03/19/96.

Definition. Suppose f is a random variable on (X, μ) taking values in a finite set, so that $(..., f \circ T^{-1}, f, f \circ T, ...) = (f)^{\infty}_{-\inf ty}$ is a stochastic process.

Let $(A_1, ..., A_n)$ be the partition of X associated with f. The **entropy** of the f-process is

$$h(f) = \int_X \sum_{i=1}^{n} \mu(A_i|\Im') \log_2 \frac{1}{\mu(A_i|\Im')} d\mu(x)$$

where $\Im' = B((f^{-1}_{-\infty}))$.

(Note: this is also equal to $\int_X \sum_{i=1}^{n} \chi_{A_i}(x) \log_2 \frac{1}{\mu(A_i|\Im')} d\mu(x)$, pp.235-236 Peterson)

Last time we took $X = (1/2, 1/2)$ the i.i.d. process.
Ex.1 $f_0(x) = x_0 :\ h(f_0) = \log_2 2 = 1\,\text{bit}$.
Ex.2 $f_1(x) = 2x_0 + x_1 :\ h(f_1) = 1$ bit.
Ex.3 $f_2(x) = x_0 + x_0 x_1 :\ h(f_2) = 1$ bit.
f_0, f_1, f_2 are generating random processes.
Ex.4 $f_3(x) = x_0 x_1 :\ h(f_3) < 1$ bit.
f_3 is not a generator.
Ex.5 $f_4(x) = x_0 + x_1 :\ h(f_4) = 1$ bit.

Today's Theorem. (Kolmogorov-Sinai). If g is $B((f)^{\infty}_{-\infty})$ measurable, then $h(g) \leq h(f)$.

Consequence. 1. If $B((f)^{\infty}_{-\infty}) = \Im$ (the full σ-algebra), then $h(f') \leq h(f)$ for all f'.
2. If $B((f)^{\infty}_{-\infty}) = \Im = B((f')^{\infty}_{-\infty})$ then $h(f') = h(f)$.

Definition. The **entropy** of a dynamicalsystem $(X, T, \Im, \mu)$ is the common value of $h(f)$, where f is any generator (i.e. any discre te r.v. such that $B((f)^{\infty}_{-\infty}) = \Im$), it is also equal to $\sup_f h(f)$. This bound is achieved by Krieger's theorem. It is manifestly an isomorphism invariant.

Outline of proof: for all $\epsilon > 0$ there exists N such that

$$h(g) - \epsilon \leq h(\bar{g}) \leq h((f)^{N}_{-N}) = h(f)$$

where $\bar{g}$ is some $(f)^{N}_{-N}$measurable function approximating g.
Why is equality true?

$$\begin{aligned} h((f)^{-1}_{-\infty}) &= H((f)^{N}_{-N}|((f)^{N}_{-N})^{-1}_{-\infty}) = H((f)^{N}_{-N}|(f)^{N-1}_{-\infty}) \\ &= H((f)^{N-1}_{-N}|(f)^{N-1}_{-\infty}) + H((f)^{N}|(F)^{N-1}_{-\infty}) \end{aligned}$$

But the first term in the sum is zero and the second is $H(f^0|(f)_{-\infty}^{-1}) = h(f)$. Note: $h_n \stackrel{\text{def}}{=} H(f|(f)_{-\infty}^{-1})$ decreases towards $h(f) = H((f)_{-\infty}^{-1})$. Hence,

$$\frac{1}{n}\sum_{k=0}^{n-1} h_k = \frac{1}{n}\sum_{k=0}^{n-1} H((f)^k)|(f)_0^{k-1}) = (\text{stationarity}) = \frac{1}{n}H((f)_0^{N-1})$$

decreases to $h(f)$ as well.

Since $H((f)_0^{N-1})$ is monotone in f (under the refinement), this proves right inequality.

Why left inequality true? For all x such that $g(x) \neq \bar{g}(x)$ let

$\delta(x) = g(x) - \bar{g}(x) mod(m)$, where we assume $range(g) = range(\bar{g}) = \{1, 2, ..., m\}$ and let $\delta(x) = 0$ otherwise.

We have $h(g) \leq h(g\, and\, \bar{g}) = h(\bar{g}\, and\, \delta) \leq h(\bar{g}) + h(\delta)$, and similarly, $h(\bar{g}) \leq h(g) + h(\delta)$.

Two important facts are used:

Prop. 2.11. If $B_n \to B_\infty$ and α is a finit e B_∞ measurable partition, then $H(\alpha|B_n) \to H(\alpha|B_\infty)$.

Prp. 2.13 For counable measurable partitions α, β

$$h(\beta) \leq h(\alpha) + H(\beta|\alpha)$$

Conclusion: two i.i.d. dynamical systems cannot be isomorphic unless the distribution vectors have equal (H) entropy.

Reading assignment: chapter 5, see 6.2.

03.21.96.

A measurable *partition* of X is a collection of pairwise disjoint measurable sets with union X

$P = \{P_1, ..., P_r\}$. Given two partitions P and Q define

$$P \cup Q = \{P_1 \cap Q_1, P_1 \cap Q_2, ..., P_r \cap Q_s\}$$

$$T^{-n}P = \{T^{-n}P_1, ..., T^{-n}P_r\}$$

Write the σ-algebra generated by $T^{-n}P$, $n \leq -1$by $\bigvee_{n\leq -1} T^{-n}P$ and $h(P) = H(P|\bigvee_{n\leq -1} T^{-n}P)$.

Restate claim: if Q is $\bigvee_{n=-\infty}^{\infty} T^{-n}P$ measurable, then $H(Q) \leq h(P)$.

Getting comfortable with entropy: reduce claims about σ-algebra to claims about partitions via Martingale Convergence Theorem (e.g.(2.3)).

Pinsker Formula.

$$H(P\bigvee Q|(P\bigvee Q)_{-\infty}^{-1}) = H(Q|Q_{-\infty}^{-1}) + H(P|P_{-\infty}^{-1}\bigvee Q_{-\infty}^{-1})$$

Define $P_a^b = \bigvee_{n=a}^b T^{-n}P$.

In general $H(Q|P_{-\infty}^{-N})$ does not converge to 0, $H(Q|Q_{-\infty}^{-1}\bigvee P_{-\infty}^{-N}) \to 0$. The entropy of a discrete stationary process (X, T, μ, P) is zero iff P is $\bigvee_{-\infty}^{-1} T^{-n}P$ measurable, modulo μ, (P countable).

E.g. the entropy of irrational rotation is zero.

A **joining** of $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ is a measure π on $(X\times Y, \mathcal{A}\times\mathcal{B})$ such that $\pi(A\times Y) = \mu(A)$, $\pi(X\times B) = \nu(B)$ for all $A \in \ calA, B\in\mathcal{B}$.

Note: we do not assume $\pi(A\times B) = \mu(A)\times\nu(B)$.

Ex.1 $\mu(diag) = 1, \mu(rest) = 0$.

2. Product measure

3. Given $f: X\to Y$, m.p., put $\pi(A\times B) = \mu(A\cap f^{-1}(B))$.

Given probability measures on a σ-algebra generated by some countable collection $C_1, C_2, \ldots$ we will say $\mu_n \to \nu$ if $\mu_n(C_k)\to\nu(C_k)$, for al l k. Equivalently,

$$d(\mu_n, \nu) = (def) \sum_{k=1}^{\infty} \frac{|\mu_n(C_k) - \nu(C_k)|}{2^k} \to 0$$

Example: take μ_n - uniform on $\{0, 1/n, 2/n, \ldots, n-1/n\}$ in $[0,1)$. $\mu_n \to \nu$- Lebesgue measure.

04.18.96

Transition to interacting particle systems via single-particle systems and non-interacting many particle systems.

Discrete Markov Chains.

Fix a finite set S. Let $P = (p(x,y))_{x,y\in S}$ be a stochastic matrix, i.e. $p(x,y)\geq 0, \forall\, x, y\in S,\ \sum_{y\in S} P(x,y) = 1$. We write the entries of P^m as $p^{(m)}(x,y)$.

Theorem. Suppose that for some $n\geq 1,\ p^{(n)}(x,y) > 0,\ \forall\, x,y\in S$. Then $\pi(y) = (def) = \lim_{n\to\infty} p^{(n)}(x,y)$ exists for $x,y\in S$, is independent of x and satisfies

$$\sum_x \pi(x)p(x,y) = \pi(y) \tag{1}$$

Proof: by coupling argument. Let (X_n, Y_n) be a Markov chain on $S \times S$ where

1. X_n and Y_n evolve independently with transition probabilities $p(\ ,\ )$, untill the first time $X_n = Y_n$.

(b) X_n and Y_n evolve identically thereafter.

Put $\epsilon = \min_{x,y} p^{(m)}(x,y) > 0$. Then

$$Pr(X_m = Y_m | X_0 = x, Y_0 = y) \geq \sum_{z \in S} p^{(m)}(x,z) p^{(m)}(y,z) \geq$$

$$\geq \epsilon \sum_{z \in S} p^{(m)}(y,z) = \epsilon > 0$$

By the Markov chain property and stationarity

$$Pr(X_{2m} \neq Y_{2m} | x_0 = x, y_0 = y) \leq$$

$$\sum_{x' \neq y'} Pr X_{2m} \neq Y_{2m} | X_m = x', Y_m = y') Pr(X_m = x', Y_m = y' | X_0 = x, Y_0 = y) \leq$$

$$(1-\epsilon) \sum_{X' \neq Y'} Pr(X_m = x', Y_m = y' | X_0 = x, Y_0 = y) \leq (1-\epsilon)^2$$

More generally

$$Pr(X_{km} \neq Y_{km} | X_0 = x, Y_0 = y) \leq (1-\epsilon)^k$$

Hence

$$Pr(X_n \neq Y_n | X_0 = x, Y_0 = y) \rightarrow 0$$

Since X and Y processes evolve under P

$$|p^{(n)}(u,y) - p^{(n)}(v,y)| =$$

$$= |P(X_n = y | X_0 = u, Y_0 = v) - P(Y_n = y | X_0 = u, Y_0 = v)| =$$

$$= |P(X_n = y \neq Y_n | X_0 = u, Y_0 = v) - P(X_n \neq Y_n = y || X_0 = u, Y_0 = v)| \leq$$

$$\leq |P(X_n = y \neq Y_n | X_0 = u, Y_0 = v)| + |P(X_n \neq Y_n = y || X_0 = u, Y_0 = v)|$$

$$\leq P(X_n \neq Y_n | X_0 = u, Y_0 = v) \to 0$$

(like $(1-\epsilon)^{\frac{n}{m}}$.)

Hence

$$|\max_x p^{(n)}(x,y) - \min_x p^{(n)}(x,y)| \to 0$$

Since $\max_x p^{(n)}(x,y)$ decreases and $\min_x p^{(n)}(x,y)$ increases, $p^{(n)}(x,y)$ tends to some limit $\pi(y)$ as $n \to \infty$.

Verifying (*) is easy: $vP^(n) \to v_\infty$ implies $v_\infty P = v_\infty$, since

$$(VP^{(n)})P = vP^{(n+1)}$$

Corollary of proof: under hypothesis of the previous theorem, there is only one measure $\pi(\ )$ on S satisfying $\sum_x \pi(x)p(x,y) = \pi(y),\ \forall\, y \in S$.

π is both a fixed point of $\mu \to \mu P = \sum \mu(x)p(x,\ )$ and an attractor of this map.

Exchangeable processes. Let I be some index set, and $(X_i)_{i\in I}$ a family of random variables. $(X_i)_{i\in I}$ is exchangeable if for all finite r, the joint distribution of $X_{i_1}, ..., X_{i_r}$ (with $i_1, ..., i_r$ distinct) is independent of $i_1, ..., i_r$.

Example.

$(X_i)_{i\in I}$ i.i.d. (coin model).

$X_i \in \{0,1\}$ for all i, $\sum_{i\in I} X_i$ fixed. (urn model)

For the rest of today we will assume that X_i-s are $0,1$ valued with I finite or countable.

Claim. If μ_1, μ_2 are measures on $\{0,1\}^I$ that define exchangeable processes, then so does $t\mu_1 + (1-t)\mu_2$ for any $0 \leq t \leq 1$. More generally, any superposition of exchangeable processes is exchangeable. So the set of exchangeable measures on $\{0,1\}^I$ is *convex.*

04/23/96

Recall: if I is an index set and $(X_i)_{i\in I}$ a family of random variables, $(X_i)_{i\in I}$ is exchangeable if for all $r < \infty$ the joint distribution of $X_{i_1}, X_{i_2}, ..., X_{i_r}$ $(i_1, ..., i_r$ distinct) is independent. Assume $I = \{1, 2, ..., n\}$ or Z, X_i-s $\{0,1\}$ valued. Give $\{0,1\}^I$ the product topology.

Claim. The set of exchangeable measures is convex and compact. Consequently (by Krein-Milman) every exchangeable measure on $\{0,1\}^I$ can be written as a superposition of indecomposable exchangeable measures on $\{0,1\}^I$.

If I is finite, then the indecomposable measures are the $|I|+1$ "urn measures" (uniform measure on the set of $|I|$ tuples $(X_1, ..., X_{|I|})$ subject to $X_1+...+X_{|I|} = k$ for a fixed $0 \leq k \leq |I|$.

If I is infinite, then the indecomposable measures are i.i.d. measures (parametrized by $(p \in [0,1])$ **(De-Finetti).**

Proof of De-Finetti: "Take 1." It suffices to show for each $0 \leq p \leq 1$ there is a unique indecomposable exchangeable process $(X_i)_{i \in Z}$ with $Pr(X_i = 1) = p$, (i.e. unique indecomposable measure μ_p on $\{0,1\}^Z$ with $\mu(\{x : x_0 = 1\} = p)$.

Given $X = \{0,1\}^I$ let b_i, $i \in Z$ be independent random bits, and generate $X^* \in \{0,2\}^Z$ by putting

$$(x^*_{2i}, x^*_{2i+1}) = (x_{2i}, x_{2i+1}) \;\; if \;\; b_i = 0,$$

and

$$(x^*_{2i}, x^*_{2i+1}) = (x_{2i+1}, x_{2i}), \;\; if \;\; b_i = 1$$

Example: let $x = (10)(01)(00)(11)$, $b = 0, 1, 1, 0$.

Then $x^* = (10)(10)(00)(11)$.

x^* is a random $\{0,1\}^Z$-valued random variable, and hence it is best described via probability distribution ν_x on $\{0,1\}^Z$. More generally, given a measure μ on $\{0,1\}^Z$ we can define a randomly permuted version of μ, $\mu^* = (def) = \int \nu_x d\mu(x)$.

Claim. If μ is exchangeable, $\mu^* = \mu$.

Proof: it suffices to consider individual strings $(x_{2i}, x_{2i+1}, ..., x_{2j-1})$, (that is the associated cylinder set C), and prove $\mu^*(C) = \mu(C)$. It is easy to check $\mu^*(C)$ is the average of measures $\mu(D)$ where D-s are cylinder sets of measure equal to $\mu(C)$. So $\mu^* = \mu$.

One can similarly define $\mu \to \mu_*$ by randomization of the pairs $x_{2i-1,2i}$ Instead of x_{2i}, x_{2i+1}. If μ is exchangeable, $\mu_* = \mu$.

04/25/96

De-Finetti:

$$(10)(01)(00)(11)$$

$$0(11)(00)(01)1$$

$$(01)(10)(01)(01)$$

Statistic will converge to i.i.d.

Assume μ is indecomposable exchangeable measure on

$\{0,1\}^Z$. Assume $\mu(\{x : x_0 = 1\}) = p$. Prove $F^n\mu$ converges in weak-* topology to i.i.d measure with parameter p. Here $F : \mu \to (\mu^*)_*$ (switched upper and lower).

Tactic: assume μ_1 and μ_2 are indecomposable exchangeable with parameter p. If we can prove $F^n\mu_1$ and $F^n\mu_2$ get arbitrarily close as $n \to \infty$ and we know $F^n\mu_1 = \mu_1$, $F^n\mu_2 = \mu_2\ \forall\, n$ then $\mu_1 = \mu_2$.

Dicrepency reduction:

Process 1:

$$A : \underbrace{(10)}_{0}\underbrace{(01)}_{1}\underbrace{(00)}_{1}\underbrace{(11)}_{0}$$

$$0(11)(00)(01)1$$

$$(01)(10)(01)(01)$$

Process 2

$$A' : \underbrace{(01)}_{0}\underbrace{(10)}_{1}\underbrace{(11)}_{1}\underbrace{(00)}_{0}$$

$$0(11)(01)(10)0$$

$$(01)(10)(11)(00)$$

Bit 0 means put in ascending order, 1- in descending order.

04/29/96.

De-Finetti resolved.

Background: every exchangeable measure on $X - \{0,1\}^Z$ is

a) shift invariant

b) invariant under "exchange dynamics." (measure-to-measure map).

Let E be the set of exchangeable measures on X (a compact convex space under the weak-* topology).

E_e - the set of extreme points of E.

If $\mu \in E_e$, μ is ergodic.

Proof: if C is a cylinder set, with indicator function f, let

$$\bar{f}(x) = (def) = \lim_{n\to\infty} \sum_{i=0}^{n-1} f(T^i x)$$

The limit is well defined since T is shift-invariant.

If $A \subset R$ with $\mu(\{x : barf(x) \in A\}), \mu(\{x : barf(x) \notin A\}) > 0$, then we can write μ as a weighted average of conditional measures, both exchangeable (condition on $x \in A, x \notin A$). This contradicts extremality of μ. Hence $\bar{f}()$ is a.s. constant. Since this is true for all the cylinder sets, it is true for all the measurable sets, so μ is ergodic.

Claim. E_e equals the set of i.i.d. measures. **(De-Finetti)**

This reduces to

Claim. If $\mu_1, \mu_2 \in E_e$ with $\mu_1(\{x : x_0 = 1\}) = p = \mu_2(\{x : x_0 = 1\})$ then $\mu_1 = \mu_2$.

Proof: let π be a shift-invariant joining of μ_1, μ_2 on the product space $\{0,1\}^Z \times \{0,1\}^Z$ that minimizes

$$\pi((x, y) : x_0 \neq y_0)$$

Since π is a convex combination of its ergodic components π_i, then almost all of these components also achieve that minimum. In addition the reduction of π on the first component of the product $\{0,1\}^Z \times \{0,1\}^Z$, which is by assumption μ_1, is a convex combination of the reductions of π_i-s - the ergodic components of π. By the previous claim μ_1 is ergodic, so all the reductions are equal to μ_1. Similarly we show that the reductions of π_i-s on the second component of the product $\{0,1\}^Z \times \{0,1\}^Z$ are all equal to μ_2. We conclude that all the ergodic components π-s of the measure π are in fact the joinings of μ_1 and μ_2. Let us, then replace π with any of its ergodic components π_i that achieves the minimum .

Suppose $\pi(\{(x, y) : x_0 \neq y_0\}) = \delta > 0$. Since with probability 1 (modulo $\pi(x, y)$) both x and y contain 1's with density p (by ergodicity), the sets $\{i : x_i = 0, y_i = 1\}, \{i : x_i = 1, y_i = 0\}$ have density $\delta/2$ for all i almost surely. It follows that for some m

$$\pi(\{(x, y) : x_0 = 0, y_0 = 1, x_m = 1, y_m = 0\}) > 0$$

But now setting $\pi_0 = \pi$ and letting π_i denote π_0 evolve forward for i steps under the coupling described last time (which decreases the number of discrepencies) we find that for a suitable n, the π_n probability of $(\frac{0}{1})$ initially at zero and $(\frac{1}{0})$ initially at m annihilating each other within n steps is > 0. Applying translation invariance, we see that the density of discrepency under π_n is a.s.$< \delta$, contradicting the choice of π_0. So $\delta = 0$ and $\mu_1 = \mu_2$.

Continuous Time Processes.

Poisson process (also exponential process).

Markov Process. Transition probability from i to j (at times in $\frac{1}{n}Z$) is given by

$$\frac{1}{n}r_{ij},\ \ if\ \ i \neq j$$

$$1-\sum_{k\neq i}\frac{1}{n}r_{ik}\ \ if\ \ i=j.$$

Send $n \to \infty$. Let $f_i(t)$ =probability that system is in state i at time t. Then

$$\dot{f}_i(t) = \sum_{j\neq i} f_j(t)r_{ji} - r_{ij}f_i(t)$$

05/01/96

Markov process: evolution of probability law is governed by

$$f_i'(t) = \sum_{j\neq i}[f_j(t)r_{ji} - f_i(t)r_{ij}],$$

(r_{ij} - rate), where $f_i(t)$=probability that the system is in state i at time t. Note if $\sum_{j\neq i} r_{ij} = \sum_{j\neq i} r_{ji}$, for all i, then the steady-state distribution is uniform.

We will assume that some power of $(r_{ij})_{i,j=1}^n$ is strictly positive. This implies uniqueness of stationary measure.

How to apply Markov process to IPS's?

Approach 1. General abstract approach.(Liggett, Ch 1).

Let us do it for the exchange process. For $x \in X \subset \{0,1\}^Z$ and $m \in Z$, define $x^{m\rightarrow}$ and $x^{\leftarrow m}$ in X as follows:

$$x^{m\rightarrow} = \begin{cases} x_i, & \text{for } i < m, i > m+1; \\ x_{i+1}, & \text{for } i = m; \\ x_i, & \text{for } i = m+1. \end{cases}$$

and

$$x^{\leftarrow m} = \begin{cases} x_i, & \text{for } i < m-1, i > m; \\ x_i, & \text{for } i = m-1; \\ x_{i-1}, & \text{for } i = m. \end{cases}$$

Let $D(X)$ be the set of all functions $X \to R$ that depend only on finitely many coordinates. Given $f \in D(X)$ we define $\Omega f : X \to R$ by

$$(\Omega f)(x) = \sum_{m\in Z}[\lambda(f(x^{m\rightarrow}) - f(x)) + \lambda(f(x^{\leftarrow m}) - f(x)]$$

(all but finitely many terms vanish). We get a linear operator $\Omega : D(X) \to D(X)$. This operator can serve as the infinitesimal generator of a semigroup.

Approach 2. Concrete approach: use countable set of independent Poisson timers.

Problem: "Well-definedness."

The evolution leads to proof of De-Finetti theorem.

Exclusion process. Let $p(x, y)$be transition probabilities for a Markov chain on a countable set S, $p(x, y) \geq 0$, $\sum_y p(x, y) = 1$. (Note shift in notation: $x, y \in S$, $\eta \in \{0, 1\}^S$.

A particle at x waits an exponentially distributed time with parameter λ, picks y location with probability $p(x, y)$ and swaps with location y (vacant or not). Given a measure μ on X and $t \geq 0$ we get a measure $\mu_t = \mu$ evolved for a duration t under exclusion dynamics. μ is stationary iff $\mu_t = \mu$ for all $t \geq 0$.

Fact (Prop.I.2.13 Liggett). μ is stationary iff

$$\int \Omega f d\mu = 0$$

for all $f \in D(X)$.

For $\eta \in X = \{0, 1\}^S$ and $x, y \in S$, let $\eta_{xy} \in X$ with

$$\eta_{xy}(u) = \begin{cases} \eta(u), & \text{for } u \neq x, y; \\ \eta(y), & \text{for } u = x; \\ \eta(x), & \text{for } u = y. \end{cases}$$

Then for $f \in D(X)$

$$(\Omega f)(\eta) = \sum_{x,y,\eta(x)=1} p(x, y)[f(\eta_{xy}) - f(\eta)] =$$

$$= \sum_{x,y,\eta(x)=1,\eta(y)=0} p(x, y)[f(\eta_{xy}) - f(\eta)]$$

05/07/96.

$D(X)$ set of functions depending on finitely many coordinates.

Exclusion processes. for $f \in D(X), \eta \in X$,

$$(\Omega f)(\eta) = \sum_{x,y,\eta(x)=1} p(x, y)[f(\eta_{xy}) - f(\eta)] =$$

$$= \sum_{x,y,\eta(x)=1,\eta(y)=0} p(x, y)[f(\eta_{xy}) - f(\eta)]$$

Claim: if $\sum_x p(x,y) = 1$ for all y then the i.i.d. measures are stable under time evolution.

Proof: Recall ν_α is a product measure on $\{0,1\}^S$ with marginals $\nu_\alpha\{\eta : \eta(x) = 1\} = \alpha$, for all x.

It suffices to show

$$\int \Omega f d\nu_\alpha = 0, \ \forall f \in D(X) \tag{2}$$

I.e.

$$\int \sum_{x,y,\eta(x)=1} p(x,y)[f(\eta_{xy}) - f(\eta)]d\nu_\alpha = 0, \ \forall f \in D(X)$$

or

$$\sum_{x \neq y} p(x,y) \int \eta(x) f(\eta_{xy}) d\nu_\alpha =$$

$$= \sum_{x \neq y} p(x,y) \int \eta(x) f(\eta) d\nu_\alpha$$

By linearity it's enough to check 2 for f of the form:

$$f_A(\eta) = \begin{cases} 1, & \text{if } \eta(x) = 1, \ \forall x \in A; \\ 0, & \text{otherwise.} \end{cases}$$

where A is any finite subset of S.

For $y \neq x$

$$\int \eta(x) f(\eta) d\nu_\alpha = \alpha^{|A \cup \{x\}|}$$

and

$$\int \eta(x) f(\eta_{xy}) d\nu_\alpha = \alpha^{|A_{xy} \cup \{x\}|} = \alpha^{|A \cup \{y\}|}$$

Thus we need to check whether

$$\sum_{x \neq y} p(x,y) \alpha^{|A \cup \{y\}|} =$$

$$\sum_{x \neq y} p(x,y) \alpha^{|A \cup \{x\}|}$$

or equivalently

$$\alpha^{|A|}[\sum_{x\neq y, y\in A} p(x,y) + \alpha \sum_{x\neq y, y\notin A} p(x,y)] =$$

$$= \alpha^{|A|}[\sum_{x\neq y, x\in A} p(x,y) + \alpha \sum_{x\neq y, x\notin A} p(x,y)]$$

Note, all pairs (x, y) with $x, y \in A$ or $x, y \notin A$ contribute equally to both sides, so we need only show

$$\sum_{x\notin A, y\in A} p(x,y) + \alpha \sum_{x\in A, x\notin A} =$$

$$= \sum_{x\in A, y\notin A} p(x,y) + \alpha \sum_{x\notin A, y\in A}$$

or equivalently we need to show

$$(1-\alpha) \sum_{x\in A, y\notin A} p(x,y) = (1-\alpha) \sum_{x\notin A, y\in A} p(x,y) =$$

$$= (1-\alpha) \sum_{x\in A, y\notin A} p(y,x)$$

But this is true, since by assumption P is doubly stochastic:

$$\sum_{x\in A, y\notin A} p(x,y) - p(y,x) = 0$$

Two flows in the proof:

1) Change of summation and integration. We should have verified summability of $\sum_{x\in A, y\in A} p(x,y)$ (A - finite set).
Note, terms of the sum

$$\sum_{x,y} p(x,y)[f(\eta_{x,y}) - f(\eta)],\ with\ x, y \in A$$

make no contribution, if $f(\eta)$ depends on $\eta(x)$ for $x \in A$.

2) Separation of summation. We write $\sum(Fred - Joe)$ as $\sum Fred = \sum Joe$ - both can diverge (e.g. $\sum_{x\neq y, x\notin A} p(x,y)f(\eta)$ typically diverges).
To fix this, it is necessary to keep all terms on the sme side of "=".

Example.
$S = Z$

$$p(x,y) = \begin{cases} 1, & \text{if } y = x+1; \\ 0, & \text{otherwise.} \end{cases}$$

Intuition for preservation of ν_α : replace Z by Z/nZ for large n. k particles, $n-k$ vacancies. Then it's a Markov chain with $\binom{n}{k}$ states. Departure rate is

$$|\{(x, x+1) : \eta(x) = 1,\ \eta(x+1) = 0\}|$$

Arrival rate is

$$|\{(x, x+1) : \eta(x) = 0, \eta(x+1) = 1\}|$$

But these are equal since the unmber of "up" transitions is equal to the number of "down" transitions. The rates balance follows. Therefore urn-measure is invariant.

Now send $n \to \infty$, $k/n \to \alpha$. Then the invariance of the urn measures intuitively follows invariance of ν_α, (urn measures $\to \nu_{alpha}$.

We will see later that these are the only invariant measures.

Caveat: this analysis applies only in continuous time.

Discrete time: put $S = Z$ and decree that at each step, a particle with a vacancy to its right moves rightward with probability 1/2. This does not preserve ν_α.

Intuition: work on Z/nZ.

Proof of 2 above: set $p(x,x) = 0$, wlog. Define

$$F(x) = \int \eta(x) f(\eta) d\nu_\alpha(\eta)$$

Need to check

$$\sum_{x,y} p(x,y) \int \eta(x) f(\eta_{xy}) d\nu_\alpha(\eta) =$$

$$= \sum_{x,y} p(x,y) \int \eta(x) f(\eta) d\nu_\alpha(\eta)$$

Claim:

$$\int \eta(x) f(\eta_{xy}) d\nu_\alpha(\eta) = F(y)$$

because ν_α is exchange invariant, therefore we need to check

$$\sum_{x,y} p(x,y) F(y) = \sum_{x,y} p(x,y) F(x)$$

or

$$< PF, 1 >_{L^2(S)} = < P^* F, 1 >_{L^2(S)}$$

where

$$PF(x) = \sum_y p(x,y) f(y)$$

$$P^* F(x) = \sum_y p(y,x) f(y)$$

05/09/96.

Claim. If $S = Z$ and $p(x, x+1) = p(x, x-1) = 1/2$, $\forall\, x$ then the only extremal translation invariant stationary under exclusion dynamics measures on $\{0,1\}^S$ are the product measures ν_α.

But take $S = Z$ and assume discrete time, where at each step an unblocked particle moves to the right w.p. p. Then we can have

system 1. $1(1)00 \to 1010$

system 2. $(1)010 \to 0101$

This creates problem - introduces extra discrepencies.

It is fixable if $p \leq 1/2$, make (1) and (1) move together. You can keep the density of discrepencies from increasing by coupling properly (see the article by Jockusch, Propp, Shor on Web page). If $p > 1/2$ what are translation invariant measures? - open problem.

Theorem. (special case of result on p.403 Liggett). Take $S = Z$, $p(x, x+1) = p(x, x-1) = 1/2$ for all $x \in Z$. Consider the exclusion process $\eta_t(x)$ with initial state given by

$$\eta^0(x) = \begin{cases} 1, & x \leq 0; \\ 0, & x > 0; \end{cases}$$

For $x \in Z$ and $\tau \geq 0$, let μ_t^x be the probability measure on $\{0,1\}^Z$ given by

$$\mu_t^x\{\psi : \psi(x_1) = \ldots = \psi(x_n) = 1\} =$$

$$= Pr\{\eta_t(x + x_1) = \ldots = \eta_t(x + x_n) = 1 | \eta_0 = \eta^0\}$$

If $f(t)/t^{1/2} \to c$ we have

$$\lim_{t\to\infty} \mu_t^{f(t)} = \nu_{\alpha(c)}$$

where

$$\nu_{\alpha(c)} = \frac{1}{(2\pi)^{1/2}\sigma}\int_c^\infty e^{-\frac{v^2}{2\sigma^2}}dv,$$

$\sigma^2 = \sum_x x^2 p(v,x) = 1$, in our case. So

$$\nu_{\alpha(c)} = \frac{1}{(2\pi)^{1/2}}\int_c^\infty e^{-\frac{v^2}{2}}$$

$\nu_{\alpha(c)}$ - product measure, and $1/t^{1/2}\sum_{y\geq ct^{1/2}}\eta_t(y)$ converges in probability to $\int_c^\infty \alpha(u)du$.

05/14/96.

Sub-additive ergodic theorem.
Suppose $S_{m,n}$, $(0 \leq m \leq n)$, are random variables such that

- (a) $S_{00} = 0$, $S_{0n} \leq S_{0m} + S_{mn}$, $0 \leq m \leq n$
- (b) $(S_{0k}, S_{k2k}, S_{2k3k}, ...)$is stationary for each k.
- (c) $(S_{mm}, S_{m+1}, S_{mm+2}, ...)$ has joint distribution independent of m.
- (d) $E[S_{01}^+] < \infty$, where $X^+ = 1/2(X + |X|)$.

Let $\alpha_n = E[S_{0n}]$ which is finite by (a),(b),(d). Then

$$\frac{\alpha_n}{n} \to \alpha = (def) = \inf_{m\geq 1}\frac{\alpha_m}{m}$$

$\frac{S_{0n}}{n}$ converges almost surely, and $E[\lim_{n\to\infty}\frac{S_{0n}}{n}] = \alpha$.
If the stationary processes in (b) are ergodic, then $\lim_{n\to\infty}\frac{S_{0n}}{n} = \alpha$ a.s. (sub-additive ergodic theorem, see Liggett, Ann.Prob.13,1279-1285).

Application 1. The standard ergodic theorem. Let $S_{mn} = X_{m+1} + X_{m+2} + ... + X_n$, $(X_1, X_2, ..., X_n, ...)$ - stationary sequance of r.v.'s with $E[|X|] < \infty$. Then the conclusion is ergodic theorem.

Application 2. Let $a_1, ..., a_n, ...$ - positive reals with $a_{r+s} \leq a_r + a_s$, $\forall\, r, s$. Put $S_{m,n} = a_{n-m}$. Then sub-additivity applies, so

$$\frac{a_n}{n} \to \inf_{m\geq 1}\frac{a_m}{m}$$

("Fekete's Lemma").

Application 3. The continuous time asymetric exclusion model on Z^1 with initial state η_0 where

$$\eta_0 = \begin{cases} 1, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Regard this as growth model.

For $u = (a, b)$ with $a, b \in Z_+$, let T_u be the first time at which both the square to the left of u and square below$u$ are on. (For example $T_{00} = 0$.) For $v = (c, d)$ with $c \geq a$, $d \geq b$ let T_{uv} denote the additional time it would take for v to turn on, after T_u, if we were to give the system a "boost" by manually turning on all squares that lie strictly to the left of u or strictly below u.

Note, that if $w = (e, f)$, $e \geq c$, $f \geq d$,then $T_{uw} \geq T_{uv} + T_{vw}$.

Put $S_{mn} = -T_{mv,nv}$ for some fixed v.

The S.E.Th applies (note $S_{01}S_{12}, S_{23}, ...$ is ergodic). Thus

$$\frac{S_{0n}}{n} \to constant\ \alpha(v)$$

(rate $= \lim \frac{n}{S_{0n}}$).

05/16/96

Boltzman's Law: given a finite physical system that can be in any of the N states η and is in thermal equilibrium with its enviroment, the probability of it being in any of the particular state η is proportional to $e^{-E(\eta)/kT}$, where T is absolute temperature, k - Boltzman's constant. Put $\beta = \frac{1}{kT}$, $Z = \sum_\eta e^{-\beta E(\eta)}$. Then

$$Prob(\eta) = \frac{e^{-\beta E(\eta)}}{Z}$$

Variational principle: given any probability measure μ on configuration space X, let

$$S(\mu) = -\sum_{\eta \in X} \mu(\{\eta\}) log \mu(\{\eta\})$$

and

$$E(\mu) = \sum_{\eta \in X} \mu(\{\eta\}) E(\eta)$$

Then maximum of $S(\mu) - \beta E(\mu)$ (Helmoholtz' free energy) over all probability measuresμ on X is $log Z$ with

$$\mu(\eta) = \frac{e^{-\beta E(\eta)}}{Z}$$

for all η.

Example 1. A finite spin system in states in $X = \{0,1\}^n$. Given $\eta \in X$, let $E(\eta) = \sum_{i=1}^n \eta(i)$ = number of 1's in $(\eta(1), ..., \eta(n))$. The Boltsmanz' distribution is a biased i.i.d distribution. Put an exponential timer on each of the n sites and if the timer at the site i goes off at time t, let $\eta_{t^+}(i) = 1$ (resp. 0) with probability p (resp. $1-p$), where

$$\frac{p}{1-p} = \frac{\mu(\eta^1)}{\mu(\eta^0)} = e^{-\beta}$$

where μ - Boltzmann's distribution and

$$\eta^r(j) = \begin{cases} \eta_{t^-}(j), & j \neq i; \\ r, & j = i. \end{cases}$$

Example 2. The Ising model on a finite graph G, with vertex set $\{1, 2, ..., n\}$ and bond set B. Let $X = \{+1, -1\}^n$,

$$E(\eta) = - \sum_{1 \leq i \leq j,\, \{ij\} \in B} \eta(i)\eta(j)$$

Heat-bath dynamics: use exponential timers at the sites and update them in accordance with $e^{-\beta\delta E}$. Under these dynamics, the law of the system at time t converges to the Boltzmann's distribution as $t \to \infty$.

Example 3. The Gaussian model. Let $X = R$, $E(x) = x^2$. $e^{-x^2/kT}$ is integrable and the Boltzmann's distribution is a mean-zero Gaussian. This generalizes to $X = R^n$, $E(\bar{x}) = ||x||^2$. More interestingly, let G be a finite graph as in Ex.2. Let $X = R^n$ and put

$$E(\eta) = \sum_{1 \leq i \leq j \leq n,\, (ij) \in B} (\eta(i) - \eta(j))^2$$

There is no way to make this integrable
(e.g. G - 2-node-1-edge, $\int\int e^{-(x-y)^2} dxdy = \infty$,) unless we mod out by the global gauge symmetry (replacing η by $\eta + c$ does not affect $E(\eta)$). Formally, we are passing to sub-σ-algebra on R^n, in which only the differences $\eta(i) - \eta(j)$ (and the functions expressible in terms of them) are measurable.

Equivalently, we are switching to a model variables $\eta(ij)$ ($\eta(ij) = \eta(i) - \eta(j)$, satisfying $\eta(ij) + \eta(jk) = \eta(ik)$).

The Boltzmann's distribution for "quotient model" is then multi-dimensional Gaussian.

Now add dynamics with exponential timers. If a timer at site i goes offat time t, replace $\eta(i)$ by the average of η values of its neighbors plus mean-zero Gaussian. Under this dynamics Law(t) $\to$ Boltzmann's Law.

Example 4. An infinite lattice model. Return to example 1, but now $X = \{0,1\}^Z$. Assume μ is translation invariant measure on X. Normalizing S and E from example 1 (by dividing by n) we get quantities

$$h(\mu) = \lim_{n\to\infty} \frac{1}{n} S(\mu_n)$$

$$e(\mu) = \lim_{n\to\infty} \frac{1}{n} E(\mu_n)$$

where μ_n restriction of μ to strings og length n.

h - asymptotic entropy of μ (March) $= \lim \frac{1}{n} H(\cap_{i=0}^{n-1} T^{-i} P)$,

e - average asymtpotic density of 1's.

One can show that i.i.d. measures with density of 1's equal to $e^\beta/(e^\beta+1)$ is unique μ that maximizes $h(\mu) - \beta e(\mu)$.

A subshift of finite type (SFT) is the set of all $\eta \in S^Z$ (S finite) such that $(\eta(i), \eta(i+1)) \in G$ for all $i \in Z (G \subset Z \times Z)$.

For ex.: if $S = \{0,1\}$ and $G = \{(00)(01)(10)\}$, then the SFT is the set of all $(..., \eta(-1), \eta(0), \eta(1), ...)$ such that $\eta(i)\eta(i+1) = 0$ for all i. Denote this by X. Note that the shift-map T sends X to itself.

Claim. if μ is shift-invariant measure on X, then $h(\mu, T) \leq log \frac{1+5^{1/2}}{2}$ and there is a unique μ that achieves this bound.

05/21/96.

Let $\Lambda = Z^d$, $(d \geq 1)$, S - finite set. For $\eta \in S^\Lambda$ and $x \in \Lambda$, define $\tau^x(\eta) \in S^\Lambda$ via $(\tau^x\eta)(y) = \eta(x+y)$ for all $y \in \Lambda$. We call (S^Λ, τ) the *full d-dimensional shift* on S. A d-dimensional SFT is the set X of all $\eta \in S^\Lambda$ such that $\tau^x\eta \in C$ for all $x \in \Lambda$ (with C fixed cylinder set).

Example. $d = 1$, $S = \{0,1\}$, $C = \{\eta \in \{0,1\}^Z : \eta(0)\eta(1) = 0\}$ - the golden-mean subshift.

The topological entropy $h(X,T)$ of such a system is

$\lim_{n\to\infty} \frac{1}{n^d} N_n$ where N_n is the number of distinct patterns that arise by restricting $\eta \in X \subset S^\Lambda$ to map from F_n to S where $F_n = [1,n] \times ... \times [1,n]$.

E.g. the topological entropy of the golden-mean subshift is

$$\lim \frac{1}{n} \log(n - th\ Fibonacci\ number) = \log(\frac{1+5^{1/2}}{2})$$

Given a τ-invariant measure μ on X the *measure-theoretic (metric) entropy*

$$h(X, \tau, \mu) = \lim_{n\to\infty} \frac{1}{n^d} H(P_n)$$

where P_n is the partition of X such that η, η' belong to the same component of P_n iff $\eta(x) = \eta'(x)$, $\forall\, x \in F_n$.

Variational Principle.

$$h(X,\tau) = \sum_{\mu\, inv} h(X,\tau,\mu)$$

and equality is achieved by some μ.

Parry's Theorem. If $d = 1$ and the topological Markov chain associated with X is irreducible, then μ is *unique.* In fact this μ is "conditionally uniform."

A measure μ on $X \subset S^{\Lambda}$ (finite type) is conditionally uniform if for any finite patch F and for a.e. $\eta_0 \in X$, the conditionl distribution of μ on the (finite) set of all η-s that agree with η_0 outside of F is uniform.

Proposition. If $h(X,\tau,\mu) = h(X,\tau) > 0$, then μ must be conditionally uniform.
(The converse is false, as we will see).

Variational Principle. Let I be the set of σ-invariant measures on X, and let $e : X \to R$ be some (local) energy function that depends only on finitely many coordinates. For any $\mu \in I$, let

$$e(\mu) = \int_X e(x)d\mu(x) = \int_X \bar{e}(x)d\mu(x)$$

where $\bar{e}(x)$ = the ergodic average of e τ-invariant. Then $\sup_{\mu\in I}[h(\mu) - \beta e(\mu)]$ = an inherently topological quantity, called the *topological pressure* of e, and this bound is achieved. (For more see Walters "Introduction to Ergodic Theory," and David Ruell,...).

A measure μ is a *Gibbs state* if its conditional distribution "are Boltzmann."
A measure μ is an *Equilibrium state* if it maximizes $h(\mu) - \beta e(\mu)$.

Theorem.
(1) Every equilibrium state is τ-invariant Gibbs state.
(2) If $X = S^{Z^d}$, every τ-invariant Gibbs state is an equilibrium state.
Why this fails for subshifts? Let $S = \{\uparrow, \downarrow, \leftarrow, \downarrow\}$, $d = 2$, with the contraint that every arrow must be "requited."

$\rightarrow$	$\leftarrow$	$\rightarrow$
$\leftarrow$	$\downarrow$	$\rightarrow$
$\leftarrow$	$\uparrow$	$\downarrow$

The point-mass concentrated on

$\leftarrow$	$\rightarrow$	$\leftarrow$	$\rightarrow$

is a Gibbs measure with nowhere near maximal entropy.