional analysis. The "graph rep." allows us to define our process without running into such technicalities.

Theorem 1 can now be stated as follows.

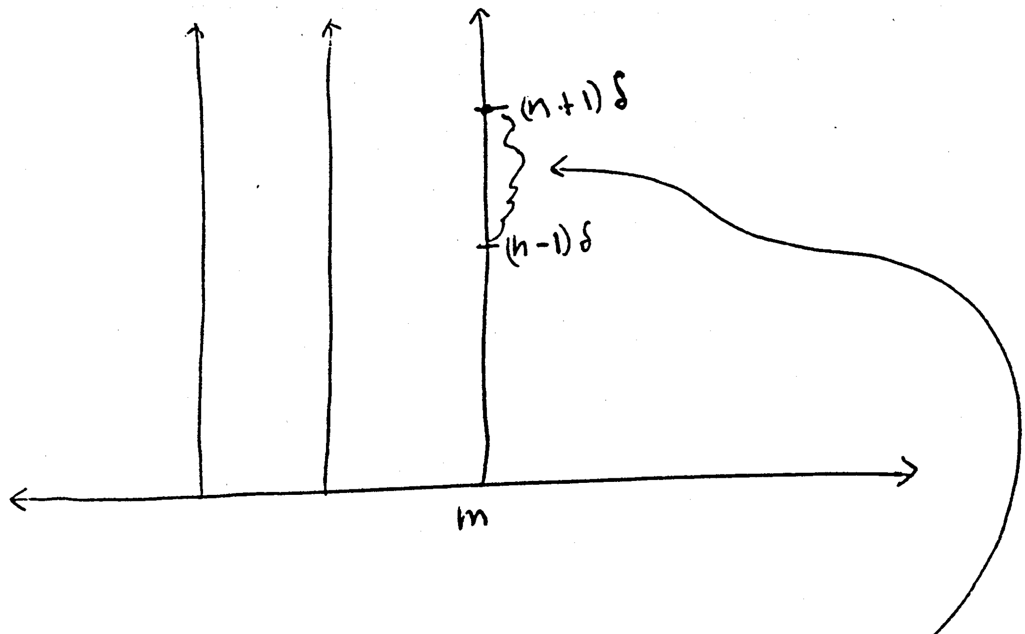
For $\lambda$ large,

$$P_\lambda(\eta_t^{\{0\}} \neq \emptyset \quad \forall t) > 0.$$

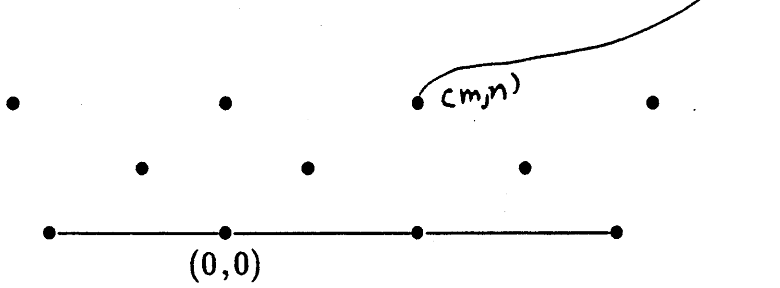
(By above, $\eta_t^{\{0\}}$ means the contact process starting with only one 1 at 0).

**Proof.** Choose $\delta$ so small that the prob that a $1 \to 0$ in time $2\delta$ is $< \frac{1}{2} 3^{-36}$ (i.e. $\delta$ is so small that the prob an exponential r.v. with parameter 1 takes a value $< 2\delta$ is $< \frac{1}{2} 3^{-36}$).

Next, choose $\lambda$ so big that the prob a point gives birth to both its neighbors in $\delta$ units of time is $> 1 - \frac{1}{2} 3^{-36}$ (i.e. $\lambda$ is so large that if we take 2 independent exponential r.v. with parameter $\lambda$, then with Prob $> 1 - \frac{1}{2} 3^{-36}$ each will take a value $< \delta$). We claim that this $\lambda$ suffices for survival (having positive probability) of the contact process. The next step is to relate OP (oriented percolation) with the contact process.




From our arrow realization we define an OP realization as follows.




(m,n) is on iff

certain "good things" occur in the graphical representation at site m between times $(n-1)\delta$ and $(n+1)\delta$. Namely

1) no $\delta$'s on this vertical segment (we have 2 $\delta$'s so be careful one was chosen small, the other is used to indicate deaths)

2) on the top half of the segment between $n\delta$ and $(n+1)\delta$ there are arrows from m to both m−1 and m+1 (when n=0, the obvious slight modification should be made).

We claim the following

1) This yields an OP realization with $p > 1 - 3^{-36}$.

2) If in this OP realization, the cluster containing the origin is infinite, then

$$\eta_t^{\{0\}} \neq \emptyset \qquad \forall t.$$

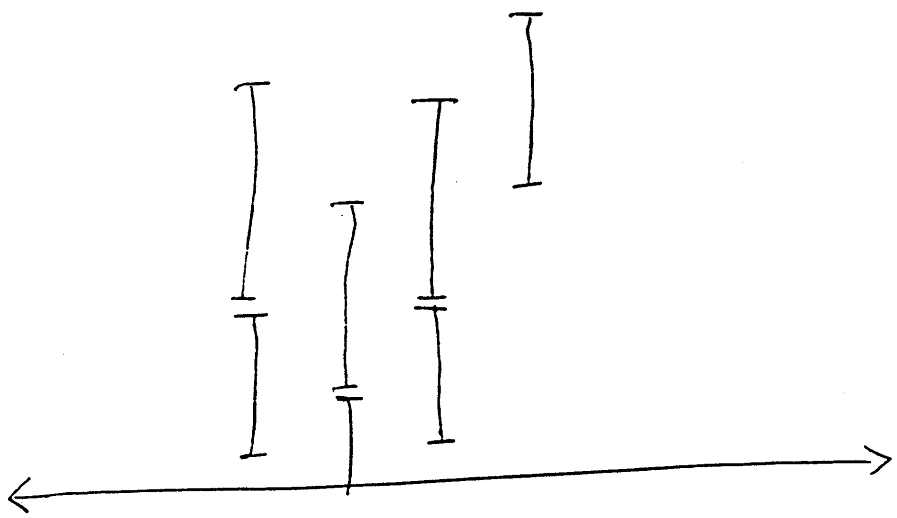
The theorem then immediately follows since by our OP theorem together with

$p > 1 - 3^{-36} \Rightarrow$ the cluster containing the origin is infinite with positive prob.

$\Rightarrow$ (by 2) $\qquad \eta_t^{\{0\}} \neq \emptyset \ \forall t$ with positive prob.

1) The fact that the (m,n)'s being on or off are independent follows from the fact that they use different parts of the graphical representation. The fact that $p > 1 - 3^{-36}$ follows from the definition of $\lambda$ and $\delta$.

2) follows from a picture and a little thought.



QED

## Self Duality

The main application of the self-duality equation is

**Theorem:** $P^{\lambda}\{\eta_t^{\{0\}} \neq 0 \ \forall t\} > 0 \Leftrightarrow \exists$ a nontrivial startionary dist (i.e. $\neq \delta_{\underline{0}}$) for $\lambda$.

Let $\eta_t^A$ be the process starting rom A occupied.

Duality Equation: $\qquad P\{\eta_t^A \cap B \neq \emptyset\} = P\{\eta_t^B \cap A \neq \emptyset\}$.

The Duality Equation also has a picture proof. (Think about it.)

**Definition:** If $\eta_t$ is any process taking values in X, we say $\eta_t \to \mu$ in distribution if

$$\forall x_1,\ldots,x_k \in Z \qquad \forall j_1,\ldots,j_k \in \{0,1\},$$
$$P(\eta_t(x_i) = j_i \text{ for } i = 1,\ldots,k\} \xrightarrow{t\to\infty} \mu(\eta\colon \eta(x_i) = j_i \text{ for } i=1,\ldots,k\}.$$

In words, the finite dimensional distributions converge.

NB: This does <u>not</u> imply $P(\eta_t \in A) \to \mu(A) \quad \forall A \subseteq X$.

**Theorem** : $\xi_t^Z \Rightarrow \bar{\nu}$ in distribution ($\Rightarrow \bar{\nu}$ is a stationary distribution).

**Proof.** $\forall A \subseteq Z$, $\forall t$

$$P\{\eta_t^A \neq \emptyset\} = P(\{\eta_t^Z \cap A \neq \emptyset\}$$

by setting $B = Z$ in the duality equation. The LHS is clearly $\downarrow$ in $t$ and so

$$P\{\eta_t^Z \cap A \neq \emptyset\} \xrightarrow[t\to\infty]{} \text{some limit.}$$

Since every event of the form $\{\eta_t^Z(x_1) = j_i \;\; i=1,\ldots,k\}$ can be expressed in terms of $\{\eta_t^Z \cap A \neq \emptyset\}$ for various A's, the theorem is proved. QED

Setting $A = \{0\}$, $B = Z$ in the duality equation gives

$$P\{\eta_t^{\{0\}} \neq \emptyset\} = P\{0 \in \eta_t^Z\}.$$

Letting $t \to \infty$ gives

$$P\{\eta_t^{\{0\}} \neq \emptyset \;\; \forall t\} = \bar{\nu}\,\{\eta\colon \eta(0) = 1\}.$$

If LHS $> 0$, then $\bar{\nu}$ is a nontrivial S.D. If $\bar{\nu}$ is a nontrivial S.D., then RHS $> 0 \Rightarrow$ LHS $> 0$.

This shows the constant process survives $\Leftrightarrow$ $\bar{\nu}$ is a nontrivial distribution. To prove our main theorem, we need to show that if $\bar{\nu}$ is trivial, then all stationary distributions are trivial. To do this, requires a detour into partial orders on measures.

Recall a partially ordered space is a set with a relation called $\leq$ such that

1) $a \leq b,\ b \leq a \Rightarrow a = b$

2) $a \leq b,\ b \leq c \Rightarrow a \leq c.$

We need to generalize the usual notion of stochstic ordering of rv's. Recall if X,Y are r.v, (not necessarily defined on the same space.),

*) $X \leq Y \quad \text{if } F_X(t) \geq F_Y(t) \quad \forall t$

"X tends to be smaller than Y".

This is equivalent to the existence of 2 rv. $\tilde{X}, \tilde{Y}$ defined on the same space so that

1) $\tilde{X} \overset{\mathcal{D}}{=} X,\ \tilde{Y} \overset{\mathcal{D}}{=} Y$

2) $\tilde{X} \leq \tilde{Y}$

Note $\{0,1\}^S$, S finite or countable, has a natural partial order defined by

$$\eta \leq \delta \text{ if } \eta(x) \leq \delta(x) \quad \forall x \in S.$$

**Definition:** $f: \{0,1\}^S \to R$ is increasing if

$$\eta \leq \delta \Rightarrow f(\eta) \leq f(\delta).$$

**Definition:** If $\nu,\mu$ are in $P(\{0,1\}^S)$, $\nu \leq \mu$ if

$$\int f d\nu \leq \int f d\mu \qquad \forall \text{ increasing } f.$$

NB: This is in spirit the analogue of (*) and says "$\mu$ concentrates its mass on larger elements than $\nu$". It can be shown to be equivalent to

**Definition.** $\nu \leq \mu$ if $\exists$ variables $\eta$ and $\delta$ (defined on the same prob space) taking values in $\{0,1\}^S \ni$

1) The distribution of $\eta$ is $\nu$
2) The distribution of $\delta$ is $\mu$
3) $\eta \leq \delta$ a.s.

In other words, $\exists$ a measure m on $\{0,1\}^S \times \{0,1\}^S \ni$

1) the first marginal (projection) of m is $\nu$
2) the second marginal (projection) of m is $\mu$.
3) $\{(\eta,\delta): \eta \leq \delta\}$ has full m-measure. i.e. $m\{(\eta,\delta): \eta \leq \delta\} = 1$.

We say $\nu$ can be coupled below $\mu$.

Back to the contact process, it is believable that if $A \subseteq B$, then $\eta_t^A \leq \eta_t^B$ as measures. In fact the graph rep. proves this since if we use the same graphical representation for the 2 initial configurations A, B then in fact

$$\eta_t^A \subseteq \eta_t^B$$

(The point is that the graph representation allows us to couple all contact processes (with different initial distributions) on the same probability space.)

Finally, if $\bar{\nu}$ is trivial, then

$$\eta_t^Z \to \delta_0 \text{ and so } \forall A$$
$$\eta_t^A \to \delta_0 \text{ since } \delta_0 \leq \eta_t^A \leq \eta_t^Z.$$

(i.e. $\eta_t^A$ is trapped below $\eta_t^Z$).

Similarly if we start with an initial distribution $\mu$ (rather than a fixed configuration), we also have $\eta_t^\mu \to \delta_0 \Rightarrow$ no nontrivial S.D.

QED

**Exercise:** Show $\eta_t^Z$ is decreasing (in t) as a sequence of measures and also that it therefore converges reproving a previous theorem.

## Voter Model

Finite voter model, $X = \{0,1\}^{Z/n}$. Each $i \in Z/n$ waits an exp (1) random time, picks a neighbor at random and forces that person to conform to the value at i. (If they are already the same value, nothing changes).

In terms of the Q-matrix,

$q(\eta,\delta) = 0$ if $n,\delta$ differ in more than 1 place,

$q(\eta,\delta) = 1/2$ if $n,\delta$ agree except at x, and

| | $\eta(x-1)$ | $\eta(x)$ | $\eta(x+1)$ |
|---|---|---|---|
| | 1 | 0 | 0 |
| either | 0 | 0 | 1 |
| | 1 | 1 | 0 |
| | 0 | 1 | 1 |

$q(\eta,\delta) = 1$

if $n,\delta$ disagree at x and

| | | | |
|---|---|---|---|
| either | 1 | 0 | 1 |
| | 0 | 1 | 0 |

NB: 1) all 0's $\underline{0}$ and all 1's $\underline{1}$ are absorbing.

2) every state except $\underline{1}$ can reach $\underline{0}$

3) every state except $\underline{0}$ can reach $\underline{1}$.

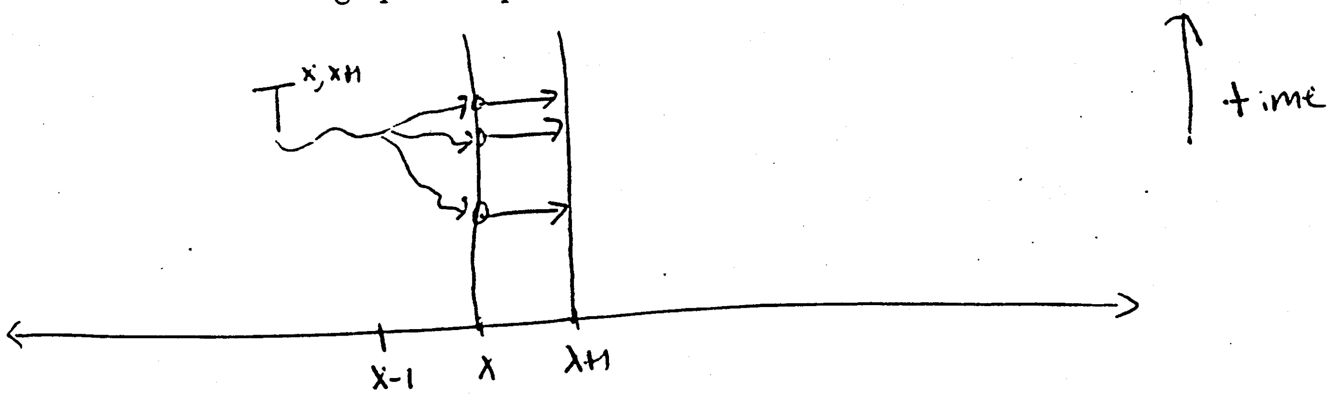
So the behavior is trivial, eventually the system will get stuck in either $\underline{0}$ or $\underline{1}$.

Hence, the only stationary distributions are $\delta_{\underline{0}}$, $\delta_{\underline{1}}$ or convex combinations $\lambda\delta_{\underline{0}} + (1-\lambda)\delta_{\underline{1}}$ $0 \le \lambda \le 1$.

Now, what happens if we do this on $\{0,1\}^{Z}$? More generally what happens in d-dimensions, $X = \{0,1\}^{Z^d}$?

$\underline{0}$ and $\underline{1}$ are still absorbing and anything except $\underline{1}$ "can reach" $\underline{0}$.
But, the space is not countable and so one cannot conclude that the only SD are $\delta_{\underline{0}}$, $\delta_{\underline{1}}$ (and their convex combinations).

**Theorem:** $\exists$ a nontrivial S.D. for the voter model in d-distribution iff $d \ge 3$.
NB: Very nice - says that there is a lot of structure.
**Proof.** Introduce the graphical representation




$\forall x \in Z^d$, $\forall y \in Z^d$ with $|y-x| = 1$, let

> $T^{x,y}$ be a Poisson process with rate $\frac{1}{2d}$ and such that the various Poisson processes are independent.

"The points of the Poisson process $T^{x,y}$ will be the "random times" and which x forces y to conform".

At the points of the $T^{x,y}$ Poisson process, we place arrows from x to y.

At each time point of $T^{x,y}$,

y will change its value to that of x.

A little thought (using basic properties of Poisson processes) shows that each lattice point x waits an exponential (1) random time and then chooses one of its 2d neighbors at random to force to conform.

The above should convince us that we should define the process as follows. Let $\eta_0 \in X$ be arbitrary which will be our intitial state. Then let $\eta_t^{\eta_0}(x)$ be the value $\eta$ obtained when sliding down the time axis, backwards along arrows until we get to time 0.

$\eta_t$ depends only on the Poisson processes where all the randomness is coming in.

NB: By using the same Poisson process realization, we get all the process $\{\eta_t^{\eta_0}\}_{\eta_0 \in X}$ defined on the same prob space (i.e. coupled).

EXTREMELY IMPORTANT OBSERVATION

Start at the point (x,t) in the space-time diagram. Let $\{Y_s^{x,t}\}_{0 \leq s \leq t}$ be the location of where we are as we slide down the time line moving backwards on arrows. Note

$Y_0^{x,t} = x$ and the process is only defined for t units of time.

NB: $\{Y_s^{x,t}\}_{0\le s\le t}$ is a continuous time random walk which jumps at rate 1 and chooses each of its 2d neighbors with equal probability. Moreover, if $z \neq x$, then $\{Y_s^{x,t}\}_{0\le s\le t}$ and $\{Y_s^{z,t}\}_{0\le s\le t}$ are independent random walks UNTIL they meet.

Hence $\{Y_s^{x,t} - Y_s^{z,t}\}_{0\le s\le t}$ is a random walk starting at x-z which jumps at rate 2 (and moves to a randomly (uniform) chosen neighbor) UNTIL it hits 0.

**Theorem:** If d = 1,2, there are no nontrivial stationary distributions; i.e. $I = \{\lambda\delta_{\underline{0}} + (1-\lambda)\delta_{\underline{1}}\}_{\lambda\in[0,1]}$ where I is the set of stationary distributions.

**Proof.** The work is in showing the following lemma. After this, it is more abstract (soft) arguing.

**Lemma.** $\forall\eta \in X$, $\forall x,y \in Z(Z^2)$

$$\text{Prob}\,(\eta_t^\eta(x) \neq \eta_t^\eta(y)) \xrightarrow{t\to\infty} 0.$$

(Of course the superscript $\eta$ is the initial configuration and so in particular $\eta_0^\eta = \eta$).

**Proof.** Fix t.

If $Y_s^{t,x} = Y_s^{t,y}$ for any $s \in [0,t]$ then $\eta_t^\eta(x) = \eta_t^\eta(y)$ by definition of the process.

$\Rightarrow P(\eta_t^\eta(x) \neq \eta_t^\eta(y)) \le P(Y_s^{t,x} - Y_s^{t,y} \neq 0 \quad \forall 0 \le s \le t)$

= Prob(a rate 2 random walk starting at x-y does not hit 0 by time t).

Finally, the last expression $\longrightarrow 0$ as $t \to \infty$ by reccurence of 1 or 2 dimensional random walk $\Rightarrow$ lemma. QED

NB1: The lemma clearly holds uniformly in $\eta$.

NB2: Once $Y_s^{t,x} = Y_s^{t,y}$, they stay $=$.

This perhaps seems to imply that $P(\eta_t^\eta(x) = \eta_t^\eta(y) \ \ \forall t \geq T) \xrightarrow{T\to\infty} 1.$

This is wrong and you should think through why.

Proof of Theorem: Let $\mu$ be a nontrivial S.D. Then $\exists\, x \neq y$ and $\mu(\eta(x) \neq \eta(y)) > 0$. Consider the stationary process $\eta_t^\mu$.

[For any measure $\nu$, $\eta_t^\nu$ means the process where the initial distribution is chosen according to $\nu$. Of course for any set $A \subseteq X$, $P(\eta_t^\nu \in A) = \int_X P(\eta_t^\eta \in A)d\nu(\eta)$.]

So $\eta_t^\mu \overset{dis}{=} \mu \ \ \forall t$ and so

$\text{Prob}(\eta_t^\mu(x) \neq \eta_t^\mu(y)) = \mu(\eta(x) \neq \eta(y)) \ \ \forall t$. But LHS $\to 0$ as $t \to \infty$ by the lemma which clearly holds for randomly chosen $\eta$ as well. (Why?). QED

Theorem: If $d \geq 3$, $\exists$ nontrivial S.D.

**Proof:** Let $\mu_{1/2}$ be product measure with density 1/2 (i.e. each $x \in Z^d$ independently flips a fair coin to decide its initial state).

* (Of course the initial configuration chosen according to $\mu_{1/2}$ should be independent of the graphical representation.)

Now $\forall xy$ $\{\eta_t^{\mu_{1/2}}(x) \neq \eta_t^{\mu_{1/2}}(y)\} = \{Y_s^{t,x} - Y_s^{t,y} \neq 0 \ \forall_{0\leq s\leq t}\} \cap \{\eta_0^{\mu_{\frac12}}(Y_t^{t,x}) \neq \eta_0^{\mu_{\frac12}}(Y_t^{t,y})\}$.

Now by * (and some thought) the ~~last two~~ prob of the 2nd event ~~are indep ⇒~~ conditioned on the 1st event = 1/2 + so

$P\{\eta_t^{\mu_{\frac12}}(x) \neq \eta_t^{\mu_{\frac12}}(y)\} =$ P(random walk starting at x-y which jumps at rate 2 does not hit 0 by time t) $\cdot \frac{1}{2}$. The RHS decreases as $t \to \infty$ to

$\frac{1}{2}$ P(random walk starting at x-y which jumps at rate 2 never hits 0)

$\equiv \frac{1}{2} g(x-y) > 0$ by transience of random walk for $d \geq 3$.

For arbitrary $\nu$, let $T(t)\nu$ denote the distribution of $\eta_t^\nu$, that is, the distribution at time t if we start with distribution $\nu$.

So $T(t)\mu_{1/2}\{\eta(x) \neq \eta(y)\} = P\{\eta_t^{\mu_{\frac{1}{2}}}(x) \neq \eta_t^{\mu_{\frac{1}{2}}}(y)\} \geq \frac{1}{2} g(x-y)$ $\forall t$. Let

$$m_T \equiv \frac{1}{T}\int_0^T T(t)\mu_{\frac{1}{2}}\, dt.$$

Then $m_T\{\eta(x) \neq \eta(y)\} \geq \frac{1}{2} g(x-y)$. Let $\nu_{\frac{1}{2}}$ be any subsequential weak limit of $\{m_T\}_{T\geq 0}$. Then (by our general theorem) $\nu$ is a stationary distribution and $\nu_{\frac{1}{2}}\{\eta(x) \neq \eta(y)\} \geq \frac{1}{2} g(x-y) > 0$ since the same is true of $m_T$ $\forall T$ <u>and</u> $\{\eta(x) \neq \eta(y)\}$ is both open and closed in the product topology. $\nu_{\frac{1}{2}}$ is then our nontrivial stationary distribution since for any $\lambda$

$$(\lambda\delta_{\underline{0}} + (1-\lambda)\delta_{\underline{1}})\{\eta(x) \neq \eta(y)\} = 0. \qquad \text{QED}$$

NB: $\nu_{\frac{1}{2}}\{\eta(x) = 1\} = 1/2$ by symmetry.

Exercise: Carry out a similar analysis starting with $\mu_\theta$, product measure with density $\theta$. Actually $\eta_t^{\mu_{\frac{1}{2}}}$ itself converges (i.e. we do not need to take Cesaro averages). The way to show this is via a (non-self) duality equation.

Let $X_t^A$ be the process obtained by intially putting down particles (1's) at points of A and letting them do independent continuous time random walks (jumps at rate 1) but such that if 2 particles land on the same spot, they coalesce (i.e. they become 1).

Duality Equation:

$$P(\eta_t^A \cap B \neq \emptyset) = P(X_t^B \cap A \neq \emptyset). \text{ (Think how to prove this!)}$$

Now let A be random chosen according to $\mu_\theta$.

Then

$$P(\eta_t^{\mu_\theta} \cap B = \emptyset) = E[(1-\theta)^{|X_t^B|}]$$

since given $|X_t^B| = K$, $P(X_t^B \cap A = \emptyset) = (1-\theta)^K$.

$$|X_t^B| \searrow \Rightarrow \text{RHS} \nearrow \Rightarrow \text{LHS} \nearrow \text{ to some limit.}$$

this implies as we saw before

$$\eta_t^{\mu_\theta} \to \text{some limit.}$$

**The interesting facts about the voter model.**

Based on what we have done, you should be able to do all of these

1. In 1-dim starting from $\mu_\theta$,

   $P(\eta_t(x) \neq \eta_t(x+1)) \sim \frac{c}{\sqrt{t}}$.

   Does this imply $P(\eta_t(x) = \eta_t(x+1) \ \forall t \geq T) \to 1$ as $T \to \infty$.

2. Show that starting with $\mu_{1/2}$ in any $d$, as each pt. changes i.o.

3. In 2-dim starting from $\mu_\theta$, $P(\eta_t(x) \neq \eta_t(x+1)) \sim c/\log t$ (This uses highly non-trivial facts about random walks in 2-dimensions).

4. In $d \geq 3$, what type of correlations does $\nu$ have?

   In other words,

   does $E^{\nu_\theta}[\eta(x)\eta(y)] - E^{\nu_\theta}[\eta(x)]\, E^{\nu_\theta}[\eta(y)] \xrightarrow{|x-y| \to \infty} 0$?

   If so, at what rate? exponential? power law?

   Take $\theta = 1/2$ for simplicity.

# Part IV Statistical Mechanics

(No Physics background needed)

From a physics point of view, the Ising Model is a sophisticated model of a magnet and puts on a rigorous foundation the interesting physical phenomenon of "spontaneous magnetization".

From a probabilistic viewpoint, it illustrates the following. First, we saw that uncountable state MP behave completely differently than countable state.

We now consider MP whose state space is $\{-1,1\}$ but parameter set is $Z^d$, so called "Markov random fields". These exhibit behavior completely different than ordinary finite state Markov Chains, i.e. with $d = 1$.

In particular, irreducibility (properly defined) does not imply uniqueness of a S.D. even though the state space is finite with only 2 states.

We move into the theory of Gibbs states in statistical mechanics.

## Boltzmann Distribution

**Definition:** Let $S$ be a finite set and $\mu$ be a measure on $S$. The entropy of $\mu$, $E(\mu)$, is

$$-\sum_{x \in S} \mu(x) \log \mu(x).$$

**Exercise:** Show $E(\mu)$ is maximized uniquely at the uniform distribution where the entropy is $\log|S|$. (Hint: Jensen's $\leq$)

Exercise: Think about what entropy "means".

Now let each $x \in S$ have a certain energy $H(x)$ (If you want, you can think of energy as an arbitrary function from $S \to R$).

"Physics says", the world wants to minimize energy and maximize entropy.

**Definition:** The energy of a measure $\mu$, $H(\mu)$, is naturally enough $\sum_{x \in S} \mu(x) H(x)$, that is the expected energy when $x \in S$ is chosen according to $\mu$. We conclude that the world naturally chooses that measure which maximizes $E(\mu) - H(\mu)$. (Note $E(\mu)$ does not depend on the energy function). We take this as an assumed physical principle but there is the following theorem.

**Theorem 1:** Consider the mapping $\mu \to E(\mu) - H(\mu)$. This is uniquely maximized at the so-called Boltzmann distribution,

$$\mu(x) = \frac{e^{-H(x)}}{Z}$$

where $Z = \sum_{x \in S} e^{-H(x)}$ is the normalization constant, also called the partition function.

**Note:** The lower the energy, the higher the probability.

The proof of this is left as an exercise using Jensens $\leq$.

The Ising Model is the simplest example of a Gibb's state and the only such we will study.

This model was introduced in the 30's as a model for the physical phenomenon of "spontaneous magnization", an experimentally verified phenomenon.

We will now move into mathematics - the connection between what we will do and magnets probably being completely unclear.

Let $X = \{-1,1\}^{Z^d}$. Each point of $Z^d$ will be an atom, and each atom will point up (+1) or point down (-1).

Adjacent atoms will try to point in the same drection (this is called ferromagnetism). There will be 2 parameters J and h, called respectively the "coupling interaction" and the "external field". For each J,h, we will have the Ising Model with parameters J and h. $d = 2$ will be assumed for simplicity. Almost all of the theory for $d = 2$ extends to $d \geq 3$ BUT not all. The important phenomenon of "translation symmetry breaking" requires $d \geq 3$ for it to occur and does not occur in $d = 2$ (a highly nontrivial theorem due to Michael Aizenman). $d=1$ is completely different.

The theory of the Ising Model is an extremely rich theory. We will be going back and forth between heuristic discussions and mathematics.

Fix J and h. We now define the Ising Model on a finite box with free boundary conditions".

Let

$$\Lambda_n = [-n,n] \times [-n,n] \subseteq Z^2.$$

We will define a probability measure on $\{-1,1\}^{\Lambda_n}$. Given $\eta \in \{-1,1\}^{\Lambda_n}$, we will define the energy $H_n^{J,h}(\eta)$ of $\eta$ and our probability measure ${}^{J,h}\mu_n^f$ will be the corresponding Boltzmann distribution

$${}^{J,h}\mu_n^f(\eta) = \frac{e^{-H_n^{J,h}(\eta)}}{Z_n^{J,h}} \text{ defined in Theorem 1,}$$

where

$$Z_n^{J,h} = \sum_{\eta\in\{-1,1\}^{\Lambda_n}} e^{-H_n^{J,h}(\eta)}.$$

In ${}^{J,h}\mu_n^f$, n refers to the size of the system, that is $\Lambda_n$, and f refers to the fact that we will use free boundary conditions as described below.

We now define $H_n^{J,h}(\eta)$.

$$H_n^{J,h}(\eta) = -J \sum_{\substack{<x,y> \\ x,y\in\Lambda_n}} \eta(x)\eta(y) - h \sum_{x\in\Lambda_n} \eta(x).$$

The notation $<x,y>$ means that we only sum over pairs $\{x,y\} \subseteq \Lambda_n$ where x and y are nearest neighbors. (This is usual physics notation).

**Exercise:** Think about the intuitive meaning of $H_n^{J,h}(\eta)$.

Here are some comments about $H_n^{J,h}(\eta)$. In the first summand, note

$$\eta(x)\eta(y) = \begin{cases} 1 \text{ if } \eta \text{ agrees at } x \text{ and } y \\ -1 \text{ if } \eta \text{ disagrees at } x \text{ and } y \end{cases}.$$

We assume $J > 0$. Since we have a-sign in front of $J$, each pair $x,y$ which are nearest neighbor (n.n.) contributes $-J$ to the sum if $\eta$ agrees at them and $+J$ if $\eta$ disagrees at them. Since the Boltzmann distribution favors "lower energy" configurations, n.n. pairs x and y will tend to agree rather than disagree.

If $J < 0$, the same argument would imply that n.n. pairs tend to disagree.

$J > 0 \Leftrightarrow$ ferromagnetism

$J < 0 \Leftrightarrow$ antiferromagnetism

The second summand is simpler to understand. h can be $> 0$ or $< 0$. If $h > 0$, each $x \in \Lambda_n$ tends to be positive since this lowers the second summand in H.

We say there are free boundary conditions since there is no interaction with any boundary terms.

"The Ising Model on a finite box with boundary conditions".

Let $\Lambda_n$ be as before.

Let $\partial\Lambda_n = \{y \in Z^2 \sim \Lambda_n : y \text{ is adjacent to some } x \in \Lambda_n\}$.

Let $\delta \in \{-1,1\}^{\partial\Lambda_n}$ be a "boundary condition". Given $\eta \in \{-1,1\}^{\Lambda_n}$, let

$$H_n^{J,h,\delta}(\eta) = -J \sum_{\substack{<x,y> \\ x,y\in\Lambda_n}} \eta(x)\eta(y) - J \sum_{\substack{<x,y> \\ x\in\Lambda_n \\ y\in\partial\Lambda_n}} \eta(x)\delta(y) - h \sum_{x\in\Lambda_n} \eta(x).$$

**Exercise:** Why do we not bother to include the term $-h \sum_{x \in \partial\Lambda_n} \delta(x)$ in the definition of $H_n^{J,h,\delta}$?

We call the corresponding Boltzmann distribution

$${}^{J,h}\mu_n^{\delta} \text{ defined by } {}^{J,h}\mu_n^{\delta}(\eta) = \frac{e^{-H_n^{J,h,\delta}(\eta)}}{Z_n^{J,h,\delta}}$$

(the denominator is simply the normalization.)

the Ising Model on $\Lambda_n$ with boundary condition $\delta$. If $\delta \equiv 1$, we write ${}^{J,h}\mu_n^{+}$ for ${}^{J,h}\mu_n^{\delta}$ and if $\delta \equiv -1$ we write ${}^{J,h}\mu_n^{-}$. These 2 special cases are very important.

In $d = 2$ (or $d \geq 3$), there is a very interesting phenomenon called "Phase transition". There are 2 different ways to describe this. We can now state one way. The other way to state it will require more development.

Let $J > 0$ and $h = 0$. It should be obvious that

* ${}^{J,0}\mu_n^{f}$ is $\{-1,1\}$ symmetric., i.e. if $T: \{-1,1\}^{\Lambda_n} \leftrightarrow$ by

$$T\eta(x) = -\eta(x) \qquad \forall x \in \Lambda_n,$$

(T just flips or reverses a configuration $\eta$) then T takes the measure ${}^{J,0}\mu_n^{f}$ to itself.

i.e. $${}^{J,0}\mu_n^{f}(\eta) = {}^{J,0}\mu_n^{f}(T\eta).$$

**Exercise:** Check or convince yourself of the above and show (or at least realize) that if $h \neq 0$, then ${}^{J,h}\mu_n^{f}$ is not $\{-1,1\}$ symmetric.

**Exercise:** Use the above to show that the expected value at the origin is 0,

i.e. $$E[\eta(0)] = 0$$

(Of course $(\Omega,\mathcal{F},P) = (\{-1,1\}^{\Lambda_n}, P(\{-1,1\}^{\Lambda_n}), {}^{J,0}\mu_n^{f})$, E is expected value and $\eta \to \eta(0)$ is a random variable on $\Omega$).

Exercise: Is the mean still 0 if $h \neq 0$ or if we introduce boundary conditions.

It is believable (and provable) that

Theorem 2: $\text{Prob}^{J,0}\mu_n^+(\eta(0) = 1) > 1/2.$

The reason that this is believable is that we are placing 1's around the boundary which should (since $J > 0$) tend to create more 1's.

One might think that this "effect coming in from the boundary" should be less and less as n (or the size of the box $\Lambda_n$) $\to \infty$ since the boundary gets further away from the origin; i.e. one might think

$$\lim_{n\to\infty} \text{Prob}^{J,0}\mu_n^+(\eta(0) = 1) = 1/2.$$

Amazingly, this is not true. (It is true however if J is small.) But

Theorem 3: $\exists\, J_c \ni \forall J > J_c$,

$$\liminf_{n\to\infty} \text{Prob}^{J,0}\mu_n^+[\eta(0) = 1] > 1/2.$$

We are also in a position to prove this. We mention (and prove later) that the above lim is actually a limit. (This uses monotonicity and couplings as in the Voter Model).

Later on, Theorem 3 will also allow us to show that although a finite state irreducible aperiodic M.C. has a unique stationary distribution, this is false when the parameter set Z for the Markov chain is replaced by $Z^d$ $d \geq 2$ and everything is appropriately defined. Such things are called Markov random fields or n.n. infinite volume Gibbs states.

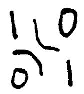
nearest neighbor

We remark that if Theorem 3 were not true, the theory of Gibbs states would probably not exist.

Theorem 3 is false for $d = 1$ (as proven by Ising). Ising also proved Theorem 3 was false in $d = 2$ but fortunately, his proof was wrong.

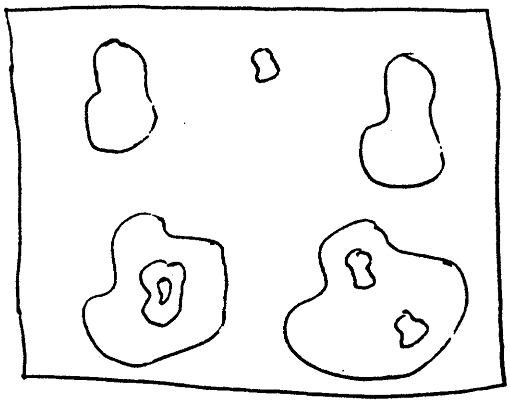
The proof we give is the so-called Peierls argument which is another type of contour argument.

**Proof of Theorem 3**: Consider edges in the dual graph. Fix $\eta \in \{-1,1\}^{\Lambda_n}$. If a dual edge lies between 2 points which have different values, we draw in this dual edge, otherwise we don't. If we have

1 0 we draw the contours 

0 1 like

We then obtain a bunch of closed contours which schematically look like:



Remember: this set of contours depends on $\eta$.

Now the event

$$\{\eta(0) = -1\} \subseteq \{\exists \text{ contour of drawn dual edges surrounding } 0\}.$$

This is simply a graph theory fact which we won't prove.

**Lemma**: Let $\gamma$ be a contour of length $\ell$ in the dual graph of $\Lambda_n$ surrounding 0. Then

$$\text{Prob}^{\mu_n^{J,0,+}}\{\text{all the edges of } \gamma \text{ are drawn}\} \leq e^{-2J\ell}.$$

Proof later.

Since the number of contours around the origin of length $\ell$ is (see percolation stuff) $\leq \ell 3^{\ell-1}$, $\text{Prob}^{\mu_n^{J,0,+}}(\eta(0) = -1) \leq \sum_{\ell=1}^{\infty} \ell 3^{\ell-1} e^{-2J\ell}$ by the lemma.

It is obvious that for large J, the sum converges, and therefore since the terms $\to 0$ as $J \to \infty$, it is clear that this sum is for large J bounded away from 1/2. Since this upper bound does not depend on n, Theorem 3 follows. QED

**Proof of Lemma:** Recall $\gamma$ is fixed. Let T: $\{-1,1\}^{\Lambda_n} \leftrightarrow$ be defined by flipping the value at each x inside $\gamma$ and not flipping the value at each x outside $\gamma$.

i.e.
$$T\eta(x) = \begin{cases} \eta(x) & x \text{ not inside } \gamma \\ -\eta(x) & x \text{ inside } \gamma. \end{cases}$$

Note T is a bijection. Let E = {all edges of $\gamma$ are drawn}. Note $T(E) \cap E = \emptyset$.

Let $\eta \in E$. We compare $H(\eta)$ and $H(T\eta)$. Since the "Hamiltonian" H gives value $-J$ for equal adjacent pairs and $+J$ for unequal adjacent pairs, $H(T\eta) = H(\eta) - 2J|\gamma|$. ($|\gamma|$ = length of $\gamma$). We now write P for $P^{\mu_n^{J,0,+}}$. Therefore $P(\eta) = \frac{e^{-H(\eta)}}{Z} = \frac{e^{-2J|\gamma|}e^{-H(T\eta)}}{Z} = \frac{e^{-2J\ell}e^{-H(T\eta)}}{Z}$. It follows that

$$P(E) = \sum_{\eta\in E} P(\eta) = e^{-2J\ell} \sum_{\eta\in E} \frac{e^{-H(T\eta)}}{Z}$$

$$\leq e^{-2\ell} \sum_{\eta\in\{-1,1\}^{\Lambda_n}} \frac{e^{-H(T\eta)}}{Z} = e^{-2J\ell},$$

the last = following from T being bijective. QED

We now take a quick excursion into Markov Random Fields and then return to the Ising Model.

First, 1-dimension, which we all know. Let $P = (P_{ij})$ be an n×n transition matrix with all positive entries. ($n < \infty$).

**Definition:** A Stochastic process $\{X_n\}_{n\in z}$ is called a Markov process relative to P if

$$* \qquad P(X_n = j \mid X_{n-1} = i, X_{n-2}, \ldots, X_{n-k}) = P_{ij} \qquad \forall i,j,n,k.$$

Everyone knows that there is only one such process (i.e. all processes satisfying $*$ have the same finite dimensional distributions).

This is false for Markov Random fields (where Z is replaced by $Z^d$) which we now define. We take $d = 2$.

Let $\qquad B = \{(1,0), (-1,0), (0,1), (0,-1)\}$.

Let $\qquad P: \{-1,1\}^B \to (0,1)$ be arbitrary. (We will call P a "specification").

(This will be the analogue of the transition matrix in higher dimensions.)

One should think of a specification as follows. It specifies what the conditional distribution (for some process $(X_n)_{n\in Z^2}$ taking values $\pm 1$) of $X_n$ given $X_{n+(1,0)}$, $X_{n+(-1,0)}$, $X_{n+(0,1)}$, $X_{n+(0,-1)}$ should be: namely:

$$\text{Prob}(X_n = 1 \mid X_{n+(1,0)}, X_{n+(-1,0)}, X_{n+(0,1)}, X_{n+(0,-1)})$$

should be $P(X_{n+(1,0)}, X_{n+(-1,0)}, X_{n+(0,1)}, X_{n+(0,-1)})$.

This is analogous to a transition matrix where the conditional distribution of $X_{n+1}$ given $X_n$ is given by the transition matrix.

**Definition:** A Stochastic process taking on values -1 and 1 and indexed by $Z^2$, $\{X_n\}_{n\in Z^2}$, is a <u>Markov random field</u> relative to P if $\forall n \in Z^2$.

$$\text{Prob}\{X_n = 1 \mid X_m, \ m \neq n\}$$

$$= P(X_m, m \in B+n).$$

The definition seems complicated but is quite simple. The idea is that if we condition on all lattice points except n, the conditional distributions of the process at n only depends on the 4 nearest neighbors of n and how this depends on these 4 points is given by P, the analogue of the transition matrix. Note P mapping into (0,1) instead of [0,1] is analogous to having a transition matrix with all positive entries. The point is that no matter what we see at points $m \neq n$, we always have positive probability of seeing either a 1 or -1 at n.

We now show that there are specifications which have more than 1 Markov Random Field associated to them.

NB: If you have not seen this before, you should find this surprising.
We now give a family of specifications indexed by J and h by returning to the Ising Model. We will for each J and h, give a specification $P^{J,h}$.
Let $\delta \in \{-1,1\}^B$. We need to define $P^{J,h}(\delta)$. Consider the Ising Model on the box consisting of only $\{0\}$ with boundary condition $\delta$. This is a measure ${}^{J,h}\mu_0^\delta$ on $\{-1,1\}$. Let $P^{J,h}(\delta) = \text{Prob}^{\,{}^{J,h}\mu_0^\delta}(\{1\})$. This defines the specification $P^{J,h}$.

Note: Unfortunately, the notation gets very messy - this is difficult to avoid. Of course, the best thing is to understand what all the definitions mean - the ideas are simpler than the notation.

**Definition:** The Stochastic process $(X_n)_{n\in Z^2}$ taking values $\pm 1$ is called an "infinite volume Gibbs state with parameter J and h" if it is a Markov random field relative to the specification $P^{J,h}$.

Let $G_{J,h}$ denote this set of processes. Untangling definitions, $\{X_n\} \in G_{J,h}$ if the conditional distribution of $X_n$ given $X_m$, $n \neq n$ only depends on the 4 nearest neighbors and this conditional distribution of $X_n$ given $\{X_m, m \in n+B\}$ is simply the Ising Model on the finite box $\{0\}$ with boundary conditions $\{X_m, m \in n+B\}$, i.e. $\text{Prob}(X_n = 1 \mid X_m, m \neq n)$

$$= \frac{e^{-JY_n - h}}{e^{-JY_n - h} + e^{JY_n + h}} \text{ where } Y_n = \sum_{m \in n+B} X_m.$$

**Exercise**: Check the above.

We finally prove there can be more than 1 MRF for a specification by showing for some $J$ and $h$, $|G_{J,h}| > 1$.

**Theorem 4**:

a) $\forall J\ \forall h \neq 0$, $|G_{J,h}| = 1$

b) $\exists\, J_c \in (0,\infty) \ni \forall J < J_c$, $|G_{J,0}| = 1$ $\quad \forall J > J_c$ $|G_{J,0}| > 1$.

In $d = 2$ it is known that $|G_{J_c,0}| = 1$. It is believed but unproven for $d \geq 3$.

The fact that for $h = 0$, $J$ large $|G_{J,0}| > 1$ follows (with some work) from Theorem 3.

The fact that for $h = 0$, $J$ small $|G_{J,0}| = 1$ requires the construction of a certain operator which for small $J$ is a contraction which will then have a unique solution demonstrating $|G_{J,0}| = 1$.

a) is the most difficult. It is due to David Ruelle and there are 2 approaches to its proof, one using convex analysis, the other using normal families in complex analysis.

The rest of the time will be spent to prove Theorem 4. However, we first give some more general facts.

We identify processes $\{X_n\}_{n \in Z^2}$ taking values $\pm 1$ with prob measures on $\{-1,1\}^{Z^2}$ (by looking at the distribution of $\{X_n\}$; i.e. by pushing the measure forward).

We can consider $G_{J,h}$ to be a set of measures. In the case $|G_{J,h}| > 1$, it is interesting to ask how big $|G_{J,h}|$ is. Since it can be shown $G_{J,h}$ is a convex set (in the space of measures), $|G_{J,h}|$ will be $\infty$. Since $G_{J,h}$ is convex, the right? to ask is how many extreme points are there in $G_{J,h}$.

Theorem 5: If $d = 2$, if $|G_{J,h}| > 1$ (in which case $h = 0$ by Theorem 4), $G_{J,h}$ has 2 extreme points.

This very deep theorem (due to M. Aizerman) will not be proven.

Theorem 6: In $d = 3$, $\exists$ values of $J$ for which $G_{J,0}$ has $\infty$ many extreme points. This is due to R. Dobrushin, one of the first developers of the mathematical theory of Gibbs states.

One can ask if the elements of $G_{J,h}$ are translation invariant.

Theorem 5 and what we will do below gives

Theorem 7: In $d = 2$, $\forall J$ $\forall h$, all $\mu \in G_{J,h}$ are translation invariant. The proof of Theorem 6 shows

Theorem 8: In $d = 3$ for $h = 0$ and large J, $\exists\ \mu \in G_{J,0}$ which are not translation invariant.

Theorem 7 and 8 allow us to say "one needs 3-dimensions for a translation symmetry breaking".

In $d = 2$, there are Gibb's states for $h = 0$ which have mostly 1's. This is a $\pm 1$ symmetry breaking since the specification is $-1,1$ symmetric (as $h = 0$) but there are associated Markov Random Fields (MRF) which are not, i.e. a symmetry of the specification is broken.

Symmetry Breaking is a general concept where a specification is invariant under a certain group action but where there are associated MRF which are not invariant under the induced group action. The above are 2 examples but we will not discuss symmetry breaking further.

We now return to proving Therorem 4.

Recall ${}^{J,h}\mu_n^+$ is a measrue on $\{-1,1\}^{\Lambda_n}$. We modify this to be a measure on $\{-1,1\}^{Z^2}$ by requiring that all $x \notin \Lambda_n$ are 1 with prob 1.

**Lemma 1:** $\{{}^{J,h}\mu_n^+\}_{n\geq 1}$ converges as $n \to \infty$ to some measure which we call ${}^{J,h}\mu^+$. ${}^{J,h}\mu^-$ is defined similarly.

The proof is fairly long: we only sketch it.

**Lemma 1:** If $\eta \leq \delta$ are in $\{-1,1\}^{\Lambda_n}$, then $\forall h, \forall J$

$$ {}^{J,h}\mu_n^\eta \leq {}^{J,h}\mu_n^\delta . $$

This lemma follows (with a little work) from Holleys $\leq$ which says

Theorem (Holley $\leq$): Let $X = \{0,1\}^S$, $|S| < \infty$. Let $\mu_1, \mu_2$ be strictly positive prob measures on $X$. If

$$ \mu_1(\eta \wedge \delta)\mu_2(\eta \vee \delta) \geq \mu_1(\eta)\mu_2(\delta) $$

$\forall \eta, \delta$, then $\mu_1 \leq \mu_2$.

$(\eta \wedge \delta(x) = \min\{\eta(x), \delta(x)\},\ \eta \vee \delta(x) = \max\{\eta(x), \delta(x)\})$.

One proves this by intelligently coupling 2 continuous time Markov chains, one with stationary distribution $\mu_1$, the other with $\mu_2$.

**Lemma 2:** Let $n > m$. Let $\delta \in \partial\Lambda_n$, $\eta \in \Lambda_n - \Lambda_m$ and $\bar{\eta}$ the restriction of $\eta$ to $\partial\Lambda_m$. Then

$$ {}^{J,h}\mu_n^\delta(\,|\,\eta)\big|_{\Lambda_m} = {}^{J,h}\mu_n^{\bar{\eta}} $$

**Proof.** Computation.

**Exercise:** Think about what this means. (It is important.)

**Theorem:** Let ${}^{J,h}\mu_n^+$ be considered to be a p.m on $\{-1,1\}^{Z^2}$ by letting the random configuration be $\equiv 1$ outside $\Lambda_n$.

Then $\{{}^{J,h}\mu_n^+\}_{n\geq 1}$ is a decreasing sequence of measures (and hence converges to some ${}^{J,h}\mu^+$).

**Proof:** This follows (with some work) from Lemmas 1 and 2. QED

**Exercise:** Why does the fact that it decreases imply that it converges?

At this point, we need on equivalent formulation of an infinite volume Gibbs state.

Let ${}^{J,h}M_n = \{{}^{J,h}\mu_n^\delta\colon \delta \in \partial\Lambda_n\}$.

Let $\overline{{}^{J,h}M_n}$ = closed convex hull of ${}^{J,h}M_n$.

EQUIVALENCE THEOREM:

$$\overline{{}^{J,h}M_{n+1}} \subseteq \overline{{}^{J,h}M_n} \quad \text{and} \quad G_{J,h} = \bigcap_{n=1}^{\infty} \overline{{}^{J,h}M_n}.$$

(We do not prove this - see Liggett.)

In view of the equivalence theorem together with Lemma 1 and 2 it follows:

**Theorem 2:** ${}^{J,h}\mu^-$ and ${}^{J,h}\mu^+$ are in $G_{J,h}$ and moreover

$$\forall\ \nu \in G_{J,h},\ {}^{J,h}\mu^- \leq \nu \leq {}^{J,h}\mu^+.$$

**Exercise:** Show ${}^{J,h}\mu^+$ is $Z^2$-invariant (i.e. stationary)

**Corollary:** In view of Lemma 2, we have that for any h and J,

$$|G_{J,h}| = 1 \text{ iff } {}^{J,h}\mu^- = {}^{J,h}\mu^+.$$

We can now prove a part of Theorem 4.

**Lemma 3:** $\exists\, J_c \ni \forall J > J_c\ |G_{J,0}| > 1$.

**Proof.** It suffices to show

$$* \qquad \text{Prob}^{\,{}^{J,0}\mu^+}(\eta(0) = 1) > 1/2$$

since by symmetry,

$$\text{Prob}^{\,{}^{J,0}\mu^-}(\eta(0) = -1) = \text{Prob}^{\,{}^{J,0}\mu^+}(\eta(0) = 1) > 1/2$$

while $\text{Prob}^{\,{}^{J,0}\mu_n^+}(\eta(0) = -1)$ would be $< 1/2 \Rightarrow {}^{J,0}\mu^+$ and ${}^{J,0}\mu^-$ give different prob to the event $\{\eta(0) = -1) \Rightarrow$ by the corollary that $|G_{J,0}| > 1$. $*$ follows from Theorem 3 as follows.

$${}^{J,0}\mu_n^+ \to {}^{J,0}\mu^+ \text{ by Lemma 1 which implies (why ?)}$$

$$\text{Prob}^{\,{}^{J,0}\mu_n^+}(\eta(0) = 1) \to \text{Prob}^{\,{}^{J,0}\mu^+}(\eta(0) = 1).$$

Theorem 3 now immediately implies $*$. QED

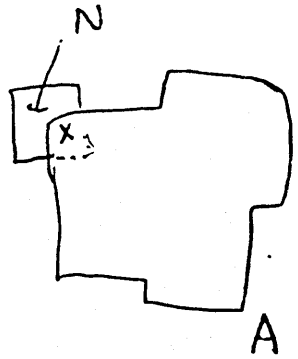
We have now shown that for $h = 0$ and large J, $|G_{J,0}| > 1$. We now show that for J small and $h = 0$.

$$|G_{J,0}| = 1.$$

**Kirkwood Salzburg Equation approach:**

Let $\mu \in G_{J,0}$. Let $\rho(A) = \mu(w \equiv 1 \text{ on } A)$ for $A \subseteq Z^2$ with $|A| < \infty$, called the correlation function for $\mu$. (Inclusion - Exclusion $\Rightarrow \mu$ is determined by $\rho$).

The KS-Equation tells us that there is some "finite integral operator" $\ni$ the correlation function for any Gibbs state is fixed by this operator. If this operator is a contraction, we get uniqueness.




**Theorem.** $\forall$ A finite and $x \in A$,

$$\rho(A) = \sum_{\substack{D \\ |D|<\infty}} K_x(A,D)\rho(D)$$

where

$$K_x(A,D) = \begin{cases} (-1)^{|C|} \sum_{B \subseteq C} (-1)^{|B|} \pi_{\{x\}}^{(A\setminus x)\cup B}(\eta(x) = 1) & \text{if } D = (A\setminus x)\cup C \text{ with } C \subseteq N(x,A) = \{y \notin A, y \sim x\} \\ 0 & \text{o.w.} \end{cases}$$

$\pi_{\{x\}}^{(A\setminus x)\cup B}$ is the Gibbs measure on the 1 point set $\{x\}$ with boundary condition given as follows:

If $z \sim x$, put a 1 at z if $z \in (A\setminus x)\cup B$ and a $-1$ otherwise.

**Theorem:** Assume $\exists\, \lambda < 1 \ni \forall A$ finite $\exists x \in A \ni$

$\sum_D |K_x(A,D)| \leq \lambda$. Then the Gibbs measure is unique.

NOTE: The kernels $K_x(A,D)$ only depend on the specification.

**Proof.** Let $\mu, \tilde{\mu} \in G_{J,0}$, and $\rho, \tilde{\rho}$ be the corresponding correllation functions. Let A be arbitrary and $x \in A$ as above. Then

$$|\rho(A) - \tilde{\rho}(A) = |\sum_D K_x(A,D)[\rho(D) - \tilde{\rho}(D)]|$$

$$\leq \sup_D |\rho(D) - \tilde{\rho}(D)|\lambda.$$

Now taking sup over $A \Rightarrow$ (as $\lambda < 1$)

$$\rho(A) = \tilde{\rho}(A) \ \forall A \Rightarrow \mu = \tilde{\mu}. \qquad \text{QED}$$

**Theorem:** If J is sufficiently small, $|G_{J,0}| = 1$.

**Proof:** Choose J so that

$$\sup_{\substack{k,\ell \\ k,\ell=0,1,2,3,4}} \left| \frac{e^{-Jk}}{e^{-Jk} + e^{Jk}} - \frac{e^{-J\ell}}{e^{-J\ell} + e^{J\ell}} \right| < \frac{1}{2^4 2^4}$$

We now apply the previous theorem.

$$\sum_D |K_x(A,D)| = \sum_{C \subseteq N(x,A)} \left| (-1)^{|C|} \sum_{B \subseteq C} (-1)^{|B|} \pi_x^{(A \setminus x) \cup B} (\eta(x) = 1) \right|$$

$$= \sum_{C \subseteq N(x,A)} \left| \sum_{B \subseteq C} (-1)^{|B|} \pi_x^{(A \setminus x) \cup B} (\eta(x) = 1) \right|$$

By the way we chose J, all the terms $\pi_x^{(A \setminus x) \cup B} (\eta(x) = 1)$ are within $\frac{1}{2^4 2^4}$ of each other. Hence (think combinatorially)

$$\left| \sum_{B \subseteq C} (-1)^{|B|} \pi_x^{(A \setminus x) \cup B} (\eta(x) = 1) \right| \leq \frac{2^{|C|}}{2} \frac{1}{2^4 2^4}$$

$$\leq \frac{2^4}{2} \frac{1}{2^4 2^4} = \frac{1}{2^5} \text{ as } |C| \leq 4.$$

Since $|N(x,A)| \leq 4 \Rightarrow$ all of the above is $\leq 2^4 \frac{1}{2^5} = \frac{1}{2}$. QED

<u>Derivation of K-S Equation</u>

**Lemma:** Let P be an arbitrary p.m. on $\{-1,1\}^{\Lambda}$. Let $R, T \subseteq \Lambda$, $R \cap T = \emptyset$, $S \subseteq T$.

Then

$$P(\eta \equiv 1 \text{ on } R \cup S,\ \eta \equiv -1 \text{ on } T\backslash S)$$
$$= \sum_{S \subseteq Q \subseteq T} P(\eta \equiv 1 \text{ on } R \cup Q)(-1)^{|S|+|Q|}.$$

**Proof.** Left the reader BUT it is easy - it is simply basic inclusion - exclusion.

QED

**Proof of K-S Equation**

$$\rho(A) = \mu[\omega \equiv 1 \text{ on } A]$$
$$= \sum_{B \subseteq N} \mu[\omega \equiv 1 \text{ on } A \cup B,\ \omega \equiv -1 \text{ on } N\backslash B]$$
$$= \sum_{B \subseteq N} \mu[\omega \equiv 1 \text{ on } (A\backslash\{x\}) \cup B,\ \omega \equiv -1 \text{ on } N\backslash B]\ \pi_{\{x\}}^{(A\backslash x)\cup(B)}(\eta(x)=1)$$
$$\overset{\text{Lemma}}{=} \sum_{B \subseteq N} \pi_{\{x\}}^{(A\backslash x)\cup B}(\eta(x)=1) \sum_{B \subseteq C \subseteq N} \rho((A\backslash\{x\}) \cup C)(-1)^{|B|}(-1)^{|C|}$$
$$= \sum_{C \subseteq N} \rho((A\backslash x) \cup C)(-1)^{|C|} \sum_{B \subseteq C} \Pi_{\{x\}}^{(A\backslash x)\cup B}(\eta(x) = 1)(-1)^{|B|}$$
$$= \sum_{D} \rho(D)\, K_x(A,D).$$

QED

Exercise: There is nothing special about $h = 0$ here. Show that $\forall h$, for small $J$, there is a unique Gibbs state.

To finish b) and prove the existence of such a $J_c$, it suffices to show

$$J_1 < J_2 \text{ and } |G_{J_1,0}| > 1 \Rightarrow |G_{J_2,0}| > 1.$$

This will follow from one of the "Griffith inequalities" which we state but do not prove. (It is not easy, not real hard and in [Liggett]).

Consider $\{-1,1\}^S$; $|S| < \infty$. We now let the coupling constant $J$ depend on the edge and the external field depend on the edge.

**Theorem:** If all couplings and external fields are positive, then if $J_B$ is the coupling at some edge B,

$$\frac{\partial}{\partial J_B} P(\eta(0) = 1) \geq 0.$$

<u>Some comments</u>:

1) The Ising model on $\{-1,1\}^S$ with + boundary conditions is the same as the Ising Model with free boundary conditions if we increase the external field by 1 unit for points in S adjacent to $\partial S$.

i.e. If we consider Ising Models with varying external field, we can always take free boundary conditions.

2) Why does Griffith's $\leq \Rightarrow J_1 < J_2$, $|G_{J_1,0}| > 1 \Rightarrow |G_{J_2,0}| > 1$.

As above, by thinking of + boundaries as increasing the external field at the boundary points of S, Griffith $\Rightarrow$

$$P_{J_1,0\,\mu_n^+}(\eta(0) = 1) \leq P_{J_2,0\,\mu_n^+}(\eta(0) = 1).$$

By assumption the LHS $\to C > 1/2 \Rightarrow$ same for RHS $\Rightarrow$ $|G_{J_2,0}| > 1$. QED

**Exercise:** Using Griffith's $\leq$, show that if $|G_{J,0}| > 1$, then the same is true in 3-dimensions. (i.e. if there is phase transition at J in d-dimensions, then there is also phase transition at J in d+1-dimensions.)

Finally Proof of Theorem 4(a). This is the most difficult part of Theorem 4, proven by Ruelle. This is very long but introduces a number of important concepts. The first is the so-called partition function

$$Z_n^{J,h} = \sum_{\eta\in\{-1,1\}^{\Lambda_n}} e^{J \sum_{\substack{\langle x,y\rangle \\ x,y\in\Lambda_n}} \eta(x)\eta(y) + h \sum_{x\in\Lambda_n} \eta(x)} .$$

We want to look at $Z_n^{J,h}$ for large $n$. This will $\to \infty$. To figure out the correct normalization, note that if $J = 0 = h$, then

$$Z_n^{J,h} = 2^{|\Lambda_n|} = 2^{(2n+1)^2}$$

We therefore look at $\frac{\log Z_n^{J,h}}{(2n+1)^2} \equiv P_n^{J,h}$.

More generally, recall ${}^{J,h}Z_n^{\delta} = \sum_{\eta\in\{-1,1\}^{\Lambda_n}} e^{-\,{}^{J,h}H_n^{\delta}(\eta)}$.

Let $P_n^{J,h,\delta} = \frac{\log {}^{J,h}Z_n^{\delta}}{(2n+1)^2}$, called the pressure.

**Theorem A:** $\lim_{n\to\infty} P_n^{J,h,\delta}$ exists and is independent of $\delta$ ($\delta$ can be free here).

Let $P(J,h)$ be this limit.

**Theorem B:** $\forall J$ $P(J,h)$ is convex (and hence continuous) in $h$.

**Theorem C:** $\forall J,h$,

$$|G_{J,h_0}| = 1 \Leftrightarrow P(J,h) \text{ is differentiable in } h \text{ at } h = h_0$$

i.e. $\frac{\partial}{\partial h} P(J,h)|_{h=h_0}$ exists.

[NB: $P_n^{J,h,\delta}$ is analytic in h for all n but the limit need not be].

**Corollary:** $\forall$ J, $|G_{J,h}| = 1$ $\forall$ but at most countably many values of h.

**Proof.** A convex function on R is differentiable at all but at most countably many points. QED

**Theorem D:** $\forall$J, P(J,h) is differentiable at all $h \neq 0$.

As far as difficulty, we give them respectively 5, 3, 7 and 10 points, that is, showing P(J,h) is differentiable in h $\forall h \neq 0$ is the most difficult part.

Time constraints (on perhaps laziness) will make this very sketchy.

**The Proof of Theorem A** we skip -- a good proof is in "Entropy, Large Deviations, and Statistical Mechanics" by Ellis.

**Proof of Theorem B:** Since a limit of convex functions is convex, it suffices to show that $P_n^{J,h,f}$ is convex in h for each n. This is algebraic manipulation plus Hölder's $\leq$. We need to show (we suppress f for free now)

$$\frac{\log\ Z_n(J,\lambda h_1+(1-\lambda)h_2)}{n} \leq \frac{\lambda_1\ \log\ Z_n(J,h_1)}{n} + \frac{(1-\lambda)\ \log\ Z_n(J,h_2)}{n}$$

or equivalently.

$$Z_n(J,\ \lambda h_1+\ (1-\lambda)h_2) \leq Z_n(J,h_1)^{\lambda} Z_n(J,h_2)^{1-\lambda}.$$

$$Z_n(J,\lambda h_1+(1-\lambda)h_2) =$$

$$\sum_{\eta\in\{-1,1\}^{\Lambda_n}} e^{J \sum_{\substack{<x,y> \\ x,y\in\Lambda_n}} \eta(x)\eta(y) + (\lambda h_1+(1-\lambda)h_2) \sum_{x\in\Lambda_n} \eta(x)}$$

$$= \sum_{\eta\in\{-1,1\}^{\Lambda_n}} e^{J \sum_{\substack{<x,y> \\ x,y\in\Lambda_n}} \eta(x)\eta(y)} e^{\lambda h_1 \sum_{x\in\Lambda_n} \eta(x)} e^{(1-\lambda)h_2 \sum_{x\in\Lambda_n} \eta(x)} .$$

Now apply Hölder's $\le$ to the measure space $\{-1,1\}^{\Lambda_n}$ with $\eta$ being given measure $e^{J \sum_{\substack{<x,y> \\ x,y\in\Lambda_n}} \eta(x)\eta(y)}$. Calling this measure m, the above is then

$$\int_{\{-1,1\}^{\Lambda_n}} e^{\lambda h_1 \sum_{x\in\Lambda_n} \eta(x)} e^{(1-\lambda)h_2 \sum_{x\in\Lambda_n} \eta(x)} dm(\eta)$$

$$\le \left[\int e^{h_1 \sum_{x\in\Lambda_n} \eta(x)} dm\right]^{\lambda} \left[\int e^{h_2 \sum_{x\in\Lambda_n} \eta(x)}\right]^{1-\lambda}$$

$$= \left[\sum_{\eta} e^{h_1 \sum_{x\in\Lambda_n} \eta(x)} e^{J \sum_{\substack{<x,y> \\ x,y\in\Lambda_n}} \eta(x)\eta(y)}\right]^{\lambda} \left[\sum_{\eta} e^{h_2 \sum_{x\in\Lambda_n} \eta(x)} e^{J \sum_{\substack{<x,y> \\ x,y\in\Lambda_n}} \eta(x)\eta(y)}\right]^{1-\lambda}$$

$$= Z_n(J,h_1)^{\lambda} Z_n(J, h_2)^{1-\lambda}.$$

QED

**Proof of Theorem C**

**Lemma 1.** $\forall n$, $\forall$ boundary conditions $\delta$,

$$\frac{\partial}{\partial h} P_n^{J,h,\delta} = E_{J,h_{\mu_n^{\delta}}} \left[\frac{\sum_{x\in\Lambda_n} \eta(x)}{|\Lambda_n|}\right].$$

(called the expected value of the magnetism).

**Proof.** $$\frac{\partial}{\partial h} P_n^{J,h,\delta} = \frac{\partial}{\partial h} \frac{\log {}^{J,h}Z_n^\delta}{|\Lambda_n|}$$

$$= \frac{\left(\frac{1}{|\Lambda_n|} \sum_\eta \left[\sum_{x\in\Lambda_n} \eta(x)\right] e^{\left[J \sum_{\substack{<x,y>\\ x,y\in\Lambda_n}} \eta(x)\eta(y) + J \sum_{\substack{<x,y>\\ x\in\Lambda_n\\ y\in\partial\Lambda_n}} + h \sum_{x\in\Lambda_n} \eta(x)\right]}\right.}{{}^{J,h}Z_n^\delta}$$

$$= E_{J,h_{\mu_n^\delta}} \left[\frac{\sum_{x\in\Lambda_n} \eta(x)}{|\Lambda_n|}\right]$$

QED

**Lemma 2.** $\lim_{n\to\infty} \frac{\partial}{\partial h} P^+(n,J,h) = \mu_{J,h}^+ (\eta(0) = 1)$.

**Proof.** By lemma 1, we need to show

$$E_{\mu_{J,h,n}^+} \left[\frac{\sum_{x\in\Lambda_n} \eta(x)}{|\Lambda_n|}\right] \to \mu_{J,h}^+ [\eta(0) = 1].$$

One has to be a little careful here - I will leave it as an exercise. The main idea is that using monotonicity (the corollary of Holley's theorem), one shows that when $n$ is large, most (NOT ALL!!) $x \in \Lambda_n$ satisfy

$$E_{J,h_{\mu_n^+}} [\eta(x)] \in [\mu_{J,h}^+[\eta(0) = 1],\ \mu_{J,h}^+[\eta(0) = 1] + \epsilon].$$

QED (sort of)

We now prove theorem (C) but only one direction, which is the direction we need to complete Theorem 4(c). Namely $\frac{\partial}{\partial h} P(J,h)\Big|_{h=h_0}$ exists $\Rightarrow |G^{J,h_0}| = 1$.

We need a lemma from elementary real analysis.

**Lemma.** Let $f_n(\lambda)$ be a sequence of differentiable convex functions on an interval $I \ni \{0\}$. Assume $f_n(\lambda) \to f(\lambda)$ $\forall \lambda$ and $f'(0)$ exists. Then $\lim_{n\to\infty} f_n'(0) = f'(0)$.

**Proof.** Left as an exercsie.

**Proof.** $P^+(n,J,h) \xrightarrow{n\to\infty} P(J,h)$. If $\frac{\partial P(J,h)}{\partial h}\Big|_{h=h_0}$ exists, the above lemma $\Rightarrow$

$\frac{\partial P^+}{\partial h}(n,J,h) \xrightarrow{n\to\infty} \frac{\partial P^+(J,h)}{\partial h} \Rightarrow$ (lemma 2) $\frac{\partial P^+(J,h)}{\partial h} = \mu^+_{J,h}(\eta(0) = 1)$.
Similarly, $\frac{\partial P^-(J,h)}{\partial h} = \mu^-_{J,h}(\eta(0) = 1) \Rightarrow \mu^-_{J,h} = \mu^+_{J,h} \Rightarrow |G_{J,h}| = 1$. QED

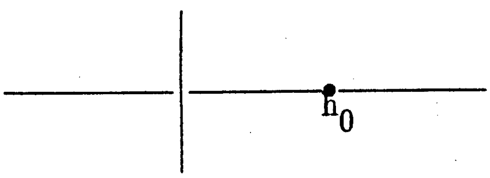
**Exercise:** In the last line of this proof, we assumed

$$\mu^-_{J,h}[\eta(0) = 1] = \mu^+_{J,h}[\eta(0)=1] \Rightarrow \mu^-_{J,h} = \mu^+_{J,h}.$$

Why is this true?

There are 2 approaches to the proof of Theorem D. The first approach will be discussed <u>very</u> roughly. The first step is to allow the external field to be complex. Then ${}^{J,h}P_n^{\delta}$ is a complex analytic function of $h$ $\forall n$. Fix $h_0 > 0$.

* One then shows that $\exists$ some neighborhood of $h_0$ in $\mathbb{C}$.




where the functions

$\{{}^{J,h}P_n^f\}$ are a normal family. Before telling you what a normal family is, one should be aware that fact * takes some work to show but once one has that the proof is easy. "Normal family" is a compactness criterion in a certain function space. We all know that in $\mathbb{R}^n$, being closed and bounded is the compactness criterion.

Ascoli's theorem gives a compactness criterion for C[0,1], the space of continuous functions on [0,1], namely, closed, bounded, and equicontinuity.

For analytic functions, the correct convergence is uniform convergence on compact sets. "Normality" then just means compactness (or actually pre-compactness) in the space of analytic functions on a domain with the above convergence. One actually allows the limit to be $\infty$. One shows that $\exists$ an open neighborhood $\mathcal{O}$ of $h_0$ and a constant $C \ni$

$$\mathrm{Re}\,{}^{J,h}P_n^f(z) \le C \quad \forall n\ \forall z \in \mathcal{O}$$

This is known to imply that $\{{}^{J,h}P_n^f(z)\}$ is a normal family and hence has a convergence sequence. Since on $\mathcal{O} \cap$ {Real line}, the sequence converges to a finite #, the limit cannot be $\equiv \infty$.

Hence some subsequence of ${}^{J,h}P_n^{\delta}(z)$ converges uniformly on compact sets in $\mathcal{O}$ to some h(z). h(z) must therefore be analytic but on the other hand must agree with P(J,h) on $\mathcal{O} \cap$ {Real line}. Hence P(J,h) has an analytic extension to some complex neighborhood of $h_0$ which certainly implies the desired differentiability.

Remember that proving this normality is some work and uses some thing called the Lee-Yang Circle Theorem.

The second approach is as follows.

Lemma 1. $\Rightarrow$
$$\frac{\partial}{\partial h} P^f(n,J,h) = E_{J,h}^{\mu_n^f}\left[\frac{\sum_{x\in\Lambda_n} \eta(x)}{|\Lambda_n|}\right]$$

which we write as M(n,h). Hence

$$* \quad P^f(n,J,h) - P^f(n,J,0) = \int_0^h M(n,s)ds.$$

We will skip the following lemma which is proven from the so-called GHS $\leq$ which is definitely non-trivial.

**Lemma.** $M(n,s)$ is for fixed $n$ concave in $s$.

Since $-1 \leq M(n,s) \leq 1$, using a standard diagonalization argument, $\exists\, n_k \to \infty \ni \forall$ rational $s \in [0,h]$,

$$M(n_k, s) \xrightarrow{k\to\infty} \text{some limit which we call } M(s).$$

Since each $M(n,s)$ is convex in $s$, a geometrical argument implies that we have that $M(n_k,s)$ converges $\forall s$. Hence $M(n_k,s) \xrightarrow{k\to\infty} M(s)$ $\forall s$.

We let $n \to \infty$ along the sequence $n_k$ in $*$ using bounded convergence on the RHS to conclude

$$P(J,h) - P(J,0) = \int_0^h M(s)ds.$$

Since $M(n_k,s)$ is concave $\forall k$, $M(s)$ is concave on $[0,h]$ and hence continuous in $(0,h)$. Since $P(J,h)$ can be represented by an integral of a continuous function on $(0,h)$, and $h$ is arbitrary, we conclude that $P(J,s)$ is differentiable in $s$ $\forall\, s > 0$, as we wanted to show. QED

Note: $M(s)$ need not be continuous at $0 \Rightarrow P(J,s)$ need not be differentiable at $0$. In fact if $J$ is large, we know $P(J,s)$ is not differentiable at $0$ and so $M(s)$ is not continuous at $0$.