LARGE RANDOM SYSTEMS

large $\quad \underbrace{\text{random system}}$

e.g. size parameter n
limit $n \to \infty$

$(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$
proba space depending on size n

limit of what?

- probabilities, expected values
- distributions

Example (UNIFORM) RANDOM PERMUTATIONS

size param. $\quad n \in \mathbb{N}$

$$\Omega_n = S_n = \{ \sigma : \{1,2,\dots,n\} \circlearrowleft \text{ bijective} \}$$

(symmetric group on n elements)

$\mathbb{P}_n$ = uniform measure on S_n

$$\mathbb{P}_n[\sigma = \pi] = \frac{1}{n!} \quad \forall \pi \in S_n$$

"random outcome"

What is this modeling?
What are natural questions?

▶ Sorting algorithms:

Common task in computer science:
sort a list of n elements (n large)

Algorithms: "mergesort", "quicksort", "insertion sort", ...

Performance of algorithm:
e.g. number N_n of pair comparisons needed
(≈ required processor time)

Input: natural to assume original list
is randomly ordered (uniform distribution)

$\rightsquigarrow$ N_n is a random variable
(deterministic function of random input)

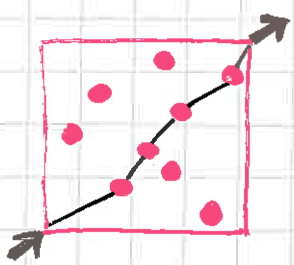
e.g. quicksort $\quad \mathbb{E}[N_n] = 2(n+1)\,H_n - 4n \sim 2n \cdot \ln(n)$

$\qquad H_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}$

see e.g. [Knuth] $\quad \mathrm{Var}[N_n] = \dots \sim 0{,}4202\, n^2$

▶ Disordered environments

A model of an interface in disordered environment:
place n dislocations independently uniformly in the unit square



consider a NE-interface from bottom left to top right corner such that a maximal number of dislocations is on its path

#(dislocations on path)
= length ℓ_n of longest increasing subsequence in a uniform random permutation

$\mathbb{E}[\ell_n] \sim 2\sqrt{n}$ $\qquad$ [Baik & Deift & Johansson]

$\mathrm{Var}[\ell_n] \sim C \cdot n^{1/3}$

$$\frac{\ell_n - 2\sqrt{n}}{n^{1/6}} \longrightarrow \text{"Tracy-Widom distribution"}$$

▶ Shuffling cards

deck of n cards

order of cards $\longleftrightarrow$ permutation $\sigma \in S_n$.

well shuffled deck $\longleftrightarrow$ $\sigma \sim$ uniform distr. μ_{unif} on S_n.

shuffling : Markov chain $(\sigma_s)_{s=0}^{\infty}$ on S_n

e.g. "top-to-middle shuffle"

insert top card to a uniform random position

irreducible aperiodic Markov chain whose stationary distribution is μ_{unif}.

$$\Longrightarrow \quad \mathbb{P}[\sigma_s = \pi] \xrightarrow[s \to \infty]{} \frac{1}{n!} = \mu_{\text{unif.}}[\sigma = \pi]$$

Distributions of the S_n-valued random variables σ_s (order after s shuffles) converge as the number of shuffles increases, $s \longrightarrow \infty$.

One notion of convergence of probability measures (= distributions) on a finite set:

Total variation distance

$\mathcal{X}$ finite set

μ, ν proba measures on $\mathcal{X}$

$$\varrho_{TV}(\mu, \nu) = \max_{E \subset \mathcal{X}} \left| \mu[E] - \nu[E] \right|$$

"maximum error in probability of any event, if μ is used instead of ν"

Exercise:

(a) Show that $\varrho_{TV}(\mu, \nu) = \frac{1}{2} \sum_{x \in \mathcal{X}} \left| \mu\{x\} - \nu\{x\} \right|$.

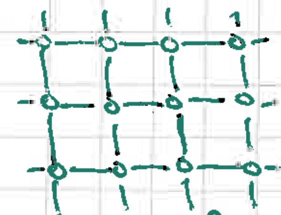
(b) Show that ϱ_{TV} is a metric on the space of proba measures on $\mathcal{X}$.

RECURRENCE AND TRANSIENCE OF RANDOM WALKS

We will consider one of the most common stochastic processes: the simple random walk on d-dimensional (hyper) cubic lattice $\mathbb{Z}^d$.

$d \in \{1, 2, 3, \ldots\}$

We ask whether the walk always returns to its starting point, or whether it can escape to infinity.



$\mathbb{Z}^2$

Lattice $\quad \mathbb{Z}^d = \{ (x_1, \ldots, x_d) \in \mathbb{R}^d \mid x_1, \ldots, x_d \in \mathbb{Z} \}$

standard basis: $\quad e_1, \ldots, e_d$

where $\quad e_k = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$ ← k-th position

nearest neighbors of origin: $\pm e_1, \pm e_2, \ldots, \pm e_d$

Random walk $\quad X = (X_n)_{n=0}^{\infty} \qquad (X_n \in \mathbb{Z}^d \quad \forall n \in \mathbb{Z}_{\geq 0})$

- $X_0 = 0$
- $\mathbb{P}[X_{n+1} = y \mid X_n = x] = \begin{cases} 1/2d & \text{if } y - x = \pm e_k \text{ for some } k. \\ 0 & \text{otherwise} \end{cases}$

Construction: $(\xi_j)_{j=1}^{\infty}$ i.i.d. steps, uniformly distributed among the $2d$ nearest neighbors of origin

$$X_n := \sum_{j=1}^{n} \xi_j$$

Theorem If $d \leq 2$ then X is recurrent, i.e.,

$$\mathbb{P}[\text{for some } n > 0 \text{ we have } X_n = 0] = 1.$$

If $d > 2$ then X is transient, i.e.,

$$\mathbb{P}[\text{for some } n > 0 \text{ we have } X_n = 0] < 1.$$

Proof: Denote $p = \mathbb{P}[\exists n > 0 \text{ s.t. } X_n = 0]$.

Consider the number of times the walk is at the origin,

$$L = \#\{n \in \mathbb{Z}_{\geq 0} \mid X_n = 0\} = \sum_{n=0}^{\infty} \mathbb{1}_{\{X_n = 0\}},$$

and its expected value

$$\mathbb{E}[L] = \sum_{n=0}^{\infty} \mathbb{P}[X_n = 0].$$

If the walk ever returns to the origin (say at time τ), then the continuation from there is a new random walk, independent of the steps before, and will thus return to the origin with the same probability p.

We get $\mathbb{P}[L \geq n] = p^{n-1}$, and then compute

$$\mathbb{E}[L] = \sum_{n=0}^{\infty} p^n = \begin{cases} \frac{1}{1-p} & \text{if } p < 1 \\ \infty & \text{if } p = 1. \end{cases}$$

Recurrence and transience can thus be characterized using the random variable L:

X recurrent $\iff$ $\mathbb{E}[L] = \infty$

X transient $\iff$ $\mathbb{E}[L] < \infty$.

We generalize to a weighted number of visits, so as to make everything finite at first.

Let $\lambda < 1$ be a weight parameter and define

$$L_\lambda(x) = \sum_{n=0}^{\infty} \lambda^n \cdot \mathbb{1}_{\{X_n = x\}}$$

and

$$G_\lambda(x) = \mathbb{E}[L_\lambda(x)] = \sum_{n=0}^{\infty} \lambda^n \cdot \mathbb{P}[X_n = x]$$

In particular $\mathbb{E}[L] = \lim_{\lambda \to 1} G_\lambda(0)$ by monotone convergence.

The point is that $G_\lambda(x)$ is not difficult to calculate.

By "first step analysis" we get a difference equation:

$$G_\lambda(x) = \sum_{n=0}^{\infty} \lambda^n \, \mathbb{P}[X_n = x]$$

$$= \delta_{x,0} + \sum_{n=1}^{\infty} \lambda^n \, \mathbb{P}[X_n = x]$$

(Kronecker delta: $\delta_{a,b} = \begin{cases} 1 & \text{if } a=b \\ 0 & \text{if } a \neq b \end{cases}$)

$$= \delta_{x,0} + \frac{1}{2d}\Big(\lambda \cdot G_\lambda(x+e_1) + \lambda \cdot G_\lambda(x-e_1) + \dots + \lambda G_\lambda(x-e_d)\Big)$$

$$= \delta_{x,0} + \frac{\lambda}{2d} \sum_{k=1}^{d} \big(G_\lambda(x+e_k) + G_\lambda(x-e_k) \big).$$

This difference equation can be solved by Fourier transform

$$\hat{G}_\lambda(\theta) = \sum_{x \in \mathbb{Z}^d} e^{-i\theta \cdot x} \, G_\lambda(x) \qquad (\theta \in \mathbb{R}^d)$$

in terms of which G_λ itself reads

$$G_\lambda(x) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} e^{i\theta\cdot x} \, \hat{G}_\lambda(\theta) \, d^d\theta$$

Multiplying the difference equation by $e^{-i\theta\cdot x}$ and summing over $x \in \mathbb{Z}^d$ we get

$$\hat{G}_\lambda(\theta) = 1 + \frac{\lambda}{2d} \sum_{k=1}^{d} \left(e^{i\theta\cdot e_k} + e^{-i\theta\cdot e_k}\right) \hat{G}_\lambda(\theta)$$

which is solved

$$\hat{G}_\lambda(\theta) = \frac{1}{1 - \frac{\lambda}{d}\sum_{k=1}^{d} \cos(\theta_k)}$$

We get $$G_\lambda(x) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} d^d\theta \; \frac{e^{i\theta\cdot x}}{1 - \frac{\lambda}{d}\sum_{k=1}^{d}\cos(\theta_k)}.$$

So far we kept $\lambda < 1$ but now we are ready to write

$$\mathbb{E}[L] = \lim_{\lambda \nearrow 1} G_\lambda(0) = \lim_{\lambda \nearrow 1} \frac{1}{(2\pi)^d} \cdot \int_{[-\pi,\pi]^d} \frac{1}{1 - \frac{\lambda}{d}\sum_{k=1}^{d}\cos(\theta_k)} \, d^d\theta.$$

Recurrence and transience thus reduces to whether this integral converges or not. Outside of any neighborhood of $\theta = 0$ the integrand is bounded*, so it is sufficient to consider integral over a small ball $B_\varepsilon(0)$.

*a continuous function on a compact set (one could of course also give explicit bounds)

If ε is small $(\varepsilon < \frac{\pi}{2})$ then $\cos(\theta_k) > 0 \quad \forall k$ if $\theta \in B_\varepsilon(0)$ and therefore

$$\frac{\lambda}{d}\sum_{k=1}^{d}\cos(\theta_k) \xrightarrow[\lambda \nearrow 1]{\text{(monotone increasing)}} \frac{1}{d}\sum_{k=1}^{d}\cos(\theta_k)$$

Monotone convergence : $\circledast = \lim_{\lambda \nearrow 1} \int_{B_\varepsilon(0)} \frac{1}{1-\frac{\lambda}{d}\sum_{k=1}^{d}\cos(\theta_k)} d^d\theta = \int_{B_\varepsilon(0)} \frac{1}{1-1/d\sum_{k=1}^{d}\cos(\theta_k)} d^d\theta$

Now note that if $|\alpha| < \frac{\pi}{4}$ then

$$\frac{1}{2\sqrt{2}}\alpha^2 \le \underbrace{1-\cos(\alpha)}_{f(\alpha)} \le \frac{1}{2}\alpha^2$$

$f(0) = 0$

$f'(0) = 0$

$\frac{1}{\sqrt{2}} \le f''(\alpha) \le 1$

so that if $\theta \in B_\varepsilon(0)$ with $\varepsilon < \frac{\pi}{4}$ we have

$$\frac{2}{\|\theta\|^2} \le \frac{1}{1-\frac{1}{d}\sum_{k=1}^{d}\cos(\theta_k)} \le \frac{2\sqrt{2}}{\|\theta\|^2}, \quad \text{where } \|\theta\|^2 = \sum_{k=1}^{d}\theta_k^2.$$

Therefore $\circledast \asymp \int_{B_\varepsilon(0)} \frac{1}{\|\theta\|^2} d^d\theta = \int_0^\varepsilon dr \frac{1}{r^2} C_d \cdot r^{d-1}$

$f \asymp g$ means $\exists\, c_1, c_2 > 0$ s.t.
$c_1 g(x) \le f(x) \le c_2 g(x)$

and in particular $\circledast < \infty \iff d > 2$.

We conclude $\mathbb{E}[L] < \infty \iff d > 2$, and thus

X transient $\iff d > 2$.

$\square$

ZERO-ONE LAWS AND APPLICATIONS

Often in an infinite system some properties hold with certainty (with probability one, i.e., almost surely).

We will recall two such "zero-one laws":

- Borel-Cantelli lemmas
- Kolmogorov's 0-1 law

As example applications we present

- the law of iterated logarithm for random walk
- phase transition (for existence of ∞ cluster) in percolation.

KOLMOGOROV'S 0-1 LAW

Let $X_1, X_2, X_3, \ldots$ be a sequence of independent random variables (taking values in an arbitrary measure space).

Denote by $\mathcal{T}_n = \sigma(X_n, X_{n+1}, X_{n+2}, \ldots)$ the σ-algebra generated by X_m, $m \geq n$

($\mathcal{T}_n$ = smallest σ-alg. with respect to which each X_m, $m \geq n$, is measurable)

Denote by $\mathcal{T} = \bigcap_{n \geq 1} \mathcal{T}_n$ the "tail sigma algebra".

Intuition: an event is in the tail σ-algebra if its occurrence is unaffected by changing just finitely many of $X_1, X_2, X_3, \ldots$

Theorem (Kolmogorov's 0-1 law)

If $E \in \mathcal{T}$ then either $\mathbb{P}[E] = 0$ or $\mathbb{P}[E] = 1$.

Proof: See course "Introduction to stochastics". $\square$

PERCOLATION

Different spacial locations are declared open or closed independently. This can model e.g.

porous material : liquid can penetrate a connected set of open locations

infectuous disease : disease can be transmitted through open locations, spreads to a connected set of open locations

etc. etc.

Bond percolation on hypercubic lattice

- $d \in \{1,2,3,\ldots\}$ dimension
- lattice $\mathbb{Z}^d$
  - sites : $x \in \mathbb{Z}^d$

    bonds: $e = \{x,y\}$ s.t. $x,y \in \mathbb{Z}^d$, $\|x-y\| = 1$

    set of bonds denoted $E(\mathbb{Z}^d)$.
- parameter $p \in [0,1]$ (probability of open bonds)
- random configuration : $\omega = (\omega_e)_{e \in E(\mathbb{Z}^d)}$

  $e \in E(\mathbb{Z}^d)$ open / closed represented by $\omega_e = 1$ / $\omega_e = 0$.

  $(\omega_e)_{e \in E}$ independent Bernoulli(p), i.e., $\mathbb{P}[\omega_e = 1] = p$

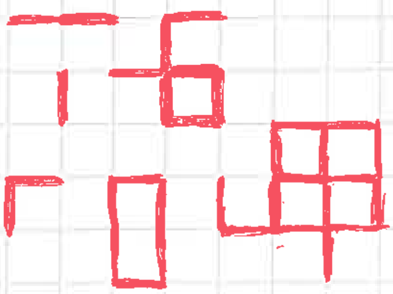
$\therefore$ $\Omega = \{0,1\}^{E(\mathbb{Z}^d)}$ countable product space

$\mathcal{F}$ = product sigma algebra = cylinder sigma algebra

(smallest σ-alg. s.t. projections $\omega \mapsto \omega_e$, $\Omega \to \{0,1\}$ are measurable for all $e \in E(\mathbb{Z}^d)$)

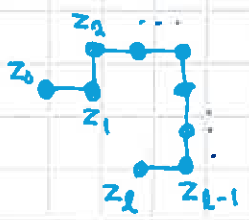
$\mathbb{P}$ = Bernoulli$(p)^{\otimes E(\mathbb{Z}^d)}$ countable product measure

(NOTE: We often denote $\mathbb{P}_p$ to emphasize the choice of the parameter p.)



Example configuration of open bonds in $\mathbb{Z}^2$

path (defined below)



A path of length ℓ on $\mathbb{Z}^d$ is a sequence $z_0, z_1, \dots, z_{\ell-1}, z_\ell \in \mathbb{Z}^d$ of distinct sites s.t. $\|z_j - z_{j-1}\| = 1 \quad \forall j = 1, \dots, \ell$

The path is open (for the configuration ω) if for all $j = 1, \dots, \ell$ we have $\omega_{\{z_{j-1}, z_j\}} = 1$.

Sites $x, y \in \mathbb{Z}^d$ are connected (in the config. ω) if there exists an open path $x = z_0, z_1, \dots, z_{\ell-1}, z_\ell = y$.

We denote $x \leftrightsquigarrow y$ (resp. $x \not\leftrightsquigarrow y$) if sites x, y are (resp. are not) connected.

The relation $\leftrightsquigarrow$ is a (random) equivalence relation, and the equivalence classes are called the components (or clusters) of the configuration ω.

The component of $x \in \mathbb{Z}^d$ is thus

$$\mathcal{C}_x = \{ y \in \mathbb{Z}^d \mid x \leftrightsquigarrow y \}$$

Exercise: Check that the following are measurable

- event $\{ x \leftrightsquigarrow y \}$ (for given $x, y \in \mathbb{Z}^d$)
- event $\{ \exists x \in \mathbb{Z}^d : \# \mathcal{C}_x = \infty \}$ = {there exists an infinite compon.}
- random variable $\# \mathcal{C}_x$, the size of the component of x.

Our main objective for this lecture is to prove the following "phase transition result"

Theorem: There exists a value $p_c \in [0,1]$ such that ← p_c depends on dimension d!

- for $p < p_c$: $\mathbb{P}_p$[there exists an infinite comp.] = 0
- for $p > p_c$: $\mathbb{P}_p$[— " —] = 1.

Moreover, if $d \geq 2$ the value is non-trivial, $p_c \in (0,1)$.

Remark: What happens at $p = p_c$? This is one of those cases which are not obvious to decide.

The event $\{ \exists \infty \text{ component} \}$ is sometimes called "percolation" and the complement $\{ \nexists \infty \text{ comp.} \}$ "non-percolation".

The first thing we note is the 0-1 law.

Proposition 1. For any p, $\mathbb{P}_p[\exists \infty \text{ comp.}] \in \{0,1\}$.

Proof: The event "there exists an infinite component" is measurable w.r.t. the information contained in $(\omega_e)_{e\in E(\mathbb{Z}^d)\setminus S}$ for any finite $S \subset E(\mathbb{Z}^d)$. Thus it is in the tail σ-algebra. Since $(\omega_e)_{e\in E(\mathbb{Z}^d)}$ are independent, we may apply Kolmogorov's 0-1 law $\square$

Since the only possible values of

$$\circledast = \mathbb{P}[\text{there exists an infinite component}]$$

are 0 and 1, it remains to show that as we increase the parameter p, the proba never jumps from 1 to 0.

(Seems obvious, right? But requires a proof...)

We will in fact prove something slightly more general, our first (very simple) monotonicity result.

Note that $\Omega = \{0,1\}^{E(\mathbb{Z}^d)}$ has a natural partial order: $\omega \preccurlyeq \omega'$ iff $\omega_e \leq \omega'_e$ $\forall e \in E(\mathbb{Z}^d)$.

We say that $f : \Omega \to \mathbb{R}$ is increasing if $\omega \preccurlyeq \omega' \Longrightarrow f(\omega) \leq f(\omega')$.

An event $E \in \mathcal{F}$ is said to be increasing if its indicator $\mathbb{1}_E$ is an increasing function.

We will prove:

Proposition 2. For any increasing function f, we have that $p \mapsto \mathbb{E}_p[f]$ is increasing.

The proof uses "coupling".

A <u>coupling</u> of two probability spaces $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ is a probability space $(\underbrace{\Omega_1 \times \Omega_2}_{\text{Carthesian product of sample spaces}}, \underbrace{\mathcal{F}_1 \otimes \mathcal{F}_2}_{\text{product sigma algebra}}, \mathbb{P}_{\text{coupling}})$ such that

- $\forall E_1 \in \mathcal{F}_1 : \mathbb{P}_{\text{coupling}}[E_1 \times \Omega_2] = \mathbb{P}_1[E_1]$
- $\forall E_2 \in \mathcal{F}_2 : \mathbb{P}_{\text{coupling}}[\Omega_1 \times E_2] = \mathbb{P}_2[E_2]$

"the two marginals of the coupling are the given probability measures"

<u>Remarks</u>: (i): The product measure $\mathbb{P}_1 \otimes \mathbb{P}_2$ is a coupling, but not a very useful one, since the two components become independent. The idea of couplings is to have some dependence between the components in order to obtain relations between $\mathbb{P}_1$ and $\mathbb{P}_2$.

(ii): Another formulation: given $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$, $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ find $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and two random variables on it $X_1 : \tilde{\Omega} \to \Omega_1$ and $X_2 : \tilde{\Omega} \to \Omega_2$ such that the law of X_1 is $\mathbb{P}_1$ and the law of X_2 is $\mathbb{P}_2$. Then the law of the pair $(X_1, X_2) \in \Omega_1 \times \Omega_2$ is a coupling.

<u>Lemma 1</u> (Monotone coupling for percolation)

Let $0 \leq p_1 \leq p_2 \leq 1$. Then there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and random variables $\omega^{(1)}, \omega^{(2)} : \tilde{\Omega} \to \{0,1\}^{E(\mathbb{Z}^d)}$ such that $\omega^{(1)} \preceq \omega^{(2)}$ (almost surely) and the distributions of $\omega^{(1)}$ and $\omega^{(2)}$ are $\mathbb{P}_{p_1}$ and $\mathbb{P}_{p_2}$, i.e., percolations with parameters p_1 and p_2, respectively.

<u>Proof</u> Define $\tilde{\Omega}$ by taking independent uniform random variables U_e on $[0,1]$ for each bond $e \in E(\mathbb{Z}^d)$.

(Formally: $\tilde{\Omega} = [0,1]^{E(\mathbb{Z}^d)}$, $\tilde{\mathcal{F}} = (\text{Borel}([0,1]))^{\otimes E(\mathbb{Z}^d)}$, $\tilde{\mathbb{P}} = (\text{Unif}([0,1]))^{\otimes E(\mathbb{Z}^d)}$)

Define random variables $\omega^{(1)} = (\omega_e^{(1)})_{e \in E(\mathbb{Z}^d)}$ and $\omega^{(2)} = (\omega_e^{(2)})_{e \in E(\mathbb{Z}^d)}$ by $\omega_e^{(1)} = \mathbb{1}_{\{U_e \leq p_1\}}$ and $\omega_e^{(2)} = \mathbb{1}_{\{U_e \leq p_2\}}$.

Clearly $\omega_e^{(1)} \leq \omega_e^{(2)}$ $\forall e \in E(\mathbb{Z}^d)$ so $\omega^{(1)} \preceq \omega^{(2)}$.

Also $\tilde{\mathbb{P}}[\omega_e^{(1)} = 1] = p_1$ and $(\omega_e^{(1)})_{e \in E(\mathbb{Z}^d)}$ are independent so $\omega^{(1)} \sim \mathbb{P}_{p_1}$. Similarly $\omega^{(2)} \sim \mathbb{P}_{p_2}$. $\square$

Proof of Proposition 2: Let $f : \Omega \to \mathbb{R}$ be increasing, and let $\tilde{\mathbb{P}}$ and $\omega^{(1)}, \omega^{(2)}$ be as in Lemma.

Then

$$\mathbb{E}_{p_1}[f(\omega)] = \tilde{\mathbb{E}}[f(\omega^{(1)})] \leq \tilde{\mathbb{E}}[f(\omega^{(2)})] = \mathbb{E}_{p_2}[f(\omega)]$$

law of $\omega^{(1)}$ is $\mathbb{P}_{p_1}$ — recall: $\omega^{(1)} \leq \omega^{(2)}$ and f increasing so $f(\omega^{(1)}) \leq f(\omega^{(2)})$ — law of $\omega^{(2)}$ is $\mathbb{P}_{p_2}$

$\square$

Examples: The following are increasing functions of p:

$p \longmapsto \mathbb{P}_p[\text{there exists an infinite connected component}]$

$p \longmapsto \mathbb{P}_p[\# C_0 \geq s]$ (for any $s \in \{0, 1, \dots, \infty\}$)

$p \longmapsto \mathbb{E}_p[\# C_0]$

At this stage we have seen that (by 0-1 law and monotonicity)

$$\circledast = \mathbb{P}_p[\exists \infty \text{ component}]$$

is increasing in p and only takes values 0 or 1. The first part of the theorem is thus proven: it suffices to take $p_c = \sup\{p \in [0,1] : \mathbb{P}_p[\exists \infty \text{ comp.}] = 0\}$.

We proceed to show the second part, that $0 < p_c < 1$ if $d \geq 2$. (Remark: in $d=1$ we have a trivial value, $p_c = 1$)

Let us characterize $\circledast = 0$ or $\circledast = 1$ differently.

Lemma 2 Let $\theta(p) = \mathbb{P}_p[\# C_0 = \infty]$.

Then $\theta(p) > 0$ iff $\mathbb{P}_p[\exists \infty \text{ component}] = 1$.

Proof: Obviously $\mathbb{P}[\exists \infty \text{ comp.}] \geq \mathbb{P}[\# C_0 = \infty] = \theta(p)$.

If $\theta(p) > 0$ then $\mathbb{P}[\exists \infty \text{ comp}] > 0$ and by 0-1 law $\mathbb{P}[\exists \infty \text{ comp}] = 1$.

Conversely, if $\theta(p) = 0$, then $\mathbb{P}[\# C_x = \infty] = 0 \quad \forall x \in \mathbb{Z}^d$ by translation invariance. Therefore

$$\mathbb{P}[\exists \infty \text{ comp.}] = \mathbb{P}[\exists x \in \mathbb{Z}^d \text{ s.t. } \# C_x = \infty] \leq \sum_{x \in \mathbb{Z}^d} \mathbb{P}[\# C_x = \infty] = \sum_x 0 = 0$$

$\square$

Let us first prove "non-percolation" for small p.

Lemma 3 If $p < \frac{1}{2d}$ then $\theta(p) = 0$.

Proof Denote by Γ_n the set of paths of n steps starting from 0. (remember that we require $n+1$ distinct sites along the path)

Obviously $\# \Gamma_n \leq (2d)^n$ since each step has at most $2d$ options (nearest nbrs on $\mathbb{Z}^d$).

We now note that for any n

$$\theta(p) = \mathbb{P}_p[\# C_0 = \infty] \leq \mathbb{P}_p[\exists \gamma \in \Gamma_n \text{ which is open}]$$

$$\leq \sum_{\gamma \in \Gamma_n} \mathbb{P}_p[\gamma \text{ is open}] = \sum_{\gamma \in \Gamma_n} p^n$$

$$= (\# \Gamma_n) \cdot p^n \leq (2d \cdot p)^n .$$

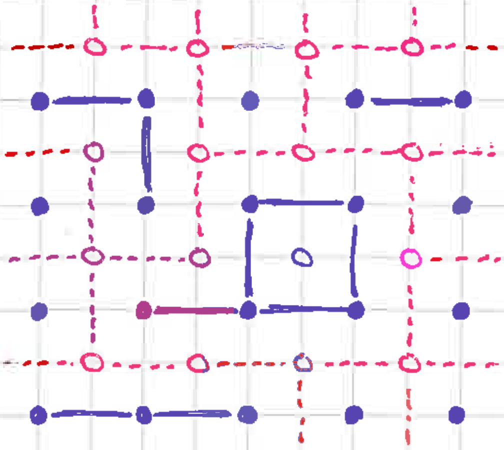
If $p < \frac{1}{2d}$ then $(2dp)^n \longrightarrow 0$ as $n \longrightarrow \infty$ and we conclude $\theta(p) = 0$. $\square$

This shows that $p_c \geq \frac{1}{2d}$.

We then prove "percolation" for large p, but for this we start by treating the special case $d=2$ where a planar duality argument can be used.

Lemma 4 In $d=2$, if $p > \frac{3}{4} + \frac{1}{4\sqrt{2}}$ then $\theta(p) > 0$.

(note $\frac{3}{4} + \frac{1}{4\sqrt{2}} < 1$)



$\bullet \in \mathbb{Z}^2$ square lattice

$\circ \in (\mathbb{Z}+\frac{1}{2})^2$ dual lattice

Define the dual of $\mathbb{Z}^2$ as $(\mathbb{Z}+\frac{1}{2})^2$, the set of midpoints of the squares (faces of $\mathbb{Z}^2$).

Bonds of the dual are pairs $\{q,r\}$ $q,r \in (\mathbb{Z}+\frac{1}{2})^2$ such that $\|q-r\|=1$. Note that each dual bond crosses a unique bond of $\mathbb{Z}^2$ and vice versa ($\{q,r\}$ crosses $\{x,y\} \in E(\mathbb{Z}^2)$ if the midpoints agree $\frac{q+r}{2}=\frac{x+y}{2}$)

A percolation configuration ω^* of the dual is defined from config. ω by setting $\omega^*_{\{q,r\}} = 1-\omega_{\{x,y\}}$

(dual bond is open $\iff$ the crossing original bond is closed)

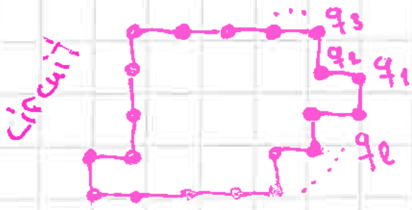
Note that if ω^* contains an open circuit surrounding the origin, then $\mathcal{C}_0$ is finite. Conversely, if $\mathcal{C}_0$ is finite, then ω^* contains such a circuit (the duals of the bonds connecting sites of $\mathcal{C}_0$ to sites of the complement of $\mathcal{C}_0$ form such a circuit).

Let Γ^0_n denote the set of circuits of n steps on the dual, which surround the origin.

We have $\#\Gamma^0_n \leq \frac{n}{2} \cdot 4^n$. (Why? Consider the minimal position at which the circuit crosses positive real axis. Given length n, this can occur at most at $\frac{n}{2}$ possible places. From there, each step has at most 4 options.)

Circuits: a circuit of length ℓ is a sequence of distinct dual sites $q_1, q_2, \dots, q_{\ell-1}, q_\ell \in (\mathbb{Z}+\frac{1}{2})^2$ such that

$$\|q_j - q_{j-1}\| = 1 \quad \forall j=2,3,\dots,\ell \quad \text{and} \quad \|q_\ell - q_1\| = 1.$$



If the sequences of sites of two circuits are obtained from each other by a cyclic permutation, we consider the two circuits the same

<u>Proof of Lemma 4:</u> Estimate

$$1-\theta(p) = \mathbb{P}_p[\# \mathcal{C}_0 < \infty]$$
$$= \mathbb{P}_p[\text{for some } n \text{ there exists dual open circuit } \gamma^* \in \Gamma_n^0]$$
$$\leq \sum_n \sum_{\gamma^* \in \Gamma_n^0} \mathbb{P}_p[\gamma^* \text{ is dual open}]$$
$$= \sum_n (\# \Gamma_n^0) \cdot (1-p)^n$$
$$\leq \frac{1}{2} \sum_n n \cdot (4(1-p))^n \leq \frac{1}{2} \frac{1}{(1-4(1-p))^2} \quad \text{if } 4(1-p)<1 \text{ i.e. } p > \tfrac{3}{4}.$$

if $|x|<1$ $\quad \frac{1}{1-x} = \sum_{n\geq 0} x^n \qquad \frac{1}{(1-x)^2} = \frac{d}{dx}\left(\frac{1}{1-x}\right) = \sum_{n\geq 0} n \cdot x^{n-1} = \sum_{n \geq 0} (n+1) x^n$

Then note that if $p > \frac{3}{4} + \frac{1}{4\sqrt{2}} \approx 0{,}927$ then the above upper bound is less than one, so $1-\theta(p) < 1$ and thus $\theta(p) > 0$. □

<u>Proof of theorem</u>

We have shown (by monotonicity and 0-1 law) that $p_c = \sup\{p \in [0,1] \mid \mathbb{P}[\exists \infty \text{ comp.}] = 0\}$ is the treshold.

From Lemma 3 we get $p_c \geq \frac{1}{2d} > 0$.

From Lemma 4 we get in $d=2$ that $p_c \leq \frac{3}{4} + \frac{1}{4\sqrt{2}} < 1$.

But if $d>2$, $\mathbb{Z}^d$ contains a sublattice $\mathbb{Z}^2$ and percolation on $\mathbb{Z}^d$ will have an infinite component if the percolation on the sublattice has one.

Therefore in $d>2$ still for $p > \frac{3}{4} + \frac{1}{4\sqrt{2}}$ we have

$$\mathbb{P}[\exists \infty \text{ comp.}] = 1 \quad \text{and thus} \quad p_c \leq \frac{3}{4} + \frac{1}{4\sqrt{2}} < 1.$$
□

ZERO-ONE LAWS AND APPLICATIONS

Infinite system (with enough independence)

$\rightsquigarrow$ some properties become certain (hold with probability one)

0-1 laws

- Borel-Cantelli lemmas
- Kolmogorov's 0-1 law

Example application

- law of iterated log for random walk
- phase transition in percolation

BOREL-CANTELLI LEMMAS

A_n events, $n = 1, 2, 3, \ldots$

$$\limsup(A_n) = \bigcap_{m\in\mathbb{N}} \bigcup_{k\geq m} A_k = \{\omega \in \Omega \mid \omega \in A_n \text{ for infinitely many } n\}$$

$$= \text{"} A_n \text{ occurs infinitely often"}$$

(abbreviation: "A_n i.o.")

Lemma (Borel-Cantelli)

Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of events. Then

(i) $\sum_{n\in\mathbb{N}} \mathbb{P}[A_n] < \infty \implies \mathbb{P}[A_n \text{ i.o.}] = 0$

(ii) (A_n) $\perp\!\!\!\perp$ and $\sum_{n\in\mathbb{N}} \mathbb{P}[A_n] = \infty \implies \mathbb{P}[A_n \text{ i.o.}] = 1$.

Proof: See course "Introduction to stochastics". $\square$

LAW OF ITERATED LOGARITHM

Consider random walk $X = (X_n)_{n \in \mathbb{Z}_{\geq 0}}$ with Gaussian steps:

$$X_n = \sum_{s=1}^{n} \xi_s \qquad \text{where} \quad (\xi_s)_{s \in \mathbb{N}} \text{ are i.i.d}$$
$$\xi_s \sim N(0,1).$$

Our main result of this lecture says that $\lambda(n) = \sqrt{n \cdot \log(\log(n))}$ is the scale of the level which is just barely reached by X_n in the long run.

Theorem (Law of iterated logarithm)

Almost surely $\limsup_{n \to \infty} \frac{X_n}{\lambda(n)} = \sqrt{2}$.

Remark: The same holds for random walk with ± 1 steps, and you should be able to modify the proof below to cover that case (with some effort).

Auxiliary estimates

We need estimates (upper and lower bounds) about the tail of the distribution of X_n. In the Gaussian steps case, these follow from the next lemma.

Lemma 1. (Tail bounds for a Gaussian random variable)

Let $\xi \sim N(0,1)$.

(a) For any $x > 0$ we have $\mathbb{P}[\xi > x] \leq e^{-\frac{1}{2}x^2}$.

(b) $\exists\, C, x_0 > 0$ s.t. $\forall x > x_0$ $\mathbb{P}[\xi > x] \geq C \cdot \frac{1}{x} e^{-\frac{1}{2}x^2}$.

Using a similar integration by parts trick as in the proof below, you can easily improve the asymptotics to

Exercise Show that for all $x > 0$

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{x} - \frac{1}{x^3}\right) e^{-\frac{1}{2}x^2} \leq \mathbb{P}[\xi > x] \leq \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{x} e^{-\frac{1}{2}x^2}$$

Proof: (a) For any $\theta > 0$ we have

$$\mathbb{P}[\xi > x] = \mathbb{P}[e^{\theta\xi} > e^{\theta x}]$$
$$\leq \frac{1}{e^{\theta x}} \mathbb{E}[e^{\theta \xi}] \quad \text{(Markov inequality)}$$
$$= \frac{1}{e^{\theta x}} e^{\frac{1}{2}\theta^2} = e^{\frac{1}{2}\theta^2 - \theta x}$$

The optimal upper bound is obtained with $\theta = x$, which gives the asserted result.

(b) Write $\mathbb{P}[\xi > x] = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{1}{2}s^2} ds$.

Integration by parts gives $\left(\frac{d}{ds}(s^{-1} \cdot e^{-\frac{1}{2}s^2}) = -s^{-2} e^{-\frac{1}{2}s^2} - e^{-\frac{1}{2}s^2} \right)$

$$\int_x^\infty e^{-\frac{1}{2}s^2} ds = x^{-1} \cdot e^{-\frac{1}{2}x^2} - \int_x^\infty s^{-2} \cdot e^{-\frac{1}{2}s^2} ds.$$

The second term is negligible w.r.t. the first (the ratio goes to zero as $x \to \infty$) so we get

$$\frac{\int_x^\infty e^{-\frac{1}{2}s^2} ds}{x^{-1} \cdot e^{-\frac{1}{2}x^2}} \xrightarrow[x \to \infty]{} 1.$$

Therefore, for any $C < 1$, for x sufficiently large, we have

$$\int_x^\infty e^{-\frac{1}{2}s^2} ds \geq C \cdot x^{-1} \cdot e^{-\frac{1}{2}x^2},$$

which shows the asserted result. $\square$

This translates to the following about $X_n = \sum_{s=1}^n \xi_s \sim N(0, n)$.

Corollary 1: We have, for all $x > 0$ and n

$$\mathbb{P}[X_n > x] \leq e^{-\frac{1}{2n}x^2} \quad \text{and} \quad \mathbb{P}[|X_n| > x] \leq 2 e^{-\frac{1}{2n}x^2}.$$

Also, there exists constants C, x_0 such that for all $x > x_0$ and any n $\mathbb{P}[X_n > x] \leq C \frac{\sqrt{n}}{x} e^{-\frac{1}{2n}x^2}$

Recall also the following exercise

Exercise If X_n is the simple random walk with ± 1 steps then $\mathbb{P}[X_n > x] \leq e^{-\frac{1}{2n}x^2}$ for any $x > 0$, $n \in \mathbb{Z}_{>0}$.

Upper bound

Recall that we want to show $\limsup_{n\to\infty} \frac{X_n}{\lambda(n)} = \sqrt{2}$ (a.s.)
with $\lambda(n) = \sqrt{n \cdot \log(\log(n))}$.
Our first goal is to show $\limsup_{n\to\infty} \frac{X_n}{\lambda(n)} \leq \sqrt{2}$ (almost surely).

Two steps: 1°) exponentially growing subsequence (n_k)
2°) gaps in the subsequence

Let $\alpha > 1$ (the value will be chosen later),
and define $n_k = \lfloor \alpha^k \rfloor$. (integer part of the number α^k)
This subsequence is exponentially growing

$$C\alpha^k \leq n_k \leq \alpha^k \quad \text{for some } C>0.$$

$$\left(\text{in fact: } \quad \alpha^k - 1 \leq n_k \leq \alpha^k \right)$$

Lemma 2.

For any $\beta > \sqrt{2}$ we have $\mathbb{P}\left[\limsup_{k\to\infty} \frac{X(n_k)}{\lambda(n_k)} \leq \sqrt{2} \right] = 1$.

Proof: By the tail upper bound (Corollary 1, first part)

$$\mathbb{P}\left[\frac{X_n}{\lambda(n)} > \beta \right] \leq \exp\left(- \frac{\beta^2 \lambda(n)^2}{2n} \right) = \exp\left(- \frac{\beta^2}{2} \log(\log(n)) \right) = \log(n)^{-\beta^2/2}.$$

By exponential growth of n_k we have $\log(n_k) \geq ak + c$
for $a = \log(\alpha) > 0$ and $c \in \mathbb{R}$. Therefore, for $\beta > \sqrt{2}$

$$\sum_{k=1}^{\infty} \mathbb{P}\left[\frac{X_{n_k}}{\lambda(n_k)} > \beta \right] \leq \sum_{k=1}^{\infty} (ak+c)^{-\beta^2/2} < \infty.$$

Borel-Cantelli lemma thus guarantees that
$\frac{X_{n_k}}{\lambda(n_k)} > \beta$ for only finitely many $k \in \mathbb{N}$ (almost surely),

i.e. $\limsup_{k\to\infty} \frac{X_{n_k}}{\lambda(n_k)} \leq \beta$. $\square$

To fill in gaps between n_k and n_{k+1} we use:

Lemma 3. (Lévy's inequality)

Suppose that $(\xi_u)_{u\in\mathbb{N}}$ are independent.
Set $S_t = \sum_{u=1}^{t} \xi_u$. Fix $m \in \mathbb{Z}_{>0}$ (and consider $\max_{1\le t\le m} |S_t|$).
Assume that for all $t \le m$ we have $\mathbb{P}[|S_t - S_m| \le \sigma] \ge \rho$
Then $\mathbb{P}[\max_{1\le t\le m} |S_t| > 2\sigma] \le \frac{1}{\rho} \mathbb{P}[|S_m| > \sigma]$.

Proof: We split according to the first time t at which $|S_t| > 2\sigma$.
Define the events $A_t = \{|S_1| \le 2\sigma, \dots, |S_{t-1}| \le 2\sigma, |S_t| > 2\sigma\}$
and $B_t = \{|S_m - S_t| \le \sigma\}$.

Then $A_1, A_2, \dots, A_m$ are disjoint and $A_t \perp B_t$.

$\sigma(\xi_1,\dots,\xi_t) \perp \sigma(\xi_{t+1},\dots,\xi_m)$

Also $\bigcup_{t=1}^{m} A_t \cap B_t \subset \{|S_m| > \sigma\}$.

Thus
$$\mathbb{P}[|S_m| > \sigma] \ge \sum_{t=1}^{m} \mathbb{P}[A_t] \underbrace{\mathbb{P}[B_t]}_{\ge \rho} \ge \rho \cdot \sum_{t=1}^{m} \mathbb{P}[A_t] = \rho \cdot \mathbb{P}[M_m > 2\sigma]$$
□

We are now ready to fill in gaps and conclude the upper bound.

Proposition 1: We have $\mathbb{P}\left[\limsup_{n\to\infty} \frac{X_n}{\lambda(n)} \le \sqrt{2}\right] = 1$.

Proof: Let $\beta > \sqrt{2}$ and $\varepsilon > 0$.
We will use a subsequence $n_k = \lfloor \alpha^k \rfloor$ with a suitably chosen $\alpha > 1$ and control the gaps between n_k and n_{k+1} with an error of order ε.

To see what happens between n_k and n_{k+1} we seek to bound the probability

$$\mathbb{P}\Big[\max_{n_k \le j \le n_{k+1}} |X_j - X_{n_k}| > 2\varepsilon \cdot \lambda(n_k)\Big]. \qquad \text{"}\sigma = 2\varepsilon\lambda(n_k)\text{"}$$

To achieve this, set $S_t = X_{t+n_k} - X_{n_k}$ and $m = n_{k+1} - n_k$.

Denote $\delta_k = \max_{n_k \le j \le n_{k+1}} \mathbb{P}\big[|X_j - X_{n_{k+1}}| > \varepsilon \cdot \lambda(n_k)\big]$.

Lévy's inequality gives

this bounds $\mathbb{P}[|X_{n_k} - X_{n_{k+1}}|>$

$$\mathbb{P}\Big[\max_{n_k \le j \le n_{k+1}} |X_j - X_{n_k}| > 2\varepsilon\lambda(n_k)\Big] \le \frac{\delta_k}{1-\delta_k}.$$

this is our ρ.

We estimate δ_k by Corollary 1

$$\delta_k \le \max_{n_k \le j \le n_{k+1}} \left(2 \cdot \exp\left(-\frac{\varepsilon^2 \cdot \lambda(n_k)^2}{2(n_{k+1}-j)}\right)\right)$$

$$= 2 \cdot \exp\left(-\frac{\varepsilon^2 \lambda(n_k)}{2(n_{k+1}-n_k)}\right) = 2\big(\log(n_k)\big)^{-\frac{\varepsilon^2}{2}\cdot\frac{n_k}{n_{k+1}-n_k}}$$

We have $\frac{n_k}{n_{k+1}-n_k} \longrightarrow \frac{1}{\alpha-1}$ and $\log(n_k) \ge ak + c$.
(with $a = \log(\alpha) > 0$).

Therefore, if we choose $\alpha < 1 + \frac{\varepsilon^2}{2}$

we have $$\sum_{k=1}^{\infty} \mathbb{P}\Big[\max_{n_k \le j \le n_{k+1}} |X_j - X_{n_k}| > 2\varepsilon\lambda(n_k)\Big] < \infty.$$

Recall from Lemma 3 that $|X_{n_k}| > \beta \cdot \lambda(n_k)$ for only finitely many k. By summability of the above probabilities and Borel-Cantelli we have also

$$\max_{n_k \le j \le n_{k+1}} |X_j - X_{n_k}| > 2\varepsilon \cdot \lambda(n_k) \quad \text{for only finitely many } k.$$

If k is now anything but one of these finitely many exceptions, then for $n_k \le j \le n_{k+1}$ we have

$$\frac{|X_j|}{\lambda(j)} \le \frac{|X_{n_k}| + |X_j - X_{n_k}|}{\lambda(n_k)} \le \beta + 2\varepsilon.$$

Since $\beta > \sqrt{2}$ and $\varepsilon > 0$ were arbitrary, we get

$$\limsup_{n\to\infty} \frac{X_n}{\lambda(n)} \le \sqrt{2}. \qquad \square$$

Lower bound

We wanted to show $\limsup_{n\to\infty} \frac{X_n}{\lambda(n)} = \sqrt{2}$ (a.s.)

We already showed "$\leq$", and now we show "$\geq$".

Note that to guarantee a large limsup, it is sufficient to consider subsequences. We will again use $n_k = \lfloor \alpha^k \rfloor$ for suitably chosen $\alpha > 1$.

The idea is to use the tail lower bound (Corollary 1, second part)

$$\mathbb{P}\left[\frac{X_n}{\lambda(n)} > \beta\right] \geq C' \cdot \frac{\sqrt{n}}{\beta\,\lambda(n)} \exp\left(-\frac{\beta^2 \lambda(n)^2}{2n}\right)$$

$$= \frac{C'}{\beta} \frac{1}{\sqrt{\log(\log(n))}} \log(n)^{-\beta^2/2}.$$

When $\beta < \sqrt{2}$ the sum along subsequence (n_k) diverges.

We'd like to use the second Borel–Cantelli, but X_{n_k} $k = 1, 2, \ldots$ are not really independent.

Slight adjustment: consider increments $R_k = X_{n_{k+1}} - X_{n_k}$

- the sequence $(R_k)_{k=1}^{\infty}$ is independent
- the law of R_k is that of X_{Δ_k} with $\Delta_k = n_{k+1} - n_k$

(note also $\frac{\Delta_k}{\alpha^k} \longrightarrow \alpha - 1 > 0$ as $k \to \infty$)

<u>Proposition 2</u> We have $\mathbb{P}\left[\limsup_{n\to\infty} \frac{X_n}{\lambda(n)} \geq \sqrt{2}\right] = 1$.

<u>Proof</u> Let $\beta < \sqrt{2}$. Recall $R_k \sim X_{\Delta_k}$ and use Corollary 1:

$$\mathbb{P}\left[\frac{R_k}{\lambda(\Delta_k)} > \beta\right] \geq \frac{C}{\beta} \frac{1}{\sqrt{\log(\log(\Delta_k))}} \log(\Delta_k)^{-\beta^2/2}$$

(for large enough k)

By $\frac{\Delta_k}{\alpha^k} \to \alpha - 1 > 0$ we get that for k large enough $\mathbb{P}\left[\frac{R_k}{\lambda(\Delta_k)} > \beta'\right] \geq \text{const} \cdot k^{-1}$

which implies that $\sum_{k=1}^{\infty} \mathbb{P}\left[\frac{R_k}{\lambda(\Delta_k)} > \beta\right] = \infty$.

By independence of $(R_k)_{k\in\mathbb{N}}$ and Borel-Cantelli we conclude that $R_k / \lambda(\Delta_k) > \beta$ for infinitely many k.

Because of the upper bound (Lemma 2) (applied to $-X_{n_k}$) we have $X_{n_k} \geq -\beta' \cdot \lambda(n_k)$ for all but finitely many k.

for any $\beta' > \sqrt{2}$

Thus for infinitely many k both hold, and we get

$$X_{n_{k+1}} = X_{n_k} + R_k \geq \beta\,\lambda(\Delta_k) - \beta' \cdot \lambda(n_k)$$

But $\dfrac{\lambda(\Delta_k)}{\lambda(n_k)} \xrightarrow[k\to\infty]{} \sqrt{\alpha - 1}$ so the second term can be made negligible as $k \to \infty$ by choosing α large enough.

$\square$

<u>Proof of theorem</u>:

Combine Propositions 1 and 2 to get: $\mathbb{P}\left[\limsup_{n\to\infty} \dfrac{X_n}{\lambda(n)} = \sqrt{2}\right] = 1$. $\square$

WEAK CONVERGENCE ON $\mathbb{R}$

Recall the idea, common in large random systems
- size parameter n, interested in $n \to \infty$
- for each n, random object X_n (here: real valued random variable)

Basic question: Do we have $X_n \longrightarrow X_\infty$ (some random object) in some suitable sense as $n \to \infty$?

Let's start with three simple example cases:

(1) Continuous dependence on parameters

(a) Suppose $X_n \sim \mathrm{Exp}(\lambda_n)$ and $\lambda_n \longrightarrow \lambda$ as $n \to \infty$.
In which sense $X_n \longrightarrow X \sim \mathrm{Exp}(\lambda)$?

(b) $X_n \sim N(\mu_n, \sigma_n^2)$
$\mu_n \to \mu$, $\sigma_n^2 \to \sigma^2$
In which sense $X_n \longrightarrow X \sim N(\mu, \sigma^2)$?

(c) etc. etc.

(2) Poisson approximation of binomial distribution

$P \sim \mathrm{Poisson}(\lambda)$ i.e. $\mathbb{P}[P=k] = \frac{\lambda^k}{k!} e^{-\lambda}$, $\forall k \in \mathbb{Z}_{\geq 0}$.

$B_n \sim \mathrm{Bin}(n, p_n)$ i.e. $\mathbb{P}[B=k] = \binom{n}{k} p_n^k (1-p_n)^{n-k}$, $k = 0, 1, \ldots, n$.

<u>Exercise</u> If $n \cdot p_n \longrightarrow \lambda$ as $n \to \infty$ then characteristic functions converge $\mathbb{E}[e^{i\theta B_n}] \to \mathbb{E}[e^{i\theta P}]$

(3) Minimum of independent uniform random variables

Let $(U_j)_{j=1}^\infty$ be i.i.d. $U_j \sim \mathrm{Unif}([0,1])$.

Denote $M_n = \min_{1 \leq j \leq n} U_j$ and $X_n = n \cdot M_n$ (rescaled minimum)

Cumulative distribution functions of X_n

$$F_n(x) = \mathbb{P}[X_n \leq x] = \mathbb{P}[M_n \leq \tfrac{x}{n}] = 1 - \mathbb{P}[M_n > \tfrac{x}{n}]$$

$$\underset{(\min)}{=} 1 - \mathbb{P}[\forall j \leq n : U_j > \tfrac{x}{n}] \overset{\perp}{=} 1 - (1 - \tfrac{x}{n})^n \xrightarrow[n \to \infty]{} 1 - e^{-x}.$$

In which sense $X_n \longrightarrow X \sim \mathrm{Exp}(1)$?

What do we want ?
- convergence of probabilities of events ?
- convergence of expected values ?
- convergence of characteristic functions ?
- convergence of cumulative distribution functions ?

We want convergence of "anything that can be reliably measured by observations of the random system"

Idealization : reliably measured observable
= expected value of bounded continuous function

Def : A sequence $(X_n)_{n\in\mathbb{N}}$ of real-valued random variables converges weakly (= "in distribution") to a real-val. r.v. X if

for all bounded continuous $f:\mathbb{R}\to\mathbb{R}$ we have $\mathbb{E}[f(X_n)] \xrightarrow[n\to\infty]{} \mathbb{E}[f(X)]$.

Remark : The r.v.'s X_n, $n\in\mathbb{N}$, do not even need be defined on the same probability space, each can have its own $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$.

Indeed, the condition only concerns the laws of X_n and X (= the distributions)

μ_n = law of X_n $\qquad\qquad$ μ = law of X

μ_n is (Borel) probability measure on $\mathbb{R}$

$\mu_n[B] = \mathbb{P}[X_n \in B]$ for all Borel sets $B\subset\mathbb{R}$

$(\mathbb{R}, \mathcal{B}, \mu_n)$ probability triple

$$\mathbb{E}[f(X_n)] = \int_{\mathbb{R}} f(x)\, d\mu_n(x) \xrightarrow{n\to\infty} \int_{\mathbb{R}} f(x)\, d\mu(x)$$

integration w.r.t. measure μ_n

We also denote $\mu_n \xrightarrow[n\to\infty]{w} \mu$, and say that the measures $(\mu_n)_{n\in\mathbb{N}}$ converge weakly to μ.

Notation for random var.'s : $X_n \xrightarrow{(d)} X$ or $X_n \xrightarrow{w} X$

The various forms of convergence in the examples are more or less equivalent, in the following precise sense.

Theorem* (Equivalent conditions for weak convergence on $\mathbb{R}$)

Let (X_n) be $\mathbb{R}$-valued r.v.'s

(μ_n) their laws — $\mu_n[B] = \mathbb{P}[X_n \in B]$

(F_n) their cumulative distribution functions — $F_n(x) = \mathbb{P}[X_n \le x]$

(χ_n) their characteristic functions — $\chi_n(\theta) = \mathbb{E}[e^{i\theta X_n}]$

and let also X be r.v., μ its law, F its c.d.f. and χ its characteristic fn.

Then the following are equivalent:

(1°) $\forall f : \mathbb{R} \to \mathbb{R}$ continuous bounded : $\mathbb{E}[f(X_n)] \xrightarrow[n\to\infty]{} \mathbb{E}[f(X)]$

(5°) for all continuity points x of F : $F_n(x) \xrightarrow[n\to\infty]{} F(x)$

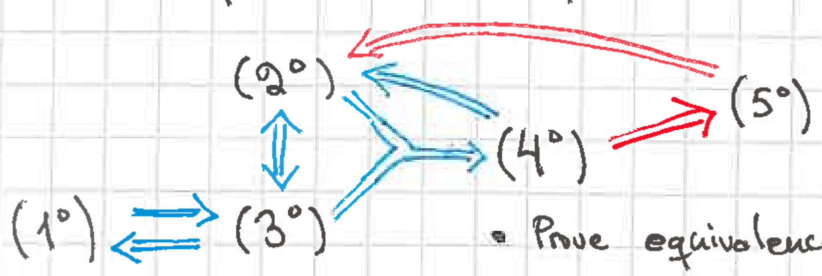
(6°) for all $\theta \in \mathbb{R}$: $\chi_n(\theta) \xrightarrow[n\to\infty]{} \chi(\theta)$

(2°) for all $V \subset \mathbb{R}$ open : $\liminf_{n\to\infty} (\mu_n[V]) \ge \mu[V]$

(3°) for all $A \subset \mathbb{R}$ closed : $\limsup_{n\to\infty} (\mu_n[A]) \le \mu[A]$

(4°) for all Borel sets $B \subset \mathbb{R}$ such that $\mu[\partial B] = 0$ we have $\mu_n[B] \xrightarrow[n\to\infty]{} \mu[B]$.

A roadmap of the proof



- Prove equivalence of (1°)-(5°) first, then separately (6°)
- The blue implications will be proven in a more general setup later.
- The red implications are proven below.
- The equivalence with (6°) is proven once we have discussed tightness. (later on in this lecture).

Proof of $(5°) \Rightarrow (2°)$:

Assume first that V is an open interval $V = (a,b)$.

Approximate (a,b) from inside:

choose (a_k), (b_k) s.t. $a_k \searrow a$, $b_k \nearrow b$ as $k \to \infty$

By monotone convergence $\mu[(a_k, b_k)] \nearrow \mu[(a,b)]$,

so for any $\varepsilon > 0$ we can find k_ε such that

$$\mu[(a_{k_\varepsilon}, b_{k_\varepsilon})] > \mu[(a,b)] - \varepsilon.$$

Note also that the increasing function F can have at most countably many points of discontinuity, so we can find continuity points a', b' s.t.

$$a < a' < b' < b \quad \text{and} \quad \mu[(a', b')] > \mu[(a,b)] - \varepsilon.$$

Then assuming $(5°)$ we have

$$\mu[(a,b)] - \varepsilon < \mu[(a', b')] = \lim_{b'' \nearrow b'} F(b'') - F(a')$$

$$\underset{\substack{\text{continuity}\\ \text{point } b'}}{=} F(b') - F(a') \underset{(5°)}{=} \lim_{n \to \infty} \left(F_n(b') - F_n(a') \right)$$

$$= \lim_{n \to \infty} \mu_n[(a', b']] \leq \liminf_{n \to \infty} \mu_n[(a,b)].$$

This proves $(2°)$ in the case V is an open interval.

In the general case V is an union of at most countably many disjoint open intervals V_j: $V = \bigcup_{j=1}^{\infty} V_j$ (or finite union, easier!)

By the argument above we have, for any $\varepsilon > 0$ that $\mu[V_j] - \varepsilon \cdot 2^{-j} \leq \liminf_{n \to \infty} \mu_n[V_j]$.

Summing over j and using Fatou's lemma we get

$$\mu[V] - \varepsilon \leq \liminf_{n \to \infty} \mu_n[V]. \qquad \square$$

Proof of $(4°) \Rightarrow (5°)$:

Note that $\partial((-\infty, x]) = \{x\}$ and $\mu[\{x\}] = 0$ iff x is a continuity point of F.

Thus by $(4°)$ for continuity points x of F we have $F_n(x) = \mu_n[(-\infty, x]] \xrightarrow[n \to \infty]{} \mu[(-\infty, x]] = F(x)$. $\square$

TIGHTNESS

Def: A collection $(\nu_\alpha)_{\alpha\in I}$ of probability measures on $\mathbb{R}$ is *tight* if for all $\varepsilon>0$ there exists an $r>0$ such that $\nu_\alpha[\mathbb{R}\setminus[-r,r]]<\varepsilon \quad \forall \alpha\in I$.

Intuition: We require that the overwhelming majority of the probability mass of all members of the collection is carried by the same compact set.

Remark For the corresponding c.d.f's $(F_\alpha)_{\alpha\in I}$ tightness means: $\forall\varepsilon>0 \ \exists r>0$ s.t. $F(r)>1-\varepsilon$, $F(-r)<\varepsilon$.

Exercise If $(F_n)_{n\in\mathbb{N}}$ is a tight sequence of c.d.f's, then there exists a subsequence $(F_{n_k})_{k\in\mathbb{N}}$ which converges pointwise to a c.d.f F at all continuity points of F.

By the exercise above we see that "tightness" is a sort of a (pre)compactness property for collection of measures: it allows us to extract subsequential limits.

Proposition Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of proba measures on $\mathbb{R}$, and let $(\chi_n)_{n\in\mathbb{N}}$ be their characteristic functions: $\chi_n(\theta)=\int_{\mathbb{R}} e^{i\theta x}\,d\mu_n(x)$. Then (μ_n) converges weakly iff $\chi_n(\theta)\longrightarrow\chi(\theta)$ for all $\theta\in\mathbb{R}$ and the limit $\chi:\mathbb{R}\to\mathbb{C}$ is continuous at 0.

Remark: This proves that (6°) is equivalent to the rest of the conditions in Theorem*. In fact the statement (6°) $\Rightarrow$ (others) is improved here: we don't need to assume a priori that χ is the characteristic function of some measure.

Remark: The proof of the non-trivial direction ("if") consists of two steps

1. Show precompactness (tightness)
   - every subsequence has a converging subsubsequence
2. Characterize explicitly any subsequential limit
   - all convergent subsequences have the same limit

It is easy to deduce from the combination of these two that the whole sequence converges.
This strategy works very very often!
(... so remember it!)

Proof of Proposition:

"only if": $x \mapsto e^{i\theta x}$ is continuous and bounded

so $\chi_n(\theta) = \int_{\mathbb{R}} e^{i\theta x} d\mu_n(x) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} e^{i\theta x} d\mu(x) =: \chi(\theta)$

assumption $\mu_n \xrightarrow{w} \mu$

continuous at $\theta = 0$ (char. fn of μ) (dominated convergence!)

"if" Suppose $\chi_n(\theta) \longrightarrow \chi(\theta)$ $\forall \theta \in \mathbb{R}$
and $\lim_{\theta \to 0} \chi(\theta) = 1$.

To show tightness, use the auxiliary calculation

$$\int_{-u}^{u} (1 - e^{i\theta x})\, d\theta = 2u - 2\,\frac{\sin(ux)}{x}.$$

Divide by u, integrate against μ_n and use Fubini

$$\frac{1}{u}\int_{-u}^{u} (1 - \chi_n(\theta))\, d\theta = 2\int_{\mathbb{R}} \underbrace{\Big(1 - \frac{\sin(ux)}{ux}\Big)}_{\geq 0 \text{ because } \left|\frac{\sin(\xi)}{\xi}\right| \leq 1} d\mu_n(x)$$

$$\geq 2 \int_{\mathbb{R}\setminus(\frac{-2}{u}, \frac{2}{u})} \Big(1 - \frac{1}{|ux|}\Big)\, d\mu_n(x)$$

we used $|\sin(\xi)| \leq 1$.

integrand is $\geq \frac{1}{2}$

$$\geq \mu_n\Big[\mathbb{R} \setminus \Big(\frac{-2}{u}, \frac{2}{u}\Big)\Big]$$

An upper bound for the mass of the complement of an interval written in terms of characteristic function

Let $\varepsilon > 0$.
By the assumption $\chi(\theta) \longrightarrow 1$ as $\theta \to 0$ we can find u small enough so that

$$\frac{1}{u}\int_{-u}^{u}(1-\chi(\theta))\,d\theta < \varepsilon .$$

By $\chi_n(\theta) \longrightarrow \chi(\theta)$ and dominated convergence for n large $\frac{1}{u}\int_{-u}^{u}(1-\chi_n(\theta))\,d\theta < 2\varepsilon$.

We get $\mu_n\left[\mathbb{R}\setminus(\frac{-2}{u},\frac{2}{u})\right] < 2\varepsilon$ for n large and tightness of $(\mu_n)_{n\in\mathbb{N}}$ follows.

By tightness we know that subsequential limits of $(\mu_n)_{n\in\mathbb{N}}$ exist. Suppose that $(\mu_{n_k})_{k\in\mathbb{N}}$ is a convergent subseq., $\nu_{n_k} \xrightarrow{w} \nu$. By the same argument as in "only if" we find that the characteristic function of ν is χ.
Therefore all convergent subsequences have the same limit. $\square$

CURIE-WEISS MODEL

- a model of ferromagnetism
  (or any phenomenon with interactions preferring alignment)
- simplified "mean field" version of the celebrated Ising model : no spacial structure of interactions
- correct qualitative phase transition
  "low temperature" $\longleftrightarrow$ "ferromagnetic"
  "high temperature" $\longleftrightarrow$ "paramagnetic"

THE CURIE-WEISS MODEL

- size parameter $n \in \mathbb{N}$: number of elementary magnets ("spins")
- sample space $\Omega_n = \{-1,+1\}^n$
- configuration $\sigma = (\sigma_j)_{j=1}^n \in \Omega_n$ declared to have energy

$$H_n(\sigma) = -\frac{1}{n}\sum_{i,j=1}^n \sigma_i\sigma_j - B\cdot\sum_{j=1}^n \sigma_j$$

(parameter $B \in \mathbb{R}$ "external magnetic field")

and the probability measure $\mathbb{P}_n$ on Ω_n defined by

$$\mathbb{P}_n[\{\sigma\}] = \frac{1}{Z}\cdot e^{-\beta H_n(\sigma)}$$

(parameter $\beta > 0$ "inverse temperature")

where the "partition function"

$$Z = Z_n(\beta, B) = \sum_{\sigma\in\Omega_n} e^{-\beta H_n(\sigma)}$$

normalizes $\mathbb{P}_n[\Omega_n] = 1$.

- if the individual elementary magnetizations σ_j are unimportant, the configuration $\sigma \in \Omega_n$ can be summarized with the random variable

$$M_n = \frac{1}{n}\sum_{j=1}^n \sigma_j \quad , \quad \text{"empirical magnetization".}$$

Note: This is the empirical mean of n elementary magnetizations, which are not independent!

THERMODYNAMICAL LIMIT OF CURIE-WEISS MODEL

$n \longrightarrow \infty$ "thermodyn. limit"

Two formulations of the phase transition at $\beta = \beta_c = \frac{1}{2}$, which indicate paramagnetic behavior for $\beta < \beta_c$ and ferromagnetic for $\beta > \beta_c$.

1°) without external magn. field, the thermodynamical limit of empirical magnetization is zero in the paramagnetic phase and a non-zero "spontaneous magnetization" with random sign in the ferromagnetic phase.

2°) with positive external magn. field $B > 0$, the thermodynamical limit of empirical magnetization is a constant $\tilde{m}(\beta, B) > 0$, which tends to zero or a positive value as the external field is removed, $B \searrow 0$, in the paramagn. and ferromagn. phases, respectively.

The most important quantitative information about the phase transition is critical exponents. We will show that near the critical point $(\beta = \beta_c, B = 0)$ the magnetization has power law behaviors: $\tilde{m}(\beta_c, B) \sim B^{1/\delta}$ $(\delta = 3)$

$$\tilde{m}(\beta, 0) \sim |\beta - \beta_c|^{b} \qquad (b = \tfrac{1}{2}).$$

Theorem 1: Fix $\beta > 0$ and $B = 0$. As $n \longrightarrow \infty$ the empirical magnetizations M_n converge in distribution to a random variable M_∞, with the following law

- if $\beta < \beta_c$ then $M_\infty = 0$ almost surely
- if $\beta > \beta_c$ then there exists a constant $\bar{m} = \bar{m}(\beta) > 0$ such that $\mathbb{P}[M_\infty = +\bar{m}] = \frac{1}{2} = \mathbb{P}[M_\infty = -\bar{m}]$.

Theorem 2: Fix $\beta > 0$ and $B > 0$. As $n \longrightarrow \infty$ the empirical magnetizations M_n converge in distribution to a constant $\tilde{m}(\beta, B)$. As $B \searrow 0$ we have

$$\lim_{B \searrow 0} \tilde{m}(\beta, B) = \begin{cases} \bar{m}(\beta) > 0 & \text{if } \beta > \beta_c \\ 0 & \text{if } \beta < \beta_c. \end{cases}$$

Theorem 3: The functions $\bar{m}$ and $\tilde{m}$ in Theorems 1 and 2 have the following asymptotics near the critical point ($\beta = \beta_c = \frac{1}{2}$, $B = 0$)

$$\lim_{\beta \searrow \beta_c} \frac{\bar{m}(\beta)}{|\beta - \beta_c|^b} \neq 0 \qquad \text{where } b = \frac{1}{2}$$

$$\lim_{B \searrow 0} \frac{\tilde{m}(\beta_c, B)}{B^{1/d}} \neq 0 \qquad \text{where } d = 3.$$

HELMHOLTZ FREE ENERGY AND LARGE DEVIATIONS

Begin by observing that energy H_n can be expressed in terms of empirical magnetization M_n:

$$H_n = -\frac{1}{n} \sum_{i,j=1}^{n} \sigma_i \sigma_j - B \cdot \sum_{j=1}^{n} \sigma_i$$

$$= -n \cdot \left(M_n^2 + B \cdot M_n \right) = n \cdot \Psi(M_n)$$

where $\Psi(m) = -m^2 - B \cdot m$,

Therefore the partition function reads

$$Z_n(\beta, B) = \sum_{m \in \mathcal{M}_n} z_n(m) \cdot e^{-\beta n \Psi(m)}$$

where $\mathcal{M}_n = \left\{ \frac{-n}{n}, \frac{2-n}{n}, \frac{4-n}{n}, \ldots, \frac{n-2}{n}, \frac{n}{n} \right\} = \begin{cases} [-1,1] \cap \frac{2}{n}\mathbb{Z} & \text{if } n \text{ even} \\ [-1,1] \cap \frac{2}{n}(\mathbb{Z} + \frac{1}{2}) & \text{if } n \text{ odd} \end{cases}$

is the set of all possible values of the empirical magnetization

and $z_n(m) = \#\left\{ \sigma \in \Omega_n \;\middle|\; \frac{1}{n}\sum_{j=1}^{n} \sigma_j = m \right\} = \binom{n}{n \cdot \frac{1+m}{2}} = \frac{n!}{(n\frac{1+m}{2})!\,(n\frac{1-m}{2})!}$

is the number of configurations whose empirical magn. equals m.

Exercise (Cramér entropy)

Show that $\log(z_n(m)) = n \cdot (\log(2) - I(m)) + o(n)$

where $I(m) := \frac{1+m}{2} \log(1+m) + \frac{1-m}{2} \log(1-m)$

Remark: $I : [-1,1] \to \mathbb{R}$ is continuous (use: $x \cdot \log(x) \to 0$ as $x \to 0$).

Using the exercise, write

$$Z_n(\beta,B) = \sum_{m\in\mathcal{M}_n} z_n(m)\, e^{-n\cdot\beta\cdot\Psi(m)}$$

$$= \sum_{m\in\mathcal{M}_n} \exp\Big(-n\cdot\beta\cdot\big(g(\beta,m) - B\cdot m\big) + o(n)\Big)$$

where $\boxed{g(\beta,m) = \frac{I(m) - \log(2)}{\beta} - m^2.}$

is called "(magnetic) Helmholz free energy" (per particle)

By the next proposition the Gibbs free energy $\boxed{f = \lim_{n\to\infty} \frac{-1}{\beta n}\log(Z_n)}$ is the Legendre transform of the Helmholz free energy.

<u>Proposition</u> Let $\beta>0$ and $B\in\mathbb{R}$. We have

$$f(\beta,B) := \lim_{n\to\infty}\left(\frac{-1}{\beta n}\log\big(Z_n(\beta,B)\big)\right) = \inf_{m\in[-1,1]}\big(g(\beta,m) - B\cdot m\big)$$

<u>Proof</u>: To prove the equality, we prove inequalities both ways.

From an upper bound for partition function, a lower bound will be derived for free energy:

$$Z_n(\beta,B) = \sum_{m\in\mathcal{M}_n} \exp\big[-n\cdot\beta\cdot(g(\beta,m) - Bm) + o(n)\big]$$

$$\leq (n+1)\cdot \max_{m\in\mathcal{M}_n}\Big(\exp\big[-n\cdot\beta\cdot(g(\beta,m) - B\cdot m)\big]\Big)$$

Take logarithms and divide by n, and collect $o(1)$ terms

$$\frac{1}{n}\log Z_n(\beta,B) \leq \beta\cdot \max_{m\in\mathcal{M}_n}\big(-g(\beta,m) - B\cdot m\big) + o(1)$$

$$\leq \beta\cdot \sup_{m\in[-1,1]}\big(-g(\beta,m) - B\cdot m\big) + o(1)$$

Now just divide by $-\beta$ and let $n\longrightarrow\infty$ to get

$$\liminf_{n\to\infty}\left(\frac{-1}{\beta n}\log Z_n(\beta,B)\right) \geq \inf_{m\in[-1,1]}\big(g(\beta,m) - B\cdot m\big).$$

To get an upper bound for free energy, estimate the partition function by the largest term of the sum

$$Z_n(\beta, B) \geq \max_{m \in \mathcal{M}_n} \left(\exp\left[-n\beta\left(g(\beta,m) - Bm\right) + o(n) \right] \right).$$

Again by taking log and dividing by n we get

$$\frac{1}{n} \log Z_n(\beta, B) \geq \beta \cdot \max_{m \in \mathcal{M}_n} \left(-g(\beta,m) + B\cdot m \right) + o(1).$$

The function $m \longmapsto g(\beta,m) + B\cdot m$ is continuous on the interval $[-1,1]$. By compactness it achieves its minimum. There are points in $\mathcal{M}_n$ at distance at most $\frac{2}{n}$ from the point which achieves the minimum, so by continuity

$$\min_{m \in \mathcal{M}_n} \left(g(\beta,m) - B\cdot m \right) \xrightarrow[n \to \infty]{} \inf_{m \in [-1,1]} \left(g(\beta,m) - Bm \right).$$

Again divide by $-\beta$ and let $n \longrightarrow \infty$ to get

$$\limsup_{n \to \infty} \left(\frac{-1}{\beta n} \log Z_n(\beta, B) \right) \leq \inf_{m \in [-1,1]} \left(g(\beta,m) - Bm \right).$$

Combining the two bounds, we conclude that the limit

$$f(\beta, B) = \lim_{n \to \infty} \left(\frac{-1}{\beta n} \log Z_n(\beta, B) \right)$$

exists and is given by the asserted formula. $\square$

Interpretation: The Helmholtz free energy gives the rate of large deviations for the empirical magnetization: roughly speaking the probability that M_n assumes approximately the value m is exponentially small:

$$\mathbb{P}_n^{(\beta,B)}\left[M_n \approx m \right] \sim \exp\left(-n\beta\left(g(\beta,m) - Bm - f(\beta,B) \right) \right).$$

A precise formulation is given next.

Proposition: For any open set $V \subset [-1,1]$ we have

$$\lim_{n\to\infty} \frac{-1}{n} \log \mathbb{P}[M_n \in V] = \beta \cdot \left(\inf_{m \in V} (g(\beta,m) - Bm) - f(\beta,B) \right)$$

Proof: The probability that $M_n \in V$ is

$$\mathbb{P}[M_n \in V] = \frac{1}{Z_n(\beta,B)} \sum_{m \in V \cap \mathcal{M}_n} z_n(m)\, e^{n\beta(m^2 + Bm)} .$$

Estimating the sum as in the previous proposition we get $\displaystyle \frac{1}{n} \log \Big(\sum_{m \in V \cap \mathcal{M}_n} z_n(m)\, e^{n\beta(m^2+Bm)} \Big) = -\beta \cdot \inf_{m \in V} \big(g(\beta,m) - Bm \big) + o(1)$.

The asserted formula follows by using simultaneously also the asymptotics of the partition function given in the previous proposition. □

We are almost ready to prove our main results: the minima of the Helmholtz free energy (i.e. the zeroes of the rate of large deviations) indicate where the probability mass of the law of M_n concentrates — elsewhere we get exponentially small probabilities.

We formulate two lemmas about these ideas.

Lemma 1: Suppose that $\mathbb{R}$-valued random variables $(X_n)_{n\in\mathbb{N}}$ converge in distribution, $X_n \xrightarrow[n\to\infty]{w} X$, and that they satisfy the large deviations upper bound

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n[X_n \in V] \leq -\inf_{x \in V} \phi(x) \qquad \forall\, V \subset \mathbb{R} \text{ open}$$

where the rate function $\phi: \mathbb{R} \to [0,\infty)$ is continuous.
Then the limit r.v. X takes values in the set $N = \{ x \in \mathbb{R} \mid \phi(x) = 0 \}$. $\left(\mathbb{P}[X \in N] = 1 \right)$

Proof: Define open sets $V_\varepsilon = \{x \in \mathbb{R} \mid \phi(x) > \varepsilon\}$, for $\varepsilon > 0$.
By continuity $\inf_{x \in \bar{V}_\varepsilon} \phi(x) \geq \varepsilon$, so we infer from large deviations upper bound that $\mathbb{P}_n[X_n \in V_\varepsilon] \xrightarrow[n\to\infty]{} 0$.
Then by characterization (ii) of Theorem* we get

$$\mathbb{P}[X \in V_\varepsilon] \leq \liminf_{n\to\infty} \mathbb{P}_n[X_n \in V_\varepsilon] = 0.$$

The complement of $N = \{x \in \mathbb{R} \mid \phi(x) = 0\}$ is a countable union of sets of form V_ε, $\varepsilon > 0$,
so we get $\mathbb{P}[X \notin N] = \mathbb{P}[X \in \bigcup_{k=1}^\infty A_{1/k}] \leq \underbrace{\sum_{k=1}^\infty \mathbb{P}[X \in A_{1/k}]}_{=0} = 0$

$\square$

Concerning the locations of the minima of the Helmholtz free energy we will use:

Lemma 2: For fixed $\beta > 0$ and $B = 0$, the rate of large deviations $\phi(m) = \beta(g(\beta,m) - f(\beta,0))$ satisfies:

(a) $\phi : [-1,1] \to \mathbb{R}$ is continuous and its minimum is 0.

(b) If $\beta \leq \frac{1}{2}$ then the unique zero of $m \mapsto \phi(m)$ occurs at $m = 0$.

(c) If $\beta > \frac{1}{2}$ then the equation $\phi'(m) = 0$ has a unique positive solution $m = \bar{m}(\beta)$ and the only zeroes of $m \mapsto \phi(m)$ are at $m = \pm\bar{m}(\beta)$.

Proof: (a) Continuity clear from that of $I : [-1,1] \to \mathbb{R}$, minimum value 0 by Proposition 1.

(b) $m \mapsto \phi(m)$ is in fact C^∞ on $(-1,1)$ so the minimum occurs at m such that $\phi'(m) = 0$ (it does not occur at $m = \pm 1$ by an easy check).
Moreover ϕ is even: $\phi(-m) = \phi(m)$, so $\phi'(0) = 0$.
An easy calculation gives $I''(m) = \frac{1}{1-m^2}$, so we read $\phi''(m) = \frac{1}{1-m^2} - 2\beta$. If $\beta < \frac{1}{2}$ then $\phi''(m) \geq 0$ and the only minimum is at $m = 0$.

(c) Similar: ϕ'' negative on an interval around 0 and positive elsewhere $\Rightarrow$ local max at 0 and minima at two other zeroes of ϕ' $\square$

Proof of Theorem 1:

We claim that the $\mathbb{R}$-valued random variables M_n converge in distribution to something. Note first that the sequence $(M_n)_{n\in\mathbb{N}}$ is tight, because $M_n \in [-1, 1]$ (all r.v.'s take values on the same bounded interval). It follows that any subsequence $(M_{n_k})_{k\in\mathbb{N}}$ has a further converging subsubsequence $(M_{n_{k_j}})_{j\in\mathbb{N}}$. It remains to be shown that any subsequential limit has the asserted form.

For $\beta \leq \beta_c = \frac{1}{2}$, if $M_{n_k} \xrightarrow{w} M$ as $k \to \infty$, we have by Lemma 1, Proposition on large deviations and Lemma 2 (b) that $\mathbb{P}[M=0] = 1$. Thus any convergent subsequence has the constant 0 as its limit, and the weak convergence to 0 follows.

For $\beta > \beta_c$, if $M_{n_k} \xrightarrow{w} M$ as $k \to \infty$, we have by Lemma 1, Proposition, and Lemma 2(c) that $\mathbb{P}[M \in \{+\bar{m}, -\bar{m}\}] = 1$. On the other hand, since $B=0$, we have symmetry $\mathbb{P}[M_n<0]=\mathbb{P}[M_n>0]$. Since $\mathbb{P}[M=0]=0$, we can use equivalent condition (iv) of weak convergence to deduce
$$\mathbb{P}[M<0] = \lim_{k\to\infty} \mathbb{P}[M_{n_k}<0] = \lim_{k\to\infty} \mathbb{P}[M_{n_k}>0] = \mathbb{P}[M>0].$$
Combining the observations, we get $\mathbb{P}[M=+\bar{m}] = \frac{1}{2} = \mathbb{P}[M=-\bar{m}]$. Any subsequential limit thus has the claimed law. We conclude the weak convergence of the entire sequence. $\square$

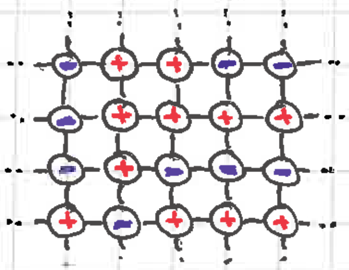
Proof of theorem 2 Similar. $\square$

Proof of theorem 3. Exercise. $\square$

WEAK CONVERGENCE ON METRIC SPACES

Most probability theory naturally needs quite general metric spaces, and luckily, most probability theory also works well on (complete separable) metric spaces.

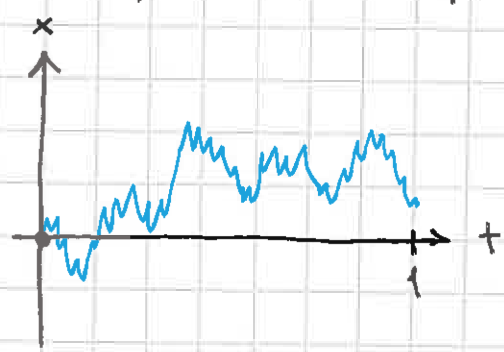
In this course the two examples to keep in mind are:



Ising model

$\Omega = \{-1,+1\}^{\mathbb{Z}^d}$

equipped with the (metrizable) product topology



Brownian motion

$\Omega = C([0,1])$

equipped with the L^∞-norm ($\rightsquigarrow$ metric)

METRIC SPACES

__Def:__ A metric space is a set $\mathfrak{X}$ equipped with a function ("metric") $\rho : \mathfrak{X} \times \mathfrak{X} \to [0,\infty)$ such that:

(Sep) $\quad \rho(x,y) = 0 \iff x = y$

(Sym) $\quad \rho(x,y) = \rho(y,x)$

(Tri) $\quad \rho(x,y) \le \rho(x,z) + \rho(z,y)$

for all $x,y,z \in \mathfrak{X}$.

Some notions in metric spaces

- <u>ball</u>: $x \in \mathcal{X}$, $r > 0$: $B_r(x) = \{y \in \mathcal{X} \mid \rho(y,x) < r\}$
- <u>open set</u>: $V \subset \mathcal{X}$ s.t. $\forall x \in V$ $\exists r > 0$ s.t. $B_r(x) \subset V$
- <u>closed set</u>: $A \subset \mathcal{X}$ s.t. $\mathcal{X} \setminus A$ is open
- <u>closure</u>: closure of $S \subset \mathcal{X}$ is $\overline{S} = \bigcap_{\substack{A \text{ closed} \\ A \supset S}} A$
- <u>interior</u>: interior of $S \subset \mathcal{X}$ is $S^\circ = \bigcup_{\substack{V \text{ open} \\ V \subset S}} V$
- <u>boundary</u>: $\partial S = \overline{S} \setminus S^\circ$
- <u>dense</u>: $S \subset \mathcal{X}$ is dense if $\overline{S} = \mathcal{X}$.
- <u>Cauchy sequence</u>: A sequence $(x_n)_{n \in \mathbb{N}}$ of points $x_n \in \mathcal{X}$ is Cauchy if $\forall \varepsilon > 0$ $\exists n_0 \in \mathbb{N}$ s.t. $\rho(x_n, x_m) < \varepsilon$ $\forall n,m \geq n_0$.

<u>Def</u>: A metric space $(\mathcal{X}, \rho)$ is

- <u>complete</u>, if all Cauchy sequences converge
- <u>separable</u>, if there exists a countable dense subset.

<u>Examples</u>

1°) $\mathbb{R}$ is complete and separable (w.r.t. $\rho(x,y) = |x-y|$)

2°) $\mathbb{Q}$ is not complete

3°) $\ell^\infty = \{ (a_n)_{n \in \mathbb{N}} \mid a_n \in \mathbb{R} \ \forall n, \ \exists M > 0 \text{ s.t. } |a_n| \leq M \ \forall n \}$

is not separable with respect to $\|(a_n)\|_\infty = \sup_{n \in \mathbb{N}} |a_n|$.

4°) $C([0,1]) = \{ f : [0,1] \to \mathbb{R} \mid f \text{ continuous} \}$

is complete and separable w.r.t.

$\|f\|_\infty = \sup_{t \in [0,1]} |f(t)|$

BOREL PROBABILITY MEASURES

In this section $\mathcal{X}$ is assumed to be a topological space (for example a metric space, and you are allowed to assume it is a metric space).

Def: The Borel σ-algebra $\mathcal{B} = \mathcal{B}(\mathcal{X})$ is the smallest σ-alg. that contains all open sets $V \subset \mathcal{X}$.

Remark: This makes sense. $\mathcal{B}$ can be "constructed" as the intersection of all σ-algebras that contain all open sets. The set of all subsets of $\mathcal{X}$ is one such σ-algebra, so the intersection is over a non-empty collection. It is easy to check that the intersection is still a σ-algebra.

Remark: $\mathcal{B}$ contains all closed sets, countable intersections of open sets, countable unions of closed sets, ...

Def: A Borel probability measure on $\mathcal{X}$ is a probability measure on the σ-algebra $\mathcal{B}(\mathcal{X})$.

Regularity of Borel probability measures on metric spaces

Now assume $(\mathcal{X}, \varrho)$ is metric space.

(This is "always" the case in practice. What sometimes happens is that no explicit metric is given, but the topology on $\mathcal{X}$ is anyway metrizable by some quite reasonable metric.)

In this setup any Borel probability measure is "regular": Borel sets can be approximated from below by closed sets and from above by open sets.

Let ν be a Borel probability measure on $(\mathcal{X}, \rho)$.

Proposition For any Borel set $E \subset \mathcal{X}$ and any $\varepsilon > 0$ there exists a closed set $F \subset \mathcal{X}$ and an open set $G \subset \mathcal{X}$ such that

$$F \subset E \subset G \quad \text{and} \quad \nu[G \setminus F] < \varepsilon.$$

Proof: Suppose first that E is closed. Then the open sets $G_\delta = \{x \in \mathcal{X} \mid \underbrace{\rho(x,E)}_{} < \delta\}$ approximate E from above

$\rho(x,E) = \inf_{y \in E} \rho(x,y)$

open as the preimage of open set of $\mathbb{R}$ under continuous $x \mapsto \rho(x,E)$.

$$G_{1/2} \supset G_{1/3} \supset G_{1/4} \supset \dots \supset E$$

and $\bigcap_{n=1}^{\infty} G_{1/n} = E$.

Therefore by monotone approximation of (finite) measures

$$\nu[E] = \lim_{n \to \infty} \nu[G_{1/n}]$$

and thus we can choose $F = E$ and $G = G_{1/n}$ for large enough n so that $\nu[G \setminus F] = \nu[G_{1/n}] - \nu[E] < \varepsilon$.

We conclude that all closed sets have the desired property. On the other hand, it is easy to check that the collection of all sets which have this property is a σ-algebra. Since $\mathcal{B}(\mathcal{X})$ is the smallest σ-alg. that contains all closed sets, the property holds for all $E \in \mathcal{B}(\mathcal{X})$. $\square$

We will also need the following trivial observation.

Lemma If $F \subset \mathcal{X}$ is closed and $\varepsilon > 0$ then there exists a continuous function $f : \mathcal{X} \to [0,1]$ such that

$$f(x) = 1 \quad \forall x \in F$$

and

$$f(x) = 0 \quad \forall x \in \mathcal{X} \text{ such that } \rho(x,F) \geq \varepsilon.$$

Proof If $F \neq \emptyset$ set $\rho(x,F) = \inf_{y \in F} \rho(x,y)$ and use for example $f(x) = \max\{0, 1 - \frac{\rho(x,F)}{\varepsilon}\}$. $\square$

Exercise Let $(\mathcal{X}, \varrho)$ be a metric space and ν_1, ν_2 two Borel probability measures on $\mathcal{X}$. Then either of the following is a sufficient condition for $\nu_1 = \nu_2$.

(i) For all closed sets $F \subset \mathcal{X}$ we have $\nu_1[F] = \nu_2[F]$.

(ii) For all bounded continuous $f : \mathcal{X} \to \mathbb{R}$ we have $\int_{\mathcal{X}} f(x)\, d\nu_1(x) = \int_{\mathcal{X}} f(x)\, d\nu_2(x)$.

Equivalent characterizations of weak convergence on $(\mathcal{X}, \varrho)$

Weak convergence is defined as in the real case.

Def Let ν_n, $n \in \mathbb{N}$, and ν be Borel probability measures on $\mathcal{X}$. Then $\nu_n \xrightarrow[n\to\infty]{w} \nu$ iff

$$\int_{\mathcal{X}} f(x)\, d\nu_n(x) \xrightarrow[n\to\infty]{} \int_{\mathcal{X}} f(x)\, d\nu(x) \quad \forall f \in C_b(\mathcal{X})$$

(Remark: Continuity of $f : \mathcal{X} \to \mathbb{R}$ depends only on topology — not on metric. So we can use whichever metric that gives the right topology.) ($C_b(\mathcal{X}) = \{ f : \mathcal{X} \to \mathbb{R}$ bounded continuous $\}$)

Remark: By part (ii) of the exercise above, if a weak limit ν of $(\nu_n)_{n \in \mathbb{N}}$ exist, it is unique.

Theorem The following are equivalent: (Portmanteau theorem)

(i) $\nu_n \xrightarrow{w} \nu$ (in the sense of the above def.)

(ii) For all open sets $G \subset \mathcal{X}$ we have $\nu[G] \leq \liminf_{n\to\infty} \nu_n[G]$.

(iii) For all closed sets $F \subset \mathcal{X}$ we have $\nu[F] \geq \limsup_{n\to\infty} \nu_n[F]$.

(iv) For all Borel sets $E \subset \mathcal{X}$ such that $\nu[\partial E] = 0$ we have $\nu[E] = \lim_{n\to\infty} \nu_n[E]$.

Map of proof

$$\begin{tikzcd}
& \text{(ii)} \arrow[d, Leftrightarrow] \arrow[dr, Rightarrow] & \\
& & \text{(iv)} \arrow[dl, Rightarrow, bend left] \\
\text{(i)} \arrow[r, Leftrightarrow] & \text{(iii)} \arrow[ur, Rightarrow] &
\end{tikzcd}$$

<u>Proof</u> (ii) $\Leftrightarrow$ (iii) : Obvious by setting $F = \mathcal{X} \setminus G$.

(ii) & (iii) $\Rightarrow$ (iv) : Suppose $E \in \mathcal{B}$ and $\nu[\partial E] = 0$.
Denote the closure by $\overline{E}$ and interior by E°.
We have $E^\circ \subset E \subset \overline{E}$ and since

$$\nu[\overline{E} \setminus E^\circ] = \nu[\partial E] = 0,$$

we have the equalities $\nu[E^\circ] = \nu[E] = \nu[\overline{E}]$.
Now E° is open and $\overline{E}$ is closed, so by the assumptions (ii) and (iii) we get

$$\begin{array}{ccc}
\liminf \nu_n[E^\circ] \geq \nu[E^\circ] = \nu[E] = \nu[\overline{E}] \geq & & \limsup \nu_n[\overline{E}] \\
\wedge\!| & & \vee\!| \\
\liminf \nu_n[E] & & \limsup \nu_n[E]
\end{array}$$

We get the existence of the limit $\lim_{n\to\infty} \nu_n[E] = \nu[E]$.

(i) $\Rightarrow$ (iii): Suppose that $F \subset \mathcal{X}$ is closed.
For any $\varepsilon > 0$ let $f_\varepsilon : \mathcal{X} \to [0,1]$ be a continuous function s.t. $f_\varepsilon(x) = 1 \quad \forall x \in F$
and $f_\varepsilon(x) = 0 \quad \forall x \notin F_\varepsilon = \{ y \in \mathcal{X} \mid \varrho(y, F) \leq \varepsilon \}$.
Then $f_\varepsilon \in C_b(\mathcal{X})$ and $\mathbb{1}_F \leq f_\varepsilon(x) \leq \mathbb{1}_{F_\varepsilon}$.
Therefore

$$\nu_n[F] = \int_{\mathcal{X}} \mathbb{1}_F \, d\nu_n \leq \int_{\mathcal{X}} f_\varepsilon \, d\nu_n \leq \int_{\mathcal{X}} \mathbb{1}_{F_\varepsilon} \, d\nu = \nu_n[F_\varepsilon].$$

We estimate, by assumption (i),

$$\limsup_{n\to\infty} \nu_n[F] \leq \limsup_{n\to\infty} \int_{\mathcal{X}} f_\varepsilon \, d\nu_n = \int_{\mathcal{X}} f_\varepsilon \, d\nu \leq \nu[F_\varepsilon].$$

Finally, $F_\varepsilon \downarrow F$ as $\varepsilon \searrow 0$ so $\nu[F_\varepsilon] \searrow \nu[F]$, and we conclude the property (iii).

(iii) $\Rightarrow$ (i): Let $f \in C_b(\mathcal{X})$. Assuming (iii) we will first prove

$$\limsup \int_{\mathcal{X}} f \, d\nu_n \leq \int_{\mathcal{X}} f \, d\nu$$

(adding a constant and multiplying by non-zero constant don't change the condition)

Assume without loss of generality that $0 < f(x) < 1 \quad \forall x \in \mathcal{X}$.

Fix $k \in \mathbb{N}$, and define the closed sets

$$F_j = \{ x \in \mathcal{X} \mid f(x) \geq \tfrac{j}{k} \} \qquad j = 0, 1, \ldots, k.$$

$F_j = f^{-1}([\tfrac{j}{k}, 1])$ preimage of closed interval under continuous f.

Estimate as follows:

$$\sum_{j=1}^{k} \frac{j-1}{k} \nu_n\left[\{x \mid \tfrac{j-1}{k} \leq f(x) < \tfrac{j}{k}\}\right] \leq \int_{\mathcal{X}} f(x) \, d\nu_n(x) \leq \sum_{j=1}^{k} \frac{j}{k} \nu_n\left[\{x \mid \tfrac{j-1}{k} \leq f(x) < \tfrac{j}{k}\}\right]$$

Rewrite the estimates using sets F_j, e.g. right hand side

$$\sum_{j=1}^{k} \frac{j}{k} \nu_n\left[\{\tfrac{j-1}{k} \leq f < \tfrac{j}{k}\}\right] = \sum_{j=1}^{k} \frac{j}{k} \left(\nu_n[F_{j-1}] - \nu_n[F_j] \right)$$

$$= \frac{1}{k} + \sum_{j=1}^{k} \nu_n[F_j]$$

Rewrite the left hand side estimate similarly to get

$$\frac{1}{k} \sum_{j=1}^{k} \nu_n[F_j] \leq \int_{\mathcal{X}} f \, d\nu_n \leq \frac{1}{k} + \frac{1}{k} \sum_{j=1}^{k} \nu_n[F_j].$$

Use assumption (iii): $\limsup_{n \to \infty} \nu_n[F_j] \leq \nu[F_j]$

$$\Longrightarrow \frac{1}{k} \sum_{j=1}^{k} \nu[F_j] \leq \limsup_{n \to \infty} \int_{\mathcal{X}} f \, d\nu_n \leq \frac{1}{k} + \frac{1}{k} \sum_{j=1}^{k} \nu[F_j]$$

As $k \to \infty$ we get $\limsup_{n \to \infty} \int_{\mathcal{X}} f \, d\nu_n \leq \int_{\mathcal{X}} f \, d\nu$.

But the same holds for the function $-f$ as well:

$$-\liminf_{n \to \infty} \int_{\mathcal{X}} f \, d\nu_n \leq - \int_{\mathcal{X}} f \, d\nu$$ so we have

$$\limsup_{n \to \infty} \int_{\mathcal{X}} f \, d\nu_n \leq \int_{\mathcal{X}} f \, d\nu \leq \liminf_{n \to \infty} \int_{\mathcal{X}} f \, d\nu_n.$$

This proves $\lim_{n \to \infty} \int_{\mathcal{X}} f(x) \, d\nu_n(x) = \int_{\mathcal{X}} f(x) \, d\nu(x).$

(iv) $\Longrightarrow$ (iii): Let F be closed and set, for $\delta > 0$

$$F_\delta = \{ x \in \mathfrak{X} \mid \varrho(x, F) \leq \delta \}.$$

Then $\partial F_\delta \subset \{ x \in \mathfrak{X} \mid \varrho(x, F) = \delta \}$.

The complement $\mathfrak{X} \setminus F$ is disjoint union $\bigcup_{\delta > 0} \{ x \in \mathfrak{X} \mid \varrho(x, F) = \delta \}$ so only countably many of these sets can have positive ν-measure.

We can thus find a sequence $(\delta_k)_{k=1}^{\infty}$, $\delta_k \searrow 0$, such that $\nu[\partial F_{\delta_k}] = 0 \quad \forall k$.

Assuming (iv) we have $\lim_{n \to \infty} \nu_n[F_{\delta_k}] = \nu[F_{\delta_k}]$.

Then

$$\limsup_{n \to \infty} \nu_n[F] \leq \lim_{n \to \infty} \nu_n[F_{\delta_k}] = \nu[F_{\delta_k}].$$

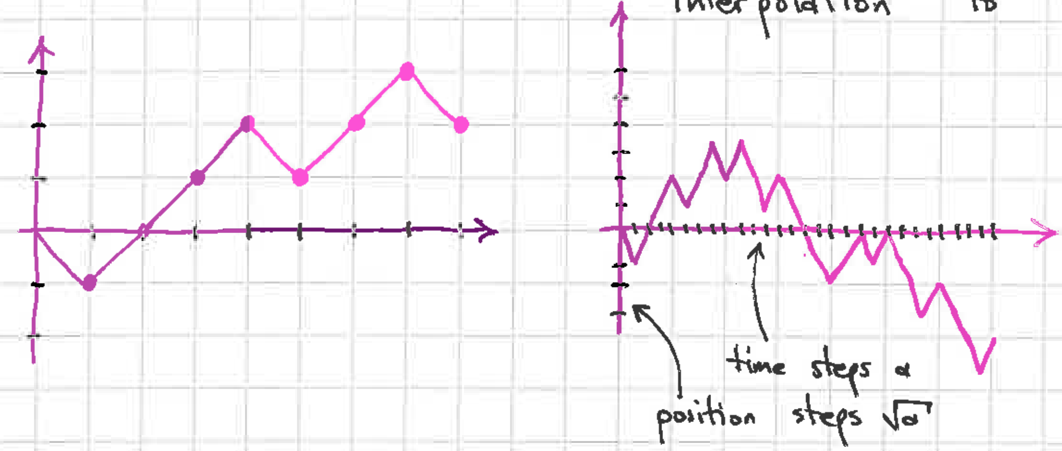
But by monotone appoximation $\left(F = \bigcap_{k=1}^{\infty} F_{\delta_k} \right)$ we have $\nu[F_{\delta_k}] \searrow \nu[F]$ as $k \to \infty$.

$\square$

DONSKER'S THEOREM

Random walk:

- steps $(\xi_j)_{j\in\mathbb{N}}$ i.i.d. $\mathbb{P}[\xi_j=+1]=\frac{1}{2}=\mathbb{P}[\xi_j=-1]$
- $S_n=\sum_{j=1}^{n}\xi_j$ walker position at integer times
- $\tilde{S}_t = S_{\lfloor t\rfloor} + (t-\lfloor t\rfloor)\cdot\xi_{\lfloor t\rfloor+1}$ piecewise linear interpolation to $t\in[0,\infty)$.



Rescaled walk: $\alpha>0$ scale parameter $(\alpha\searrow 0)$

$$X^{(\alpha)}_t = \sqrt{\alpha}\,\tilde{S}_{t/\alpha}$$

Our goal is the following scaling limit result for random walks:

<u>Donsker's theorem</u> As $\alpha\searrow 0$, the rescaled random walks $X^{(\alpha)}$ converge weakly to Brownian motion, on a space of continuous functions.

Proof strategy: 1°) precompactness (tightness)
2°) identification of limit

Before the proof we define and study Brownian motion.

ON BROWNIAN MOTION

Recall: Gaussian $Z \sim N(\mu, \sigma^2)$

characteristic function $\mathcal{X}(\theta) = \mathbb{E}[e^{i\theta Z}] = e^{i\theta\mu - \frac{1}{2}\sigma^2\theta^2}$.

Remark: naturally allows $\mu \in \mathbb{R}$, $\sigma \geq 0$

if $\sigma = 0$ then $Z = \mu$ (a.s.)

Gaussian vector $V = (V_1, \dots, V_n)$ vector valued r.v.

is Gaussian if for all vectors $A = (a_1, \dots, a_n) \in \mathbb{R}^n$

the "projection" $A \cdot V = \sum_{i=1}^n a_i V_i$ is $\mathbb{R}$-valued Gaussian.

Note: Characterized by mean $M = (\mathbb{E}[V_1], \dots, \mathbb{E}[V_n]) \in \mathbb{R}^n$

and covariance $C \in \mathbb{R}^{n \times n}$, $C_{ij} = \mathbb{E}[V_i V_j] - \mathbb{E}[V_i]\mathbb{E}[V_j]$.

Remark: similarly, allows for "degenerate" C if the random vector a.s. belongs to an affine subspace of $\mathbb{R}^n$.

Note: Components $(V_i)_{i=1}^n$ independent $\iff$ $C_{ij} = 0$ for $i \neq j$.

Gaussian process: $(X_t)_{t \in [0,\infty)}$ stochastic process

= collection of $\mathbb{R}$-valued r.v.'s indexed by "time" $t \in [0,\infty)$.

is Gaussian if

for any finite set of times, $0 \leq t_1 < t_2 < \dots < t_n$

the values $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ form a Gaussian vector.

Stationary process: $(X_t)_{t \in [0,\infty)}$ is stationary, if for

any $s \geq 0$ and any $0 \leq t_1 < t_2 < \dots < t_n$

the law of $(X_{s+t_1} - X_s, \dots, X_{s+t_n} - X_s)$ is the same

as that of $(X_{t_1} - X_0, \dots, X_{t_n} - X_0)$.

Independent increments $(X_t)_{t \in [0,\infty)}$ has independent increments

if for any $0 = t_0 < t_1 < \dots < t_n$, the

collection $\big(\underbrace{X_{t_j} - X_{t_{j-1}}}_{\text{"increment from time } t_{j-1} \text{ to } t_j\text{"}}\big)_{j=1,2,\dots,n}$ of r.v.'s is independent.

Example (Poisson process) $N = (N_t)_{t \in [0,\infty)}$ s.t.

- $N_0 = 0$
- $t \mapsto N_t$ is non-decreasing, right-continuous
- $N_{t+s} - N_s \sim \text{Poisson}(\lambda t)$ $\forall\, t,s \geq 0$.

Can be constructed from $(\tau_j)_{j \in \mathbb{N}}$ i.i.d. $\tau_j \sim \text{Exp}(\lambda)$ waiting times as $N_t = \sup\{ n \in \mathbb{Z}_{\geq 0} \mid \sum_{j=1}^{n} \tau_j \leq t \}$.
The Poisson process $(N_t)_{t \in [0,\infty)}$ is a stationary process with independent increments

Proposition For a stochastic process $X = (X_t)_{t \in [0,\infty)}$ the following are equivalent:

(i) X is stationary and has independent increments, and $X_t \sim N(0,t)$.

(ii) X is a Gaussian process with $\mathbb{E}[X_t] = 0$ and $\mathbb{E}[X_s X_t] = \min(s,t)$.

Notation: $\min(s,t) = s \wedge t$ and $\max(s,t) = s \vee t$
(common and convenient notation for stoch. processes)

Proof: (i) $\Rightarrow$ (ii): Let $0 = t_0 < t_1 < \dots < t_n$ and $\alpha_1, \dots, \alpha_n \in \mathbb{R}$.
Then $\sum_{k=1}^{n} \alpha_k X_{t_k} = \sum_{k=1}^{n} b_k (X_{t_k} - X_{t_{k-1}})$ $\leftarrow X_{t_0} = X_0 = 0$ since $X_0 \sim N(0,0)$
for certain coefficients b_k. $(b_n = \alpha_n,\ b_{n-1} = \alpha_{n-1} + \alpha_n, \dots)$
Since these increments $(X_{t_k} - X_{t_{k-1}})$ independent Gaussians, the linear combination is Gaussian.
Using independence we calculate the covariance, $s<t$,

$$\mathbb{E}[X_s X_t] = \mathbb{E}[X_s \cdot (X_s + X_t - X_s)] = \mathbb{E}[X_s]^2 + \mathbb{E}[X_s]\,\mathbb{E}[X_t - X_s]$$
$$= s + 0 \cdot 0 = s = \min(s,t).$$

Thus (ii) follows.

(ii) $\Rightarrow$ (i): Suppose $(X_t)_{t\in[0,\infty)}$ is a Gaussian process with $\mathbb{E}[X_t]=0$ $\mathbb{E}[X_t X_s] = s\wedge t$.

Then the increment $X_t - X_s$ is Gaussian with mean $\mathbb{E}[X_t - X_s] = 0-0=0$ and variance

$$\mathbb{E}[(X_t-X_s)^2] = \mathbb{E}[X_t^2]+\mathbb{E}[X_s^2]-2\,\mathbb{E}[X_s X_t]$$
$$= t+s-2(s\wedge t) = t-s \text{ .}$$

say $s\le t$.

So indeed $X_t \sim N(0,t)$, and $X_t - X_s \sim X_{t-s}$.

Finally consider, for $0\le t_1<t_2<\dots<t_n$ the vector $(X_{t_2}-X_{t_1}, X_{t_3}-X_{t_2}, \dots, X_{t_n}-X_{t_{n-1}})$ which is Gaussian. The covariances are (assume $1<j\le k\le n$)

$$\mathbb{E}[(X_{t_j}-X_{t_{j-1}})(X_{t_k}-X_{t_{k-1}})] = (t_j\wedge t_k)+(t_{j-1}\wedge t_{k-1}) - (t_{j-1}\wedge t_k) - (t_j\wedge t_{k-1})$$
$$= t_j + \cancel{t_{j-1}} - \cancel{t_{j-1}} - (t_j\wedge t_{k-1})$$
$$= \begin{cases} 0 & \text{if } j\neq k \\ t_k - t_{k-1} & \text{if } j=k \text{ .}\end{cases}$$

Since the non-diagonal covariances vanish, the Gaussian components (here increments) are independent. The same calculation shows stationarity. $\square$

<u>Continuous process</u>: $X=(X_t)_{t\in[0,\infty)}$ is a continuous process if $\mathbb{P}[\{\omega\in\Omega \mid t\mapsto X_t(\omega) \text{ is continuous } [0,\infty)\to\mathbb{R}\}]=1$.

<u>Def</u>: A process $B=(B_t)_{t\in[0,\infty)}$ is a standard Brownian motion if the following conditions hold

(1) B is a continuous process

(2) B satisfies the equivalent conditions of the previous proposition.

<u>Comments</u>:
- $\{\omega \mid t\mapsto X_t(\omega) \text{ continuous}\}$ must be measurable. It depends on uncountably many values of the process, so $(\Omega,\mathcal{F},\mathbb{P})$ can't be chosen carelessly!
- It is not a priori obvious that the conditions can hold simultaneously. For example differentiability [illegible]

Comments :

- The event $\{\omega \in \Omega \mid t \mapsto B_t(\omega) \text{ continuous}\}$ must first of all be measurable. It depends on X_t for uncountably many t, so $(\Omega, \mathcal{F}, \mathbb{P})$ can not be chosen carelessly !
- It is not a priori clear that the two requirements (continuity & law of increments) can hold simultaneously. For example differentiability can not !
- We have so far only talked about "finite dimensional distributions" of the process, i.e., joint laws of process values at finitely many times, $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$. These alone might not be enough to determine the law of the process.

A counterexample Suppose that $B = (B_t)_{t \in [0,\infty)}$ is a "good" Brownian motion, for which the two conditions (continuity & law of increments) hold. Let τ be positive r.v. with continuous law, independent of B. Define $\tilde{B} = (\tilde{B}_t)_{t \in [0,\infty)}$ by

$$\tilde{B}_t = \begin{cases} B_t & \text{if } t \neq \tau \\ B_t + 1 & \text{if } t = \tau \end{cases}$$

Then $\tilde{B}$ has the same finite-dimensional distributions as B, since $\mathbb{P}[\tau \in \{t_1, t_2, \dots, t_n\}] = 0$.
However, if $\mathbb{P}[\{t \mapsto B_t \text{ cont.}\}] = 1$ then

$$\mathbb{P}[\{t \mapsto \tilde{B}_t \text{ cont.}\}] = 0.$$

This should serve as a warning :
we must be careful !

SPACE OF CONTINUOUS FUNCTIONS

Donsker's theorem concerns weak convergence of scaled random walks $X^{(d)}$ to Brownian motion on the space

$$C([0,1]) = \{ f : [0,1] \to \mathbb{R} \mid f \text{ continuous} \}$$

equipped with

norm $\quad \|f\|_\infty = \sup_{t \in [0,1]} |f(t)|$

metric $\quad \rho(f_1, f_2) = \|f_1 - f_2\|_\infty$.

Proposition: $C([0,1])$ is a complete separable metric space.

Sketch of proof: "separability": countable dense set e.g. by piecewise linear interpolations of functions having rational values at finite sets of rational points

"completeness" $(f_n)_{n\in\mathbb{N}}$ Cauchy $\Rightarrow$ $(f_n(t))_{n\in\mathbb{N}}$ Cauchy $\forall t \in [0,1]$
$\Rightarrow$ pointwise limit $\lim_{n\to\infty} f_n(t) = \varphi(t)$ exists.
Standard argument (uniform limit of continuous functions) shows $\varphi \in C([0,1])$ and $\|f_n - \varphi\|_\infty \to 0$. $\square$

Finite dimensional distributions and Borel σ-algebra

We equip $C([0,1])$ with $\mathcal{B}$, the Borel σ-algebra.

Consider the σ-algebra $\mathcal{F}$ generated by finite-dimensional events of type $\{ f \in C([0,1]) \mid f(t_1) \in A_1, \dots, f(t_n) \in A_n \}$

for $0 \le t_1 < t_2 < \dots < t_n \le 1$ and $A_1, \dots, A_n \subset \mathbb{R}$ Borel.

Fix e.g. $\varphi \in C([0,1])$ and $r > 0$, and let $A_j = B_r(\varphi(t_j))$

$\Rightarrow$ set $\{ f \in C([0,1]) \mid |f(t_j) - \varphi(t_j)| \le r \ \ \forall j = 1, \dots, n \}$ is in $\mathcal{F}$.

Countable intersections, $t_j \in \mathbb{Q} \cap [0,1]$,

$\Rightarrow$ set $\{ f \in C([0,1]) \mid |f(t) - \varphi(t)| \le r \ \ \forall t \in \mathbb{Q} \cap [0,1] \}$

continuity $\rightarrow$ $= \{ f \in C([\quad \mid \|f - \varphi\|_\infty \le r \} = \overline{B}_r(\varphi) \subset C([0,1])$

These generate the Borel σ-alg. $\mathcal{B}$ of separable $C([0,1])$.

COMPACTNESS AND TIGHTNESS

$\mathcal{X}$ topological space

Def: $\mathcal{X}$ is

- sequentially compact if every sequence $(x_n)_{n\in\mathbb{N}}$ of points $x_n \in \mathcal{X}$ has a convergent subseq. $(x_{n_k})_{k\in\mathbb{N}}$ $x_{n_k} \longrightarrow x \in \mathcal{X}$.
- compact if any open cover ($(U_i)_{i\in I}$ $U_i \subset \mathcal{X}$ open $\bigcup_{i\in I} U_i = \mathcal{X}$) has a finite subcover $\mathcal{X} = U_{i_1} \cup \dots \cup U_{i_n}$ for some $i_1,\dots,i_n \in I$.

Proposition: A metric space $(\mathcal{X}, \varrho)$ is compact iff it is sequentially compact.

Proof: Standard (see other courses or textbooks) $\square$

It is convenient to call a subset $A \subset \mathcal{X}$ precompact if any sequence $(x_n)_{n\in\mathbb{N}}$, $x_n \in A$, has a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ which converges in $\mathcal{X}$ (not necessarily in A), $x_{n_k} \longrightarrow x \in \mathcal{X}$.
Equivalently, A is precompact iff the closure $\bar{A} \subset \mathcal{X}$ is sequentially compact.

Def: A collection $(\mu_\alpha)_{\alpha\in\mathcal{A}}$ of probability measures on $\mathcal{X}$. is precompact if any sequence $(\mu_{\alpha_n})_{n\in\mathbb{N}}$ has a weakly convergent subsequence.
The collection $(\mu_\alpha)_{\alpha\in\mathcal{A}}$ is tight if for any $\varepsilon > 0$ there exists a compact subset $K \subset \mathcal{X}$ such that

$$\mu_\alpha[K] \geq 1-\varepsilon \quad \forall \alpha \in \mathcal{A}$$

We will later show that tightness implies precompactness for (μ_α) and the converse holds if $\mathcal{X}$ is a complete separable metric space

(PRE) COMPACTNESS AND TIGHTNESS IN $C([0,1])$

The proof of Donsker's theorem consists of two steps

(1°) Show precompactness of the laws of the scaled random walks $X^{(a)}$, $a>0$.

$\Longrightarrow$ can extract weakly converging subsequences

(2°) Identify any subsequential limit as Brownian motion.

Part (1°) amounts to checking tightness.
(of the laws of $X^{(a)}$, viewed as proba measures on $C([0,1])$)

We thus need to understand compact subsets $K \subset C([0,1])$.

For this purpose, introduce the "modulus of continuity"

for $f \in C([0,1])$ and $\delta > 0$ set $\boxed{w_f(\delta) = \sup_{\substack{t,s \in [0,1] \\ |t-s| \leq \delta}} |f(t) - f(s)|}$

Note: for any $f \in C([0,1])$ $\lim_{\delta \searrow 0} w_f(\delta) = 0$, since f is uniformly continuous on the compact set $[0,1]$.

Def: A family $\Phi \subset C([0,1])$ of functions is equicontinuous if

$$\lim_{\delta \searrow 0} \sup_{f \in \Phi} w_f(\delta) = 0.$$

Arzelà - Ascoli theorem says that Φ is precompact iff it is equicontinuous and uniformly bounded.

<u>Theorem</u> (Arzelà - Ascoli theorem)

A subset $\Phi \subset C([0,1])$ is precompact iff the following conditions hold

1°) $\sup_{f\in\Phi} |f(0)| < \infty$

2°) $\lim_{\delta\searrow 0} \sup_{f\in\Phi} \omega_f(\delta) = 0$.

<u>Proof</u>: "only if": Suppose Φ is precompact.
Since $f \longmapsto f(0)$ is continuous $C([0,1]) \to \mathbb{R}$, it is bounded on the compact $\overline{\Phi}$, so 1° holds.
Likewise, $f \longmapsto \omega_f(\delta)$ is continuous, so the sets $\{f \in C([0,1]) \mid \omega_f(\delta) < \varepsilon\}$ are open, for $\varepsilon > 0$.
Also, for any f, $\omega_f(\delta) \longrightarrow 0$ as $\delta \searrow 0$, so for any $\varepsilon > 0$ these open sets cover $C([0,1])$.
By compactness, a finite number of them suffices to cover $\overline{\Phi}$. Let δ_0 be the minimal of the needed $\delta > 0$. Then $\sup_{f\in\mathcal{F}} \omega_f(\delta_0) < \varepsilon$.
We see that $\lim_{\delta\searrow 0} \sup_{f\in\Phi} \omega_f(\delta_0) = 0$.

"if": Assume Φ satisfies 1° and 2°.
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence, $f_n \in \Phi$.
Let $(t_j)_{j\in\mathbb{N}}$ be a countable dense set of points $t_j \in [0,1]$, for example an enumeration of $\mathbb{Q} \cap [0,1]$.
By boundedness of $(f_n(t_j))_{n\in\mathbb{N}}$ and diagonal extraction, we can choose a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ such that for all $j \in \mathbb{N}$ the sequence $(f_{n_k}(t_j))_{k\in\mathbb{N}}$ converges, $f_{n_k}(t_j) \longrightarrow \varphi(t_j)$. We then use equicontinuity of $\mathcal{F}$.
Let $\varepsilon > 0$. For some $\delta > 0$ we have $\omega_f(\delta) < \varepsilon$ $\forall f \in \mathcal{F}$. Choose finitely many points $s_1, \dots, s_m$ among $t_1, t_2, \dots$ so that $[0,1] \subset \bigcup_{j=1}^m (s_j - \delta, s_j + \delta)$, which is possible by compactness of $[0,1]$.
For large enough k_0 we have $|f_{n_k}(s_j) - f_{n_{k'}}(s_j)| < \varepsilon$ $\forall k, k' \geq k_0$, and for all $j = 1, \dots, m$. Therefore, for any $t \in [0,1]$ finding s_j s.t. $|t - s_j| < \delta$ we have
$$|f_{n_k}(t) - f_{n_{k'}}(t)| \leq |f_{n_k}(t) - f_{n_k}(s_j)| + |f_{n_k}(s_j) - f_{n_{k'}}(s_j)| + |f_{n_{k'}}(s_j) - f_{n_{k'}}(t)| < 3\varepsilon.$$

Arzelà - Ascoli theorem gives the following explicit conditions for tightness of probability measures on $C([0,1])$.

<u>Proposition</u> A sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ on $C([0,1])$ is tight if and only if the following conditions hold

1°) $\forall \varepsilon>0 \;\; \exists M>0, n_0>0$ s.t. $\mu_n\left[\{f \mid |f(0)|\geq M\}\right] \leq \varepsilon \;\; \forall n\geq n_0$

2°) $\forall \varepsilon_1,\varepsilon_2>0 \;\; \exists \delta>0$ and $n_0>0$ s.t.

$$\mu_n\left[\{f \mid w_f(\delta)\geq \varepsilon_1\}\right] \leq \varepsilon_2 \qquad \forall n\geq n_0 .$$

<u>Proof</u>: "only if": Assume $(\mu_n)_{n\in\mathbb{N}}$ is tight. Choose $K\subset C([0,1])$ compact such that $\mu_n[K]\geq 1-\varepsilon \;\; \forall n\in\mathbb{N}$.
Then by Arzelà - Ascoli theorem for large enough $M>0$
$K\subset\{f \mid |f(0)|\leq M\}$ and for small enough $\delta>0$
$K\subset\{f \mid w_f(\delta)\leq\varepsilon_1\}$. The conditions 1° and 2° follow.

"if": Assume 1° and 2°, and prove tightness.
(We may use $n_0=1$ by increasing M and decreasing δ, if necessary.)
Let $\varepsilon>0$. Choose $M>0$ s.t. $B=\{f \mid |f(0)|\leq M\}$ satisfies $\mu_n[B]\geq 1-\varepsilon \;\; \forall n$. For every $k\in\mathbb{Z}_{>0}$ choose $\delta_k>0$ s.t. $D_k=\{f \mid w_f(\delta_k)\leq 1/k\}$ satisfies $\mu_n[D_k]\geq 1-\varepsilon\cdot 2^{-k}$. Then the closure K of the set $B\cap\bigcap_{k\in\mathbb{Z}_{>0}} D_k$ satisfies $\mu_n[K]\geq 1-2\varepsilon$.
The set $B\cap\bigcap_{k\in\mathbb{Z}_{>0}} D_k$ satisfies the conditions of Arzelà - Ascoli theorem, so its closure K is compact

$\square$

PROKHOROV'S THEOREM

Recall that a family $(\mu_\alpha)_{\alpha\in\mathcal{A}}$ of probability measures on a topological space $\mathcal{X}$ is called

- tight: if for every $\varepsilon > 0$ there exists a compact subset $K \subset \mathcal{X}$ such that (T)
$$\mu_\alpha[K] \geq 1-\varepsilon \qquad \forall \alpha\in\mathcal{A}$$

- precompact: if for any sequence $(\mu_{\alpha_n})_{n\in\mathbb{N}}$ of measures from the family, there exists a subsequence $(\mu_{\alpha_{n_k}})_{k\in\mathbb{N}}$ which converges weakly. (P)

Prokhorov's theorem relates tightness and precompactness.

Theorem (Prokhorov's theorem)

Let $(\mathcal{X}, \varrho)$ be a metric space, and $(\mu_\alpha)_{\alpha\in\mathcal{A}}$ a family of probability measures on $\mathcal{X}$.
If $(\mu_\alpha)_{\alpha\in\mathcal{A}}$ is tight, then it is precompact.
If moreover $(\mathcal{X}, \varrho)$ is complete and separable, then if $(\mu_\alpha)_{\alpha\in\mathcal{A}}$ is precompact, then it is tight.

Remark: For proving weak convergence, the more important direction is $(T) \Longrightarrow (P)$.

We consider the following special case of $(T)\Rightarrow(P)$ known. It follows, e.g., by combining Alaoglu's theorem and Riesz representation theorem, which are standard topics in functional analysis.

Fact: If $(\mathcal{X}, \varrho)$ is a compact metric space and $(\nu_\alpha)_{\alpha\in\mathcal{A}}$ is a family of Borel measures of bounded total mass (i.e., $\exists M > 0$ s.t. $\nu_\alpha[\mathcal{X}] \leq M \quad \forall \alpha\in\mathcal{A}$) then any sequence from the family has a weakly convergent subseq.

Proof of $(T) \Longrightarrow (P)$: Assume $(\mu_\alpha)_{\alpha \in \mathcal{A}}$ is tight.

Choose compact sets $K_1 \subset K_2 \subset K_3 \subset \dots \subset \mathcal{X}$ s.t.

$$\mu_\alpha[K_m] \geq 1 - \frac{1}{m} \qquad \forall \alpha \in \mathcal{A}, \quad m \in \mathbb{Z}_{>0}.$$

Consider a sequence $(\mu_{\alpha_n})_{n \in \mathbb{N}}$ of measures from the collection. By the previous fact and diagonal extraction we can find a subsequence $(\mu_{\alpha_{n_k}})_{k \in \mathbb{N}}$ such that the restrictions to K_m converge, for all m, to some finite measure $\mu^{(m)}$ on K_m. Note that

$$\mu^{(m)}[K_m] \geq \limsup_{k \to \infty} \mu_{\alpha_{n_k}}[K_m] \geq 1 - \frac{1}{m}.$$

Let $E \subset \mathcal{X}$ be a Borel set. We will define $\mu[E] = \lim_{m \to \infty} \mu^{(m)}[E \cap K_m]$. To show that the limit exists, we show that $(\mu^{(m)}[E \cap K_m])_{m \in \mathbb{N}}$ is an increasing sequence.

So for $m_1 < m_2$ we want to show

$$\mu^{(m_1)}[E \cap K_{m_1}] \leq \mu^{(m_2)}[E \cap K_{m_2}].$$

We may assume $E \subset K_{m_1}$. Let $\varepsilon > 0$. Use regularity of measures to find (relatively) open subsets $G_1 \subset K_{m_1}$, $G_2 \subset K_{m_2}$ and closed $F_1 \subset K_{m_1}$, $F_2 \subset K_{m_2}$ such that

$$F_1 \subset E \subset G_1, \qquad \mu^{(m_1)}[G_1 \setminus F_1] < \varepsilon$$

$$F_2 \subset E \subset G_2, \qquad \mu^{(m_2)}[G_2 \setminus F_2] < \varepsilon.$$

Note that $G_1 = \hat{G}_1 \cap K_{m_1}$, where $\hat{G}_1 \subset K_{m_2}$ is open.

Set $G = \hat{G}_1 \cap G_2$ and $F = F_1 \cup F_2$.

Take a continuous function $f : K_{m_2} \to \mathbb{R}$ s.t. $\mathbb{1}_F \leq f \leq \mathbb{1}_G$. Then

$$\left| \mu^{(m_1)}[E] - \int_{K_{m_1}} f \, d\mu^{(m_1)} \right| < \varepsilon$$

$$\left| \mu^{(m_2)}[E] - \int_{K_{m_2}} f \, d\mu^{(m_2)} \right| < \varepsilon$$

Note that $\int_{K_{m_1}} f \, d\mu_{\alpha_{n_k}} \leq \int_{K_{m_2}} f \, d\mu_{\alpha_{n_k}}$

$$\begin{array}{ccc} \int_{K_{m_1}} f \, d\mu_{\alpha_{n_k}} & \leq & \int_{K_{m_2}} f \, d\mu_{\alpha_{n_k}} \\ \downarrow {\scriptstyle k\to\infty} & & \downarrow {\scriptstyle k\to\infty} \\ \int_{K_{m_1}} f \, d\mu^{(m_1)} & \leq & \int_{K_{m_2}} f \, d\mu^{(m_2)} \end{array}$$

Therefore $\mu^{(m_1)}[E] \leq \mu^{(m_2)}[E] + 2\varepsilon$, and taking $\varepsilon \searrow 0$ we get that $(\mu^{(m)}[E \cap K_m])_{m\in\mathbb{N}}$ is increasing.

We then show that μ defined by $\mu[E] = \lim_{m\to\infty} \mu^{(m)}[E \cap K_m]$ is a probability measure. The total mass is one, since $\mu[\mathcal{X}] = \lim_{m\to\infty} \mu^{(m)}[K_m] = 1$. We need to prove σ-additivity. Let $E = \bigcup_{i=1}^{\infty} E_i$ be a disjoint union. Then

$$\mu[E] = \lim_{m\to\infty} \mu^{(m)}\left[\left(\bigcup_{i=1}^{\infty} E_i\right) \cap K_m\right]$$

$$\geq \lim_{m\to\infty} \mu^{(m)}\left[\left(\bigcup_{i=1}^{n} E_i\right) \cap K_m\right] = \lim_{m\to\infty} \sum_{i=1}^{n} \mu^{(m)}[E_i \cap K_m]$$

$$= \sum_{i=1}^{n} \mu[E_i] \qquad \text{for any } n,$$

so $\mu[E] \geq \sum_{i=1}^{\infty} \mu[E_i]$. On the other hand, for $\varepsilon > 0$ fixed, for large enough m

$$\mu[E] \leq \mu^{(m)}[E \cap K_m] + \varepsilon = \sum_{i=1}^{\infty} \mu^{(m)}[E_i \cap K_m] + \varepsilon \leq \sum_{i=1}^{\infty} \mu[E_i] + \varepsilon.$$

Letting $\varepsilon \searrow 0$ we have shown $\mu[E] = \sum_{i=1}^{\infty} \mu[E_i]$.

It remains to show $\mu_{\alpha_{n_k}} \xrightarrow{w} \mu$ as $k \to \infty$. Let $F \subset \mathcal{X}$ be closed, and $\varepsilon > 0$. Then for $m > \frac{1}{\varepsilon}$

$$\limsup_{k\to\infty} \mu_{\alpha_{n_k}}[F] \leq \limsup_{k\to\infty} \mu_{\alpha_{n_k}}[F \cap K_m] + \varepsilon \leq \mu^{(m)}[F] + \varepsilon \leq \mu[F] + \varepsilon$$

Letting $\varepsilon \searrow 0$ we have shown $\mu[F] \geq \limsup_{k\to\infty} \mu_{\alpha_{n_k}}[F]$, which by characterization (iii) of weak convergence implies $\mu_{\alpha_{n_k}} \xrightarrow{w} \mu$. $\square$

Lemma: A collection $(\mu_\alpha)_{\alpha\in\mathcal{A}}$ of probability measures on a complete separable metric space $(\mathfrak{X},\varrho)$ is tight iff for any $\varepsilon>0$ and $r>0$ there exists a finite number of points $x_1,\dots,x_n\in\mathfrak{X}$ s.t.

$$\mu_\alpha\Big[\bigcup_{i=1}^n B_r(x_i)\Big]\geq 1-\varepsilon.$$

Pf: "only if": If $(\mu_\alpha)_{\alpha\in\mathcal{A}}$ is tight, choose $K\subset\mathfrak{X}$ compact s.t. $\mu_\alpha[K]\geq 1-\varepsilon$ $\forall\alpha\in\mathcal{A}$. The collection $\{B_r(x)\mid x\in K\}$ is an open cover of K, and a finite subcover $B_r(x_1),\dots,B_r(x_n)$ has $\mu_\alpha\Big[\bigcup_{i=1}^n B_r(x_i)\Big]\geq\mu_\alpha[K]\geq 1-\varepsilon$.

"if": Assume the condition. Let $\varepsilon>0$.
For any $k\in\mathbb{Z}_{>0}$ there is by assumption a finite collection of balls $B_{1/k}(a_1^{(k)}),\dots,B_{1/k}(a_{n_k}^{(k)})$ s.t. $\mu_\alpha\Big[\bigcup_{j=1}^{n_k}B_{1/k}(a_j^{(k)})\Big]\geq 1-2^{-k}\varepsilon$. $\forall\alpha\in\mathcal{A}$
Then $A=\bigcap_{k=1}^\infty\bigcup_{j=1}^{n_k}B_{1/k}(a_j^{(k)})$ satisfies $\mu_\alpha[A]\geq 1-\varepsilon$ $\forall\alpha$.
Also, A is totally bounded, so $\overline{A}$ is compact, because $\mathfrak{X}$ is complete.

Proof of (P) ⟹ (T) ("converse Prokhorov theorem"): Let $(\mathfrak{X},\varrho)$ be compl., separ.

Assume, by contradiction, that $(\mu_\alpha)_{\alpha\in\mathcal{A}}$ is precompact for weak convergence but not tight. By lemma above, for some $\varepsilon>0$ and $r>0$, any finite collection $x_1,\dots,x_n$ is such that $\mu_\alpha\Big[\bigcup_{j=1}^n B_r(x_j)\Big]\leq 1-\varepsilon$ for some $\alpha\in\mathcal{A}$. Since $\mathfrak{X}$ is separable, choose a dense sequence $x_1,x_2,\dots$ and for any n choose $\alpha_n\in\mathcal{A}$ s.t. $\mu_{\alpha_n}\big[\bigcup_{j=1}^n B_r(x_j)\big]\leq 1-\varepsilon$. By precompactness, some subsequence $(\mu_{\alpha_{n_k}})_{k\in\mathbb{N}}$ converges weakly to some μ. Since $\bigcup_{j=1}^m B_r(x_j)$ is open,

$$\mu\Big[\bigcup_{j=1}^m B_r(x_j)\Big]\leq\liminf_{k\to\infty}\mu_{\alpha_{n_k}}\Big[\bigcup_{j=1}^m B_r(x_j)\Big]\leq 1-\varepsilon.$$ But LHS $\uparrow 1$ as $m\to\infty$.

Contradiction. $\square$

(DONSKERS THEOREM, CONTINUED...)

We saw that to prove tightness of a sequence $(\mu_n)_{n\in\mathbb{N}}$ of probability measures on $C([0,1])$, we need to check that

$$(B): \quad \forall \varepsilon > 0 \quad \exists M > 0, n_0 \geq 0 \quad \text{s.t.} \quad \mu_n\left[\{ f \mid |f(0)| \geq M \}\right] \leq \varepsilon \quad \forall n \geq n_0$$

$$(E.C.): \quad \forall \varepsilon_1, \varepsilon_2 > 0 \quad \exists \delta > 0, n_0 \geq 0 \quad \text{s.t.} \quad \mu_n\left[\{ f \mid w_f(\delta) \geq \varepsilon_1 \}\right] \leq \varepsilon_2 \quad \forall n \geq n_0.$$

Verifying condition (EC) becomes easier with the following:

__Lemma 1:__ Suppose that $0 = t_0 < t_1 < t_2 < \dots < t_k = 1$ are such that $\min_{1\leq j\leq k} (t_j - t_{j-1}) \geq \delta$. Then for any $f \in C([0,1])$ we have

$$w_f(\delta) \leq 3 \max_{1 \leq j \leq k} \max_{s \in [t_{j-1}, t_j]} |f(s) - f(t_{j-1})| = 3M$$

__Proof:__ Denote by M the maximum on RHS.

If both $s,t \in [t_{j-1}, t_j]$ then

$$|f(s) - f(t)| \leq |f(s) - f(t_{j-1})| + |f(t_{j-1}) - f(t)| \leq 2M.$$

If $s \in [t_{j-1}, t_j]$ and $t \in [t_j, t_{j+1}]$ then

$$|f(s) - f(t)| \leq |f(s) - f(t_{j-1})| + |f(t_{j-1}) - f(t_j)| + |f(t_j) - f(t)| \leq 3M.$$

The above cases are the only possibilities if $s,t \in [0,1]$ are such that $s < t$ and $|s-t| \leq \delta$. $\square$

Still, to verify (E.C.) in the context of rescaled random walks

$$(R.W.): \quad X_t^{(a)} = \sqrt{a} \left(\sum_{l=1}^{\lfloor t/a \rfloor} \xi_l + \left(\frac{t}{a} - \left\lfloor \frac{t}{a} \right\rfloor\right) \xi_{\lfloor t/a \rfloor + 1} \right)$$

we use two other lemmas, formulated with slightly different assumptions about the steps $(\xi_l)_{l\in\mathbb{N}}$.

<u>Lemma 2</u>: If $(\xi_\ell)_{\ell\in\mathbb{N}}$ is a stationary sequence, i.e. $(\xi_{L+1},\ldots,\xi_{L+n}) \sim (\xi_1,\ldots,\xi_n)$ $\forall L\in\mathbb{N}$, and $(a_n)_{n\in\mathbb{N}}$ is a sequence of scales $a_n>0$ tending to zero $a_n \searrow 0$, then the laws of $X^{(a_n)}$ defined as in (R.W.) are tight if

$$\lim_{\lambda\to\infty}\ \limsup_{h\to\infty}\ \lambda^2\, \mathbb{P}\Big[\max_{1\le \ell'\le h}\Big|\sum_{\ell=1}^{\ell'}\xi_\ell\Big| \ge \lambda\sqrt{h}\Big] = 0.$$

<u>Proof</u>: Condition (B) holds trivially since $X^{(a_n)}_0 = 0$ $\forall n$.

We only need to verify (E.C.), i.e., that for all $\varepsilon>0$

$$\lim_{\delta\searrow 0}\ \limsup_{n\to\infty}\ \mathbb{P}\big[w_{X^{(a_n)}}(\delta)\ge\varepsilon\big] = 0.$$

By Lemma 1 we can estimate

$$\mathbb{P}\big[w_{X^{(a_n)}}(\delta)\ge 3\varepsilon\big] \le \sum_{j=1}^{k}\mathbb{P}\Big[\max_{s\in[t_{j-1},t_j]}\big|X^{(a_n)}_s - X^{(a_n)}_{t_{j-1}}\big|\ge\varepsilon\Big]$$

when $0=t_0<t_1<\cdots<t_k\ge 1$ are such that $t_j-t_{j-1}\ge\delta$ $\forall j$.

For a given n and δ, we let $h=h_n=\lceil \delta/a_n\rceil$ and $k=\lceil 1/\delta\rceil$ and choose $t_1,\ldots,t_k$ so that $t_j = j\cdot h_n\cdot a_n$.

Note that $h_n\cdot a_n \longrightarrow \delta$ as $n\to\infty$ so for large enough n we have $a_n \le \frac{2\delta}{h_n}$ and $k\le\frac{2}{\delta}$.

The maximum $\max_{s\in[t_{j-1},t_j]}\big|X^{(a_n)}_s - X^{(a_n)}_{t_{j-1}}\big|$ is attained at some $s\in a_n\mathbb{Z}$. The estimate of Lemma 1 then reads

$$\mathbb{P}\big[w_{X^{(a_n)}}(\delta)\ge 3\varepsilon\big] \le \sum_{j=1}^{k}\mathbb{P}\Big[\max_{\substack{\ell'\in\mathbb{Z}\\ jh_n<\ell'\le(j+1)h_n}}\Big|\sqrt{a_n}\sum_{\ell=jh_n+1}^{\ell'}\xi_\ell\Big|\ge\varepsilon\Big]$$

$$= k\cdot\mathbb{P}\Big[\max_{\substack{\ell'\in\mathbb{Z}\\ 0<\ell'\le h_n}}\Big|\sum_{\ell=1}^{\ell'}\xi_\ell\Big|\ge\frac{\varepsilon}{\sqrt{a_n}}\Big] \qquad \leftarrow \text{stationarity (and some rearrangements)}$$

$$\le \frac{2}{\delta}\,\mathbb{P}\Big[\max_{\ell'\le h_n}\Big|\sum_{\ell=1}^{\ell'}\xi_\ell\Big|\ge\frac{\varepsilon\sqrt{h_n}}{\sqrt{2\delta}}\Big] \qquad \text{for large enough } n \text{ (choice of } k, h_n)$$

For a given $\varepsilon>0$ set $\lambda = \varepsilon/\sqrt{2\delta}$ (so $\delta\searrow 0 \Leftrightarrow \lambda\nearrow\infty$)

and rewrite $\mathbb{P}\big[w_{X^{(a_n)}}(\delta)\ge 3\varepsilon\big] \le \frac{4\lambda^2}{\varepsilon^2}\,\mathbb{P}\Big[\max_{1\le\ell'\le h_n}\Big|\sum_{\ell=1}^{\ell'}\xi_\ell\Big|\ge\lambda\sqrt{h_n}\Big]$.

By the assumption, we can make this arbitrarily small for all n large by choosing a sufficiently small δ (large λ). $\square$

We also use the following Etemadi's inequality (very similar to Lévy's inequality used in LLN).
For notational convenience, denote $S_k = \sum_{\ell=1}^{k} \xi_\ell$.

<u>Lemma 3</u> If $(\xi_\ell)_{\ell\in\mathbb{N}}$ are independent, then for all $h, \lambda > 0$ we have

$$\mathbb{P}\Big[\max_{1\le k\le h} |S_k| \geq 3\lambda\Big] \leq 3 \max_{1\le k\le h} \mathbb{P}\big[|S_k| \geq \lambda\big].$$

<u>Proof</u> Define events $A_k = \{ |S_k| \geq 3\lambda \text{ and } |S_{k'}| < 3\lambda \ \forall k' < k \}$.

Then $A_1, \ldots, A_h$ are disjoint and $\bigcup_{k=1}^{h} A_k = \{\max_{1\le k\le h} |S_k| \geq 3\lambda\}$.

Thus
$$\begin{aligned}
\mathbb{P}\Big[\max_{1\le k\le h} |S_k| \geq 3\lambda\Big] &\leq \mathbb{P}\big[|S_h| \geq \lambda\big] + \sum_{k=1}^{h} \mathbb{P}\big[A_k \cap \{|S_h| < \lambda\}\big] \\
&\leq \mathbb{P}\big[|S_h| \geq \lambda\big] + \sum_{k=1}^{h} \underbrace{\mathbb{P}\big[A_k \cap \{|S_h - S_k| > 2\lambda\}\big]}_{= \mathbb{P}[A_k]\, \underbrace{\mathbb{P}[|S_h - S_k| > 2\lambda]}_{\leq \mathbb{P}[|S_h| > \lambda] + \mathbb{P}[|S_k| > \lambda]}} \\
&\leq \mathbb{P}\big[|S_h| \geq \lambda\big] + \Big(\sum_{k=1}^{h} \mathbb{P}[A_k]\Big) \max_{1\le k\le h} \big(\mathbb{P}[|S_h| > \lambda] + \mathbb{P}[|S_k| > \lambda]\big) \\
&\leq 3 \max_{1\le k\le h} \mathbb{P}\big[|S_k| > \lambda\big]. \qquad \square
\end{aligned}$$

Now assume steps $(\xi_\ell)_{\ell\in\mathbb{N}}$ are i.i.d. with $\mathbb{P}[\xi_\ell=+1]=\frac{1}{2}=\mathbb{P}[\xi_\ell=-1]$.

Proof of Donsker's theorem:

Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of scales, $a_n \searrow 0$. We claim that $\bar{X}^{(a_n)}$ converges weakly in $C([0,1])$ to Brownian motion, as $n\longrightarrow\infty$.

The proof strategy is the usual:

1°) show all subsequences have convergent subsubsequences (tightness, which by Prokhorov's thm implies precompactness)

2°) show any subsequential limit is Brownian motion.

Step 1°:

Tightness of laws of $(\bar{X}^{(a_n)})_{n\in\mathbb{N}}$ in $C([0,1])$ is equivalent to (B) and (EC), which we verify using Lemma 2.

Note that since $\mathbb{E}[\xi_\ell]=0$ and $(\xi_\ell)_{\ell\in\mathbb{N}}$ are i.i.d,

$$\mathbb{E}[S_h^4] = \sum_{\ell=1}^{h}\mathbb{E}[\xi_\ell^4] + 6\sum_{1\le \ell<k\le h}\mathbb{E}[\xi_\ell^2]\,\mathbb{E}[\xi_k^2]$$

$$= h\cdot\mathbb{E}[\xi_1^4] + 3\cdot h\cdot(h-1)\,\mathbb{E}[\xi_1^2]^2$$

and with $|\xi_1|=1$ we get $\mathbb{E}[S_h^4]\le 3h^2$.

Then by Lemma 3 and Chebyshev inequality

$$\mathbb{P}\Big[\max_{1\le k\le h}|S_k|\ge\lambda\sqrt{h}\Big]\le 3\max_{1\le k\le h}\mathbb{P}\Big[|S_k|\ge\tfrac{1}{3}\lambda\sqrt{h}\Big]$$

$$\le 3\cdot\max_{1\le k\le h}\frac{3k^2}{(\frac{1}{3}\lambda\sqrt{h})^4} = 3^6\lambda^{-4}.$$

This shows the estimate

$$\lim_{\lambda\to\infty}\limsup_{h\to\infty}\lambda^2\,\mathbb{P}\Big[\max_{1\le k\le h}|S_k|\ge\lambda\sqrt{h}\Big]=0$$

so Lemma 2 implies $(\bar{X}^{(a_n)})_{n\in\mathbb{N}}$ is tight in $C([0,1])$.

Step 2°:

such subsequential limits are known to exist by 1°

finite-dim. distributions do determine a measure on $C([0,1])$

Characterization of subsequential limits is reduced to the central limit theorem (CLT) as follows

Suppose that along some subsequence $(a_{n_k})_{k\in\mathbb{N}}$ of scales we have a limit $X^{(a_{n_k})} \xrightarrow{w} X^*$.

We will show that X^* is Brownian motion by calculating the finite-dimensional distributions.

Note first that we don't need to worry about the piecewise linear interpolation: if

$$\overline{X}^{(a_n)}_t = \sqrt{a_n}\, S_{\lfloor t/a_n \rfloor}, \quad \text{then} \quad |\overline{X}^{(a_n)}_t - X^{(a_n)}_t| \leq \sqrt{a_n} \to 0$$

so the weak limits of the finite-dim. distributions

$$\left(X^{(a_{n_k})}_{t_1}, \dots, X^{(a_{n_k})}_{t_m}\right) \quad \text{and} \quad \left(\overline{X}^{(a_{n_k})}_{t_1}, \dots, \overline{X}^{(a_{n_k})}_{t_m}\right)$$

for any $0 \leq t_1 < t_2 < \dots < t_m \leq 1$, are equal.

Now $\overline{X}^{(a_n)}_{t_j} - \overline{X}^{(a_n)}_{t_{j-1}} = \sqrt{a_n} \sum_{\ell \in \mathbb{Z} \cap (t_{j-1}/a_n, t_j/a_n]} \xi_\ell$

which by CLT tends weakly to $N(0, t_j - t_{j-1})$.

The increments

$$\left(\overline{X}^{(a_{n_k})}_{t_1} - \overline{X}^{(a_{n_k})}_{t_0}, \dots, \overline{X}^{(a_{n_k})}_{t_m} - \overline{X}^{(a_{n_k})}_{t_{m-1}}\right)$$

are independent, and thus tend weakly to a sequence of m independent Gaussians with variances $t_1 - t_0, \dots, t_m - t_{m-1}$.

But knowing the law of these increments is clearly equivalent to knowing the finite dimensional distribution so we see that finite-dim. distributions of the subsequential limit X^* are those of Brownian motion.

Recall that the σ-algebra of $C([0,1])$ generated by finite-dimensional events coincides with the Borel σ-algebra, and finite dimensional events form a π-system (stable under finite intersections) so the finite dimensional distributions determine the distribution of $X^* \in C([0,1])$. $\square$

ISING MODEL

- a model of ferromagnetism (or any other random phenomenon with interactions preferring local alignment)

  Used also e.g. in:
  liquid-gas transition, neural networks,
  image processing, social sciences,
  geology, finance, ...

  NOTE: Comparison with simpler Curie-Weiss model: Ising includes "geometric" aspects (notion of "local")

- has a qualitative phase transition
  (in dimensions $d \geq 2$, in the thermodynamical limit)

  "low temperature" $\longleftrightarrow$ "ferromagnetic"
  "high temperature" $\longleftrightarrow$ "paramagnetic"
  (like for Curie-Weiss)

- has also correct quantitative properties of critical behavior of d-dimensional uniaxial ferromagnets (unlike Curie-Weiss)

ISING MODEL ON A FINITE GRAPH

Fix: finite graph $\mathcal{G}$ "(discretization of) the physical domain"

$\mathcal{G}$ = set of sites
$E(\mathcal{G})$ = set of bonds
a bond = a pair of sites that are "neighbors"

$\beta > 0$ "inverse temperature"

$B \in \mathbb{R}$ "external magnetic field"

- sample space $\Omega = \{-1,+1\}^{\mathcal{G}} \ni \sigma = (\sigma_x)_{x \in \mathcal{G}}$ "spin configuration"

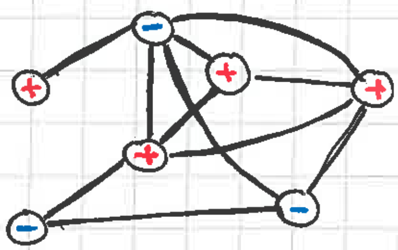
  $\sigma_x \in \{-1,+1\}$ "spin at location x"

- energy $$H(\sigma) = -\sum_{\{x,y\} \in E(\mathcal{G})} \sigma_x \sigma_y - B \cdot \sum_{x \in \mathcal{G}} \sigma_x$$

- probability measure $\mathbb{P} = \mathbb{P}^{(\mathcal{G})}_{\beta,B}$ on $\{-1,+1\}^{\mathcal{G}}$

$$\mathbb{P}[\{\sigma\}] = \frac{1}{Z} e^{-\beta H(\sigma)}$$ where $$Z = \sum_{\sigma \in \{-1,+1\}^{\mathcal{G}}} e^{-\beta H(\sigma)}$$ "partition function"

Any graph $\mathcal{G}$, $E(\mathcal{G})$

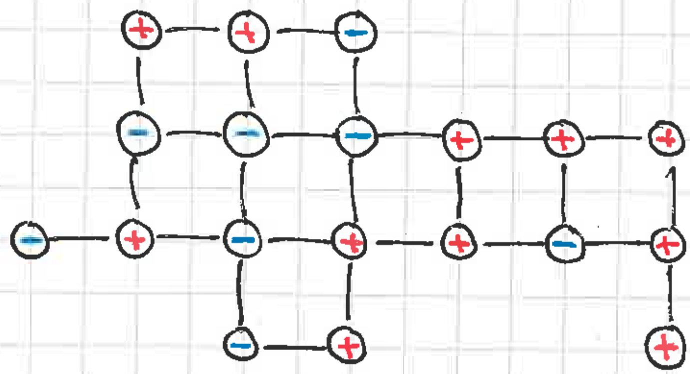


Subgraphs of a lattice

(the most important case)

$$\mathcal{G} \subset \mathbb{Z}^d$$

$$E(\mathcal{G}) = \big\{ \{x,y\} \;\big|\; x,y \in \mathcal{G} \,,\; \|x-y\| = 1 \big\}$$



Interpretation :

- if a larger number of neighbors $\{x,y\}$ have aligned spins, $\sigma_x = \sigma_y$, then lower energy $H(\sigma)$ and higher probability $\mathbb{P}[\{\sigma\}]$
- if a larger number of sites x have σ_x with same sign as B, then lower $H(\sigma)$ and higher $\mathbb{P}[\{\sigma\}]$

Thermodynamical limit

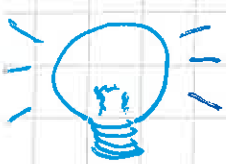
Ising model on infinite graph $\mathbb{Z}^d$ is the physically relevant "large random system". How to define?

Direct definition as above?

Obvious problems: convergence of sums in $H(\sigma)$, infinite Z, zero $\mathbb{P}[\{\sigma\}]$ $\rightsquigarrow$ DOES NOT WORK.

Kolmogorov's extension theorem?

No explicit expression for the (consistent) finite-dimensional marginal distributions $\rightsquigarrow$ DOES NOT WORK.



Approximate $\mathbb{Z}^d$ by increasing finite subgraphs $\mathcal{G}_n \uparrow \mathbb{Z}^d$ and take weak limit of $\mathbb{P}^{(\mathcal{G}_n)}_{\beta,B}$.

$\rightsquigarrow$ WORKS!

ASIDE : DYNAMICAL ISING MODEL

$\mathcal{G}$ finite graph, $\mathbb{P}^{(\mathcal{G})}_{\beta,B}$ as above

$\mathbb{P}^{(\mathcal{G})}_{\beta,B}$ "equilibrium distribution of statistical mechanics"

dynamics "(random) thermal motion"

Glauber dynamics

Continuous time Markov process $(X_t)_{t\geq 0}$ on $\{-1,+1\}^{\mathcal{G}}$ with transitions from σ to σ' at rate

$$\lambda_{\sigma,\sigma'} = \begin{cases} 0 & \text{if } \sigma \text{ and } \sigma' \text{ differ at more than one site} \\ \dfrac{e^{-\beta H(\sigma')}}{e^{-\beta H(\sigma)} + e^{-\beta H(\sigma')}} & \text{if } \#\{x \in \mathcal{G} \mid \sigma_x \neq \sigma'_x\} = 1 \end{cases}$$

(Meaning: $\mathbb{P}[X_{t+\varepsilon} = \sigma' \mid X_t = \sigma] = \varepsilon \cdot \lambda_{\sigma,\sigma'} + o(\varepsilon)$.)

Theorem: The unique stationary distribution of $(X_t)_{t\geq 0}$ is the Ising measure $\mathbb{P}^{(\mathcal{G})}_{\beta,B}$, and we have $X_t \xrightarrow[t\to\infty]{} X_\infty \sim \mathbb{P}^{(\mathcal{G})}_{\beta,B}$ from any initial state X_0

Proof: (Uses theory of Markov processes on finite state space.) Clearly $(X_t)_{t\geq 0}$ is an irreducible continuous time Markov process, so there exists a unique stationary distribution, and the process converges to it as $t\to\infty$. To see that $\mathbb{P}^{(\mathcal{G})}_{\beta,B}$ is the stationary distribution, use reversibility (i.e. detailed balance)

$$\mathbb{P}^{(\mathcal{G})}_{\beta,B}[\{\sigma\}] \cdot \lambda_{\sigma,\sigma'} = \mathbb{P}^{(\mathcal{G})}_{\beta,B}[\{\sigma'\}] \cdot \lambda_{\sigma',\sigma}$$

$\square$

Relevance of dynamics

- thermal motion modeling
- simulation of Ising model

Remark How to sample $\sigma \sim \mathbb{P}^{(\mathcal{G})}_{\beta,B}$?

WARNING: $\#\mathcal{G} \approx 10^4 \Rightarrow \#\Omega \approx 2^{10000}$

$\#\mathcal{G} \approx 10^{23} \Rightarrow \#\Omega \approx 2^{10^{23}}$

COUNTABLE PRODUCT OF DISCRETE SPACES

Ising model on the infinite graph $\mathbb{Z}^d$ will be a probability measure on $\Omega = \{-1,+1\}^{\mathbb{Z}^d}$. Let us take a look at properties of such spaces.

S a finite set e.g. $S = \{-1,+1\}$

I a countable index set e.g. $I = \mathbb{Z}^d$

$$S^I = \{ (\omega_i)_{i\in I} \mid \omega_i \in S \quad \forall i \in I \}$$

To define a metric on $\mathcal{X} = S^I$, choose an enumeration $I = \{i_1, i_2, i_3, \ldots\}$ of the index set, and define

$$\rho(\omega, \omega') = \sum_{j=1}^{\infty} 2^{-j} \mathbb{1}_{\{\omega_j \neq \omega'_j\}} = \sum_{j:\, \omega_j \neq \omega'_j} 2^{-j} .$$

<u>Exercise</u> ρ is a metric on S^I.

Moreover, (S^I, ρ) is complete and separable.

This says that (S^I, ρ) is a nice space. It is in fact very nice:

<u>Proposition</u> (S^I, ρ) is compact.

<u>Proof</u> Let $(\omega^{(n)})_{n\in\mathbb{N}}$ be a sequence, $\omega^{(n)} = (\omega_i^{(n)})_{i\in I} \in S^I$.
The finite set S is compact, so for any $i \in I$ we find subsequences $(\omega^{(n_k)})_{k\in\mathbb{N}}$ s.t. $\omega_i^{(n_k)}$ converges.
Since I is countable, by diagonal extraction we find a subsequence s.t. $\omega_i^{(n_k)}$ converges for all $i \in I$.
A componentwise limit is a limit. $\square$

<u>Remark</u>: More generally, any product of compact spaces is compact by Tychonoff's theorem. However, one then needs to worry about metrizability and whether compactness and sequential compactness are equivalent.

For $s \in S$ and $i \in I$ define

$$C_i(s) = \{\omega \in S^I \mid \omega_i = s\} \subset S^I$$

This is called a <u>simple cylinder set.</u>

More generally, finite intersections

$$C_{i_1}(s_1) \cap C_{i_2}(s_2) \cap \cdots \cap C_{i_n}(s_n) \qquad \begin{matrix} s_1, \dots, s_n \in S \\ i_1, \dots, i_n \in I \end{matrix}$$

are called <u>basic cylinder sets</u>. Finally

$$\{\omega \in S^I \mid \omega_{i_1} \in S_1, \dots, \omega_{i_n} \in S_n\}$$

$S_1, \dots, S_n \subset S$

are called <u>cylinder sets</u>. (These are finite unions of basic cylinders)

(or "cylinder events" or "cylinders")

Lemma 1

(i) There is a countable number of cylinder sets.

(ii) If $s_1, \dots, s_n \in S$ and $i_1, \dots, i_n \in I$, then $\exists r > 0$ s.t. the basic cylinder $C_{i_1}(s_1) \cap \dots \cap C_{i_n}(s_n)$ contains the balls of radius r around each of its points.

(iii) Cylinder sets are open.

(iv) The complement of a cylinder set is a finite union of cylinder sets, and in particular cylinder sets are closed.

(v) Any open set $G \subset S^I$ is a union of cylinder sets.

Corollary The σ-algebra generated by cylinder sets coincides with the Borel σ-algebra.

Proof: Every cylinder set is open by Lemma 1 (iii), so the σ-algebra generated by cylinders is contained in the Borel σ-algebra. On the other hand, any open $G \subset S^I$ is a countable union of cylinder sets by (v) and (i) in Lemma 1, so the σ-algebra generated by cylinder sets contains the Borel σ-algebra $\square$

Proof of Lemma 1   Assume for notational simplicity $I = \mathbb{N}$.

(i) For each $n \in \mathbb{N}$ the number of cylinders of form
$$\{\omega \mid \omega_1 \in S_1, \ \omega_2 \in S_2, \dots, \omega_n \in S_n\}$$
is finite. The number of choices of n is countable (and all cylinders are thus obtained) so the number of cylinders in countable.

(ii) Consider $C_{j_1}(s_1) \cap \dots \cap C_{j_n}(s_n) \ni \omega$.
Let $m = \max\{j_1, \dots, j_n\}$ and $r < 2^{-m}$.
Then for any $\omega' \in B_r(\omega)$ we have $\omega'_{j_1} = \omega_{j_1} = s_1$, $\dots$, $\omega'_{j_n} = \omega_{j_n} = s_n$ since otherwise
$$\rho(\omega, \omega') = \sum_{j=1}^{\infty} \mathbb{1}_{\{\omega_j \neq \omega'_j\}} 2^{-j} \quad \text{is at least } 2^{-m}.$$

(iii) Basic cylinders are open by (ii) and any cylinder is a union of basic cylinders, thus open.

(iv)
$$\mathcal{X} \setminus \{\omega \mid \omega_{i_1} \in S_1, \dots, \omega_{i_n} \in S_n\} = \bigcup_{k=1}^{n} \{\omega \mid \omega_{i_k} \in S \setminus S_k\}$$

(v) Let $G \subset \mathcal{X}$ be open and $\omega \in G$.
Then $\exists r_\omega > 0$ s.t. $B_{r_\omega}(\omega) \subset G$. If $r_\omega > 2^{-m}$
then $\{\tilde{\omega} \mid \tilde{\omega}_1 = \omega_1, \tilde{\omega}_2 = \omega_2, \dots, \tilde{\omega}_{m+1} = \omega_{m+1}\} \subset B_{r_\omega}(\omega) \subset G \subset \mathcal{X}$.
The open set G is the union of such cylinders ("one for each $\tilde{\omega} \in G$", in principle). $\square$

A criterion for weak convergence

Let $(\mathcal{X}, \varrho)$ be a metric space, and $\mathcal{B} = \mathcal{B}(\mathcal{X})$ its Borel σ-algebra.

Proposition Suppose that $\mathcal{E} \subset \mathcal{B}$ is a collection s.t.

(FIP) $\mathcal{E}$ is stable under finite intersections: if $E_1, E_2 \in \mathcal{E}$ then $E_1 \cap E_2 \in \mathcal{E}$.

(OCU) Any open set $G \subset \mathcal{X}$ is a countable union of sets from $\mathcal{E}$: $G = \bigcup_{i \in \mathbb{N}} E_i$ where $E_i \in \mathcal{E}$.

Then, a sequence $(\nu_n)_{n \in \mathbb{N}}$ of probability measures on $\mathcal{X}$ converges weakly to a proba meas. ν if for all $E \in \mathcal{E}$ we have $\nu_n[E] \longrightarrow \nu[E]$.

Proof: If $E_1, \ldots, E_m \in \mathcal{E}$ then by (FIP) the intersections of these are also in $\mathcal{E}$.

By inclusion-exclusion formula

$$\nu_n\Big[\bigcup_{i=1}^m E_i\Big] = \sum_{\substack{J \subset \{1,\ldots,m\} \\ J \neq \emptyset}} (-1)^{\#J-1} \, \nu_n\Big[\underbrace{\bigcap_{j \in J} E_j}_{\in \mathcal{E} \text{ by (FIP)}}\Big]$$

$$\xrightarrow{\text{assumption}} \sum_{\substack{J \subset \{1,\ldots,m\} \\ J \neq \emptyset}} (-1)^{\#J-1} \, \nu\Big[\bigcap_{j \in J} E_j\Big] = \nu\Big[\bigcup_{i=1}^m E_i\Big].$$

If $G \subset \mathcal{X}$ is open, then by (OCU) $G = \bigcup_{i=1}^{\infty} E_i$ for some $E_i \in \mathcal{E}$. Then the above gives, for any m,

$$\nu\Big[\bigcup_{i=1}^m E_i\Big] = \lim_{n \to \infty} \nu_n\Big[\bigcup_{i=1}^m E_i\Big] \leq \liminf_{n \to \infty} \nu_n[G].$$

On the other hand, $\bigcup_{i=1}^m E_i \uparrow G$ as $m \to \infty$, so the LHS tends to $\nu[G]$ by monotone approximation of measures. We conclude condition (ii) of Portmanteau theorem, and thus $\nu_n \xrightarrow{w} \nu$. $\square$

Weak convergence with cylinders

Let S be finite and I countable, and consider the space (S^I, ϱ) as above.

Theorem A sequence of probability measures $(\nu_n)_{n\in\mathbb{N}}$ on S^I converges weakly if and only if for every cylinder set C the limit $\lim_{n\to\infty} \nu[C]$ exists.

Proof: "only if": Suppose $\nu_n \xrightarrow{w} \nu$, and let C be a cylinder set. By Lemma 1, (iii)&(iv) C is both open and closed, and therefore $C^\circ = C = \overline{C}$ and $\partial C = \phi$. Thus by condition (iv) of Portmanteau theorem $\nu_n[C] \longrightarrow \nu[C]$ as $n \to \infty$.

"if": The collection of cylinder sets is stable under finite intersections, and any open set is a countable union of cylinder sets by Lemma 1, (i) & (v). By Proposition above, it is therefore sufficient for weak convergence that $\nu_n[C] \longrightarrow \nu[C]$ for all cylinders C, where ν is a probability measure on S^I. We assume $\alpha[C] = \lim_{n\to\infty} \nu_n[C]$ exists for all cylinders C, we only need to show that α is a probability measure. Recall that S^I is compact, and therefore $(\nu_n)_{n\in\mathbb{N}}$ is automatically tight. By Prokhorov's theorem there exists a subsequence $(\nu_{n_k})_{k\in\mathbb{N}}$ s.t. $\nu_{n_k} \xrightarrow{w} \nu$ as $k \to \infty$. But clearly $\nu[C] = \alpha[C]$ (by the same argument as in "only if"). This shows α is a probability measure. $\square$

ISING MODEL, CONTINUED

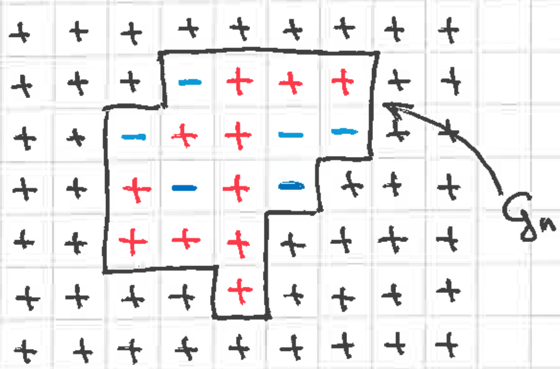
THERMODYNAMICAL LIMIT OF ISING MODEL

For definiteness, we consider the infinite volume limit (thermodynamical limit) of the Ising model with plus boundary conditions. (Imagine spins outside the domain to be frozen as $+1$)

Fix:
- dimension $d \in \mathbb{N}$, inverse temp. $\beta > 0$, ext. magn. field $B \in \mathbb{R}$.
- an increasing sequence $(G_n)_{n \in \mathbb{N}}$ of finite subgraphs $G_n \subset \mathbb{Z}^d$ such that $\bigcup_{n \in \mathbb{N}} G_n = \mathbb{Z}^d$.

A spin configuration $\sigma^{(n)} = (\sigma^{(n)}_x)_{x \in G_n} \in \{-1,+1\}^{G_n}$ is extended to a spin configuration $\sigma = (\sigma_x)_{x \in \mathbb{Z}^d} \in \{-1,+1\}^{\mathbb{Z}^d} = \mathcal{X}$ by setting

$$\sigma_x = \begin{cases} \sigma^{(n)}_x & \text{if } x \in G_n \\ +1 & \text{if } x \in \mathbb{Z}^d \setminus G_n . \end{cases}$$



We will interpret Ising model measures $\mathbb{P}^+_{G_n}$ on $\{-1,+1\}^{G_n}$ for all $n \in \mathbb{N}$ as measures on

$$\mathcal{X} = \{-1,+1\}^{\mathbb{Z}^d}$$

and study weak limit $\mathbb{P}^+_{G_n} \xrightarrow{w} \mathbb{P}^+_{\mathbb{Z}^d}$.

Recall: A sufficient condition for weak convergence is that all cylinder probabilities $(x_1, \ldots, x_n \in \mathbb{Z}^d, \quad \epsilon_1, \ldots, \epsilon_n \in \{-1,+1\}.)$

$$\mathbb{P}^+_{G_n}[\sigma_{x_1} = \epsilon_1, \ldots, \sigma_{x_n} = \epsilon_n]$$ have limits as $n \to \infty$.

Further simplification: We may assume $\epsilon_1 = \cdots = \epsilon_n = +1$.

The general case can be obtained by inclusion-exclusion:

$$\mathbb{P}[\sigma_x = +1 \ \forall x \in P, \sigma_y = -1 \ \forall y \in M] = \mathbb{P}[\sigma_x = +1 \ \forall x \in P] - \mathbb{P}[\sigma_x = +1 \ \forall x \in P, \exists y \in M : \sigma_y = +1]$$

$$= \sum_{J \subset M} (-1)^{\# J} \, \mathbb{P}[\sigma_z = +1 \ \forall z \in P \cup J].$$

Here the Ising model on G_n is taken with plus boundary condition, meaning that the measure $\mathbb{P}^+_{G_n}$ is

$$\mathbb{P}^+_{G_n}[\{\sigma\}] = \frac{1}{Z^+_{G_n}} e^{-\beta H^+_{G_n}(\sigma)} \qquad \forall \sigma \in \{-1,+1\}^{G_n}$$

where

$$H^+_{G_n}(\sigma) = -\sum_{\substack{\{x,y\}: \\ \|x-y\|=1,\ \{x,y\}\cap G_n \neq \phi.}} \sigma_x \sigma_y - B \sum_{x \in G_n} \sigma_x$$

(Note: for $x \notin G_n$ σ_x is interpreted as $+1$)

and

$$Z^+_{G_n} = \sum_{\sigma \in \{-1,+1\}^{G_n}} e^{-\beta H^+_{G_n}(\sigma)}$$

normalizes $\mathbb{P}^+_{G_n}[\{-1,+1\}^{G_n}] = 1$.

Lemma: Let $n > 1$. Then the conditional law of $\sigma \sim \mathbb{P}^+_{G_n}$ given the event $\sigma|_{G_n \setminus G_{n-1}} \equiv +1$ is $\mathbb{P}^+_{G_{n-1}}$.

Proof: Let $\tau \in \{-1,+1\}^{G_{n-1}}$ (extended as $+1$ to $\mathbb{Z}^d \setminus G_{n-1}$). Then

$$H^+_{G_n}(\tau) = -\sum_{\substack{\{x,y\} \\ \|x-y\|=1,\ \{x,y\}\cap G_n \neq \phi}} \tau_x \tau_y - B \sum_{x \in G_n} \tau_x$$

$$= \Bigg(-\sum_{\substack{\{x,y\} \\ \|x-y\|=1,\ \{x,y\}\cap G_{n-1}}} \tau_x \tau_y - B \sum_{x \in G_{n-1}} \tau_x \Bigg) \underbrace{- \sum_{\substack{\{x,y\} \\ \|x-y\|=1, \\ x,y \notin G_{n-1} \\ \{x,y\}\cap G_n \neq \phi}} (+1)(+1) - B \sum_{x \in G_n \setminus G_{n-1}} (+1)}_{\text{constant } c}$$

$$= H^+_{G_{n-1}}(\tau) + c$$

Therefore, we have

$$\mathbb{P}^+_{G_n}\left[\sigma = \tau \,\middle|\, \sigma|_{G_n \setminus G_{n-1}} \equiv +1\right] = \frac{e^{-\beta H^+_{G_n}(\tau)}}{\mathbb{P}^+_{G_n}[\sigma|_{G_n \setminus G_{n-1}} \equiv +1]} = C \cdot e^{-\beta H^+_{G_{n-1}}(\tau)}$$

where $C = e^{-\beta c} / \mathbb{P}^+_{G_n}[\sigma|_{G_n \setminus G_{n-1}} \equiv +1]$ is a constant.

But also

$$\mathbb{P}^+_{G_{n-1}}[\sigma = \tau] = \frac{1}{Z^+_{G_{n-1}}} e^{-\beta H^+_{G_{n-1}}(\tau)}.$$

We must have $C = \frac{1}{Z^+_{G_{n-1}}}$ since both are probability measures on $\{-1,+1\}^{G_{n-1}}$ $\square$.

We want to prove:

Theorem (Existence of infinite volume Ising model)

The probability measures $\mathbb{P}^+_{G_n}$ (of Ising model with plus boundary conditions on G_n) converge weakly to a probability measure $\mathbb{P}^+_{\mathbb{Z}^d}$ on $\mathcal{X} = \{-1,+1\}^{\mathbb{Z}^d}$. This limit does not depend on the choice of the increasing sequence $(G_n)_{n\in\mathbb{N}}$ of finite subgraphs $G_n \subset \mathbb{Z}^d$ s.t. $\bigcup_{n\in\mathbb{N}} G_n = \mathbb{Z}^d$.

To prove the existence of infinite volume Ising model, we use:

Proposition (Positive association)

Let $A, B \subset G_n$. Then we have

$$\mathbb{P}^+_{G_n}[\sigma_{|A} \equiv +1 \mid \sigma_{|B} \equiv +1] \geq \mathbb{P}^+_{G_n}[\sigma_{|A} = +1]$$

Interpret: Conditioning some spins to be plus increases the probability of other spins to be plus.
Very natural for a ferromagnet!

Proof of existence of infinite volume limit:

By earlier observations, it is sufficient to show that for all finite $A \subset \mathbb{Z}^d$ the limit $\lim_{n\to\infty} \mathbb{P}^+_{G_n}[\sigma_{|A} \equiv +1] = \alpha[A]$ exists. We show that the sequence $(\mathbb{P}^+_{G_n}[\sigma_{|A} \equiv +1])_{n\in\mathbb{N}}$ is decreasing, and therefore must have a limit.

Indeed, by Lemma above and positive association we have:

$$\mathbb{P}^+_{G_{n-1}}[\sigma_{|A} \equiv +1] = \mathbb{P}^+_{G_n}[\sigma_{|A} \equiv +1 \mid \sigma_{|G_n \setminus G_{n-1}} \equiv +1] \geq \mathbb{P}^+_{G_n}[\sigma_{|A} \equiv +1]$$

To show the independence of the limit of the choice of (G_n), let (G_n') be another sequence.
Fix $\varepsilon > 0$. For n large enough we have $\mathbb{P}^+_{G_n}[\sigma_{|A} \equiv +1] \leq \alpha[A] + \varepsilon$.
But for large enough m, $G_n \subset G_m'$ so $\mathbb{P}^+_{G_m'}[\sigma_{|A} \equiv +1] \leq \mathbb{P}^+_{G_n}[\sigma_{|A} \equiv +1]$.
We get $\lim_{m\to\infty} \mathbb{P}^+_{G_m'}[\sigma_{|A} \equiv +1] \leq \lim_{n\to\infty} \mathbb{P}^+_{G_n}[\sigma_{|A} \equiv +1]$. Also vice versa. $\square$

So we need to prove positive association, and for this we use the more general result.

<u>Theorem</u> (FKG inequality for Ising model)

Suppose that $f, g : \{-1,+1\}^{\mathcal{G}} \longrightarrow \mathbb{R}$ are increasing functions. Then we have

$$\mathbb{E}_{\mathcal{G}}^{+}[f(\sigma)g(\sigma)] \geq \mathbb{E}_{\mathcal{G}}^{+}[f(\sigma)] \cdot \mathbb{E}_{\mathcal{G}}^{+}[g(\sigma)].$$

<u>Recall</u>: $\{-1,+1\}^{\mathcal{G}}$ is a partially ordered set: $\sigma \leq \sigma'$ means $\forall x \in \mathcal{G}: \sigma_x \leq \sigma'_x$. A function $f : \{-1,+1\}^{\mathcal{G}} \to \mathbb{R}$ is increasing if $f(\sigma) \leq f(\sigma')$ $\forall \sigma \leq \sigma'$.

<u>Proof of positive association</u>:

Let $A, B \subset \mathcal{G}$. Define f, g as indicators

$$f = \mathbb{1}_{\{\sigma|_A \equiv +1\}}, \quad g = \mathbb{1}_{\{\sigma|_B \equiv +1\}}.$$

These are clearly increasing functions, and

$$\mathbb{E}_{\mathcal{G}}^{+}[f] = \mathbb{P}_{\mathcal{G}}^{+}[\sigma|_A \equiv +1]$$

$$\frac{\mathbb{E}_{\mathcal{G}}^{+}[f \cdot g]}{\mathbb{E}_{\mathcal{G}}^{+}[g]} = \frac{\mathbb{P}_{\mathcal{G}}^{+}[\sigma|_A \equiv +1 \text{ and } \sigma|_B \equiv +1]}{\mathbb{P}_{\mathcal{G}}^{+}[\sigma|_B \equiv +1]} = \mathbb{P}_{\mathcal{G}}^{+}[\sigma|_A \equiv +1 \mid \sigma|_B \equiv +1].$$

By FKG inequality $\mathbb{E}[f \cdot g]/\mathbb{E}[g] \geq \mathbb{E}[\;]$ so we get

$\mathbb{P}_{\mathcal{G}}^{+}[\sigma|_A \equiv +1 \mid \sigma|_B \equiv +1] \geq \mathbb{P}_{\mathcal{G}}^{+}[\sigma|_A \equiv +1]$, which proves positive association. $\square$

So we need to prove FKG inequality. To this end, we use the following more general result, Holley's criterion, about monotonicity / domination in a stochastic setting.

To state the result we some definitions and observations.

Definitions: Let Ω be a partially ordered set with partial order $\preceq$, and let $\mathcal{F}$ be a σ-algebra on Ω.

A function $f: \Omega \to \mathbb{R}$ is said to be <u>increasing</u> if for any $\omega, \omega' \in \Omega$ s.t. $\omega \preceq \omega'$ we have $f(\omega) \leq f(\omega')$.

If $\nu, \tilde{\nu}$ are two probability measures on Ω (same σ-algebra $\mathcal{F}$), then $\tilde{\nu}$ is said to <u>stochastically dominate</u> ν, denoted $\nu \preceq \tilde{\nu}$, if for all increasing $f: \Omega \to \mathbb{R}$ we have

$$\int_\Omega f(\omega)\, d\nu(\omega) \leq \int_\Omega f(\omega)\, d\tilde{\nu}(\omega).$$

Observation: If there exists a coupling μ of ν and $\tilde{\nu}$ (i.e. μ proba meas. on $\Omega \times \Omega$ s.t. $\mu[E \times \Omega] = \nu[E]$, $\mu[\Omega \times E] = \tilde{\nu}[E]$) such that for $\mu[\{(\omega,\omega') \in \Omega \times \Omega \mid \omega \preceq \omega'\}] = 1$, then $\nu \preceq \tilde{\nu}$. Indeed, for any increasing $f: \Omega \to \mathbb{R}$

$$\int_\Omega f\, d\nu = \int_{\Omega\times\Omega} f(\omega)\, d\mu(\omega,\omega') \leq \int_{\Omega\times\Omega} f(\omega')\, d\mu(\omega,\omega') = \int_\Omega f\, d\tilde{\nu}.$$

$\omega \preceq \omega'$ μ-a.s. $\Rightarrow f(\omega) \leq f(\omega')$

Exercise Suppose F_1, F_2 are two cumulative distribution functions such that $F_1(x) \geq F_2(x)$ $\forall x \in \mathbb{R}$. Let ν_1 and ν_2 be the corresponding probability measures on $\mathbb{R}$. Show $\nu_1 \preceq \nu_2$.

Theorem (Holley's criterion) Let G be a finite set, and $\nu, \tilde{\nu}$ two probability measures on $\{-1,+1\}^G$ such that $\nu[\{\sigma\}] > 0$ and $\tilde{\nu}[\{\tau\}] > 0$ for all $\sigma, \tau \in \{-1,+1\}^G$. Assume that whenever $\sigma \preceq \tau$ and $z \in G$, we have

$$\frac{\tilde{\nu}[\{\tau^{z\to+1}\}]}{\tilde{\nu}[\{\tau^{z\to-1}\}]} \geq \frac{\nu[\{\sigma^{z\to+1}\}]}{\nu[\{\sigma^{z\to-1}\}]}.$$

Then there exists a coupling μ of ν and $\tilde{\nu}$ such that $\mu[\{(\sigma,\tau) \mid \sigma \preceq \tau\}] = 1$. In particular we have $\nu \preceq \tilde{\nu}$.

Notation: For $\sigma \in \{-1,+1\}^G$, $z \in G$ denote by $\sigma^{z\to+1}, \sigma^{z\to-1}$ the configurations obtained by redefining the spin at z:

$$\sigma^{z\to\pm1}_x = \begin{cases} \sigma_x & \text{if } x \neq z \\ \pm 1 & \text{if } x = z \end{cases}$$

Proof of FKG inequality. Denote the Ising proba measure by $\nu = \mathbb{P}^+_{\mathcal{G}}$.

Let f, g be two increasing functions $\{-1,+1\}^{\mathcal{G}} \to \mathbb{R}$.
We may assume without loss of generality that $g > 0$ and $\int g \, d\nu = 1$, since adding a constant to g and multiplying g by positive constant does not affect the inequality we want to prove.

Define another measure $\tilde{\nu}$ on $\{-1,+1\}^{\mathcal{G}}$ by

$$\tilde{\nu}[\{\sigma\}] = g(\sigma) \cdot \nu[\{\sigma\}],$$

in which case $\int f \, d\tilde{\nu} = \int f \cdot g \, d\nu$, and we seek to prove that for all increasing f we have $\int f \, d\tilde{\nu} \geq \int f \, d\nu$.

The properties $g > 0$ and $\int g \, d\nu = 1$ imply that $\tilde{\nu}$ is also a probability measure. We seek to prove $\tilde{\nu} \geq \nu$.
To achieve this, we verify Holley's criterion.
Let $\sigma \preceq \tau$. We calculate the ratio appearing in the criterion directly by the definitions of measures $\nu, \tilde{\nu}$:

$$\frac{\tilde{\nu}[\{\tau^{z \to +1}\}] \cdot \nu[\{\sigma^{z \to -1}\}]}{\tilde{\nu}[\{\tau^{z \to -1}\}] \cdot \nu[\{\sigma^{z \to +1}\}]} = \frac{g(\tau^{z \to +1})}{g(\tau^{z \to -1})} \cdot \frac{\nu[\{\tau^{z \to +1}\}]}{\nu[\{\tau^{z \to -1}\}]} \, \frac{\nu[\{\sigma^{z \to -1}\}]}{\nu[\{\sigma^{z \to +1}\}]}$$

$$= \underbrace{\frac{g(\tau^{z \to +1})}{g(\tau^{z \to -1})}}_{\substack{\geq 1 \text{ since } g \\ \text{is increasing}}} \cdot \underbrace{\exp\Big(2\beta \cdot \sum_{x \,:\, \|x - z\| = 1} (\tau_x - \sigma_x)\Big)}_{\geq 1 \text{ since } \sigma \preceq \tau} \geq 1.$$

This verifies Holley's criterion and we get that $\nu \preceq \tilde{\nu}$. The FKG inequality follows:

$$\int f \, d\nu \leq \int f \, d\tilde{\nu} \quad \longleftarrow \text{ since } \nu \preceq \tilde{\nu} \text{ and } f \text{ increasing}$$

$$\int f \, d\nu = \mathbb{E}^+_{\mathcal{G}}[f] = \mathbb{E}^+_{\mathcal{G}}[f] \cdot \underbrace{\mathbb{E}^+_{\mathcal{G}}[g]}_{=1}$$

$$\int f \, d\tilde{\nu} = \int f \cdot g \, d\nu = \mathbb{E}^+_{\mathcal{G}}[f \cdot g] \qquad \square$$

We still need to prove Holley's criterion.
Let's start with a simpler case.

__Lemma__ Let $\mathcal{G}$ be a finite set, and ν a probability measure on $\{-1,+1\}^{\mathcal{G}}$ s.t. $\nu[\{\sigma\}] > 0 \quad \forall \sigma \in \{-1,+1\}^{\mathcal{G}}$.
Define a continuous time Markov process $(X_t)_{t\geq 0}$ on $\{-1,+1\}^{\mathcal{G}}$ by transition rates

for all $z \in \mathcal{G}$ and $\sigma \in \{-1,+1\}^{\mathcal{G}}$

$$\begin{cases} \lambda(\sigma^{z\to+1}, \sigma^{z\to-1}) = 1 \\ \lambda(\sigma^{z\to-1}, \sigma^{z\to+1}) = \dfrac{\nu[\{\sigma^{z\to+1}\}]}{\nu[\{\sigma^{z\to-1}\}]} \end{cases}$$

$$\lambda(\sigma,\tau) = 0 \quad \text{if } \#\{x \in \mathcal{G} \mid \sigma_x \neq \tau_x\} > 1.$$

- spins flip down at unit rate
- spins flip up at a rate depending on ν
- only one spin can flip at a time

Then ν is stationary for the process $(X_t)_{t\geq 0}$ and $X_t \longrightarrow X_\infty \sim \nu$ as $t \to \infty$.

__Proof__: The process is irreducible on $\{-1,+1\}^{\mathcal{G}}$ (can flip all spins one at a time to get to any configuration).
The measure ν satisfies reversibility (detailed balance)

$$\nu[\{\sigma\}] \cdot \lambda(\sigma,\tau) = \nu[\{\tau\}] \cdot \lambda(\tau,\sigma) \qquad \forall \sigma,\tau \in \{-1,+1\}^{\mathcal{G}}.$$

Thus ν is the unique stationary distribution and $X_t \longrightarrow X_\infty \sim \nu$. $\square$

__Proof of Holley's criterion:__

We define a continuous time Markov process with two components, $(X_t, \tilde{X}_t)_{t\geq 0}$, taking values on the subset $S = \{(\sigma,\tau) \mid \sigma \preceq \tau\} \subset \{-1,+1\}^{\mathcal{G}} \times \{-1,+1\}^{\mathcal{G}}$ by transition rates

$$\lambda\big((\sigma, \tau^{z\to+1}), (\sigma^{z\to-1}, \tau^{z\to-1})\big) = 1$$

for all $z \in \mathcal{G}$ and $(\sigma,\tau) \in S$

spins of second component flip down at unit rate, and whenever this happens, the same spin in the first component flips down (if it wasn't down already)

$$\lambda\big((\sigma^{z\to-1}, \tau), (\sigma^{z\to+1}, \tau^{z\to+1})\big) = \frac{\nu[\{\sigma^{z\to+1}\}]}{\nu[\{\sigma^{z\to-1}\}]}$$

spins of the first component flip up at a rate given by ν, and whenever this happens, the same spin in the second component flips up (if it wasn't up already)

$$\lambda\left((\sigma^{z\to-1},\tau^{z\to-1}),(\sigma^{z\to-1},\tau^{z\to+1})\right) = \frac{\tilde\nu[\{\tau^{z\to+1}\}]}{\tilde\nu[\{\tau^{z\to-1}\}]} - \frac{\nu[\{\sigma^{z\to+1}\}]}{\nu[\{\sigma^{z\to-1}\}]} \qquad (\star)$$

spins of the second component can flip up while the same spin remains down in the first component

$\lambda\left((\sigma,\tau),(\sigma',\tau')\right) = 0$ for all other pairs $(\sigma,\tau),(\sigma',\tau')\in S$.

Note that all the rates are non-negative — the non-negativity of $(\star)$ is guaranteed by the condition in Holley's criterion about ν and $\tilde\nu$. Therefore the rates indeed define a process $(X_t,\tilde X_t)_{t\geq 0}$.

The process is irreducible on the subset S.
(we can first flip everything down, then flip both components up at all $x\in G$ for which $\sigma_x=+1$ and then flip only the second component up at all $x\in G$ for which we have $\tau_x=+1$ but $\sigma_x=-1$.)

Therefore $(X_t,\tilde X_t)\longrightarrow(X_\infty,\tilde X_\infty)\sim\mu$ as $t\to\infty$, where μ is the unique stationary distribution of the process on $S\subset\{-1,+1\}^G\times\{-1,+1\}^G$.

On the other hand, the first component $(X_t)_{t\geq 0}$ alone is a Markov process on $\{-1,+1\}^G$, whose transition rates are exactly those in Lemma above. Therefore $X_t\longrightarrow X_\infty\sim\nu$.

Finally, the second component $(\tilde X_t)_{t\geq 0}$ is also a Markov process, with transition rates as in Lemma above except ν is replaced by $\tilde\nu$.
Therefore $\tilde X_t\longrightarrow\tilde X_\infty\sim\tilde\nu$.

We get that the marginals of μ on the first and second component are ν and $\tilde\nu$, i.e., μ is a coupling of ν and $\tilde\nu$.
Moreover μ is supported on $S=\{(\sigma,\tau)\mid\sigma\leq\tau\}$.
In particular if $f:\{-1,+1\}^G\to\mathbb{R}$ is increasing, then

$$\int f\,d\nu = \int f(\sigma)\,d\mu(\sigma,\tau) \leq \int f(\tau)\,d\mu(\sigma,\tau) = \int f\,d\tilde\nu. \qquad \square$$

$\sigma\leq\tau$ μ-a.s. $\Rightarrow$ $f(\sigma)\leq f(\tau)$ μ-a.s.

PHASE TRANSITION IN THE ISING MODEL

We will formulate precisely a result, by which the Ising model has a phase transition if the spatial dimension d is at least two. The proof is omitted in this course. The result should be compared with the Curie-Weiss model. Note in particular that again, seeing the qualitative phase transition requires that we take the limit of a large system (here: graph tending to the infinite lattice $\mathbb{Z}^d$.)

<u>Theorem</u> Let $d \in \mathbb{N}$ and $B = 0$, and let $\mathbb{P}^{+;\beta}_{\mathbb{Z}^d}$ denote the weak limit of Ising probability measures with plus-boundary conditions, at inverse temperature $\beta > 0$, on finite subgraphs $\mathcal{G}_n \subset \mathbb{Z}^d$ increasing to $\mathbb{Z}^d$ as $n \longrightarrow \infty$. Then if $d \geq 2$, there exists a critical value $\beta_c = \beta_c(d) > 0$ such that for any $x \in \mathbb{Z}^d$

$$\mathbb{E}^{+;\beta}_{\mathbb{Z}^d}[\sigma_x] = 0 \qquad \text{if} \quad \beta > \beta_c$$

$$\mathbb{E}^{+;\beta}_{\mathbb{Z}^d}[\sigma_x] > 0 \qquad \text{if} \quad \beta > \beta_c.$$

"spontaneous magnetization"

If $d = 1$, then for all β we have $\mathbb{E}^{+;\beta}_{\mathbb{Z}}[\sigma_x] = 0$.

INTERACTING PARTICLE SYSTEMS

The term "interacting particle system" refers to a class of continuous time Markov processes, which share some common features, such as

- state space $S^{\mathbb{Z}^d}$, where S is a finite set
- flip rates of components are translation invariant and depend only on finitely many other components

We will give precise definitions and examples later, after first having reviewed the much simpler case of continuous time Markov processes on finite state space.

REVIEW OF POISSONIAN JUMP PROCESSES ON FINITE STATE SPACE

S finite set, "state space"

Transition rates ("jump rates")

$$\lambda(\eta,\eta') \geq 0 \qquad \eta,\eta' \in S \qquad \eta \neq \eta'$$

given

we should construct

Process $X = (X_t)_{t\geq 0}$ s.t.

$$X_t \in S \qquad \forall t \geq 0$$

$$\mathbb{P}[X_{t+s} = \eta' \mid X_t = \eta] = \lambda(\eta,\eta') \cdot s + o(s)$$

More specifically, for any initial state $\eta_0 \in S$ we want a proba meas. $\mathbb{P}_{\eta_0}$ on such $X = (X_t)_{t\geq 0}$ with the property that $\mathbb{P}_{\eta_0}[X_0 = \eta_0] = 1$. If μ is a proba meas. on S, then denote $\mathbb{P}_\mu = \sum_{\eta_0 \in S} \mu[\{\eta_0\}]\, \mathbb{P}_{\eta_0}$.

Markov property: denote by Θ_s $(s \geq 0)$ the time shift operation on processes $X = (X_t)_{t\geq 0}$ def. by $(\Theta_s X)_t = X_{s+t}$.

The Markov property states that the law of $\Theta_s X$ is $\mathbb{P}_{\mu_s}$, where μ_s is the law of X_s.

Poisson process

Let $\lambda \geq 0$.

Take $(\tau_n)_{n\in\mathbb{Z}_{>0}}$ i.i.d. $\tau_n \sim \mathrm{Exp}(\lambda)$. "waiting times"

Define $(T_n)_{n\in\mathbb{Z}_{\geq 0}}$ by $T_n = \sum_{k=1}^{n} \tau_k$, "arrival times"

and $P = (P_t)_{t\geq 0}$ by $P_t = \sup\{n\in\mathbb{Z}_{\geq 0} \mid T_n \leq t\}$.

This is the Poisson process of intensity λ. (Below we mainly use arrival times $(T_n)_{n\in\mathbb{N}}$.)

Exercise: If, independently of P, each arrival is kept independently with probability p, then the kept arrivals form a Poisson process of intensity λp.

"Thinning of Poisson process"

Remark: $\mathbb{P}[P_{t+s} = n+1 \mid P_t = n] = \lambda s + o(s)$.

"jump rate from n to $n+1$ is λ".

Construction of jump processes:

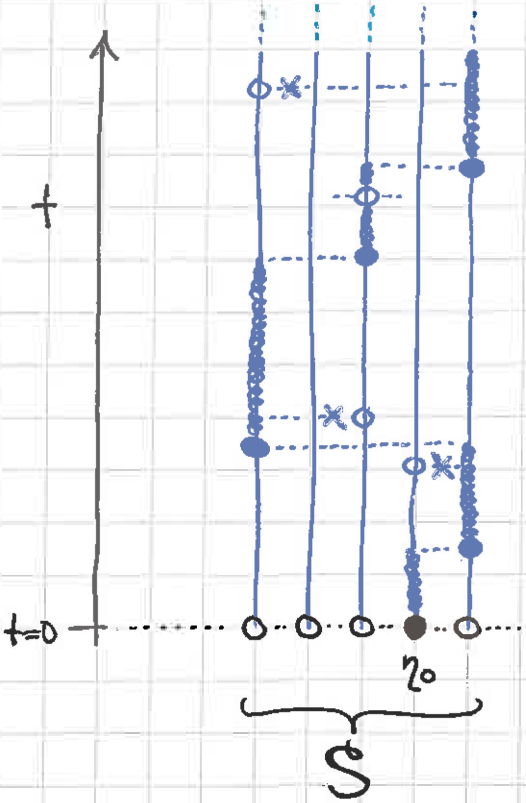
We may assume (and do) that $\lambda(\eta,\eta') \leq 1 \quad \forall \eta,\eta'\in S \quad \eta\neq\eta'$.

(Do a linear time change if necessary)

To construct a process with given rates, we take for each $\eta\in S$ an independent unit rate Poisson process $(T_n^\eta)_{n\in\mathbb{N}}$ and independent i.i.d. sequence $(U_n^{(\eta)})_{n\in\mathbb{N}}$ of uniform r.v.'s $U_n^\eta \sim \mathrm{Unif}([0,1])$.

We define $X = (X_t)_{t\geq 0}$ by setting $X_0 = \eta_0$ and prescribing that there is a jump at time $t = T_n^\eta$ to state η if $U_n^\eta \leq \lambda(X_{t-},\eta)$. Since there are (almost surely) only finitely many arrival times T_n^η, $n\in\mathbb{N}, \eta\in S$ in any bounded time interval, the prescription gives rise to a well-defined right-continuous random function $t \longmapsto X_t \in S$.

This is best illustrated graphically:



$\circ$: arrival times T_n^η

$\text{-}\text{-}\times\text{-}\text{-}$: $U_n^\eta > \lambda(X_{T_n^\eta -}, \eta)$ and thus no jump

- - - - : jump at time T_n^η to state η

In this construction we obtain

- $\mathbb{P}_{\eta_0}[X_0 = \eta_0] = 1$
- $\mathbb{P}_{\eta_0}[X_{t+s} = \eta' \mid X_t = \eta] = \lambda(\eta, \eta')s + o(s)$
- law of $\Theta_s X$ is $\mathbb{P}_{\mu_s}$ if $X_s \sim \mu_s$.

Markov semigroup : $f : S \to \mathbb{R}$ function

Define $T_t f : S \to \mathbb{R}$ by $(T_t f)(\eta_0) = E_{\eta_0}[f(X_t)]$.

By Markov property, we get $T_t \circ T_s = T_{t+s}$,

i.e. $(T_t)_{t \geq 0}$ is a semigroup of operators defined on the space of functions on S.

Stationary distribution :

A proba meas. μ on S is said to be stationary for the process $(X_t)_{t \geq 0}$ if under $\mathbb{P}_\mu$ we have $X_t \sim \mu$. $(X_0 \sim \mu \implies X_t \sim \mu)$

Equivalenty : $$\mu[\{\eta\}] \sum_{\eta' \in S \setminus \{\eta\}} \lambda(\eta, \eta') = \sum_{\tilde\eta \in S \setminus \{\eta\}} \mu[\{\tilde\eta\}] \cdot \lambda(\tilde\eta, \eta) \quad \forall \eta \in S.$$

Example "Dynamical Ising model"

$(\mathcal{G}, E(\mathcal{G}))$ finite graph, $S = \{-1, +1\}^{\mathcal{G}}$

$$\lambda(\sigma, \sigma') = \begin{cases} \dfrac{e^{-\beta H(\sigma')}}{e^{-\beta H(\sigma)} + e^{-\beta H(\sigma')}} & \text{if } \#\{x \in \mathcal{G} \mid \sigma(x) \neq \sigma'(x)\} = 1 \\ 0 & \text{if } \#\{x \in \mathcal{G} \mid \sigma(x) \neq \sigma'(x)\} \neq 1 \end{cases}$$

The Ising measure $\mu[\{\sigma\}] = \frac{1}{Z_\beta} e^{-\beta H(\sigma)}$ is stationary.

Remark simplifies to $(\ldots (2\beta\sigma(x))(B + \Sigma \ldots \sigma(y)))^{-1}$

INTERACTING PARTICLE SYSTEMS

Process $(\xi_t)_{t\geq 0}$ $\quad \xi_t \in S^{\mathbb{Z}^d}$ such that for $x \in \mathbb{Z}^d, \alpha \in S$

$\xi \in S^{\mathbb{Z}^d}$, $\xi(x) \neq \alpha$ $\quad \mathbb{P}\left[\xi_{t+s}(x) = \alpha \mid \xi_t = \xi\right] = s \cdot c_\alpha(x,\xi) + o(s)$ $\quad \circledast$

Note that we only look at change in $\xi_t(x)$ for one x, not the entire ξ_t in the uncountable space $S^{\mathbb{Z}^d}$.

where the flip rates $c_\alpha(x,\xi)$ are of the form

$$c_\alpha(x,\xi) = g_\alpha(\xi(x+z_1), \dots, \xi(x+z_n))$$

for some fixed finite subset $\mathcal{N} = \{z_1, \dots, z_n\} \subset \mathbb{Z}^d$.

The flip rates are translation invariant and depend on the configuration ξ only locally.

Dynamical infinite volume Ising model

$S = \{-1, +1\}$

$$c_\varepsilon(x,\xi) = \left[1 + \exp\left(-2\beta\varepsilon\left(B + \sum_{y: \|y-x\|=1} \xi(y)\right)\right)\right]^{-1}$$

$x \in \mathbb{Z}^d$
$\varepsilon \in \{-1,+1\}$
$\xi \in \{-1,+1\}^{\mathbb{Z}^d}$

Here $\mathcal{N} = \{z \in \mathbb{Z}^d \mid \|z\| = 1\}$

Interpretation: at unit rate, each site x resets its spin to local thermal equilibrium with its neighbors.

The voter model

$\mathbb{Z}^d$ "array of houses"

S a finite set of "opinions"

e.g. $S = \{$republican, democrat$\}$

or $S = \{$Niinistö, Haavisto, Väyrynen, Soini, Lipponen, Arhinmäki, Biaudet, Essayah$\}$

$\xi_t \in S^{\mathbb{Z}^d}$ $\quad \xi_t(x) =$ "opinion of the inhabitant of house x at time t"

$\alpha \in S$, $x \in \mathbb{Z}^d$, $\xi \in S^{\mathbb{Z}^d}$, rate:

$$c_\alpha(x, \xi) = \#\{y \in \mathbb{Z}^d \mid \|y-x\| = 1,\ \xi(y) = \alpha\}$$

The inhabitant of any house x adopts at a constant rate the opinion of any one of her neighbors.

The contact process

$\lambda, \delta > 0$ parameters

$\mathbb{Z}^d$ "array of possible locations of plants"

$S = \{0,1\}$

$\xi_t \in \{0,1\}^{\mathbb{Z}^d}$ $\quad \xi_t(x) = \begin{cases} 1 & \text{"a living plant at } x \text{ at time } t\text{"} \\ 0 & \text{"no plant at } x \text{ at time } t\text{"} \end{cases}$

$x \in \mathbb{Z}^d$, $\xi \in \{0,1\}^{\mathbb{Z}^d}$ rates

$$c_0(x, \xi) = \delta \qquad \text{if} \quad \xi(x) = 1$$

$$c_1(x, \xi) = \lambda \cdot \sum_{y\,:\,\|y-x\|=1} \xi(y) \qquad \text{if} \quad \xi(x) = 0$$

Plants die at rate δ and are born at rate proportional to the number of plants at neighboring sites.

Question 0: Do these rates specify a unique Markov process $(\xi_t)_{t\geq 0}$ on $S^{\mathbb{Z}^d}$?

In this lecture we will give a positive answer to this basic question

Question 1: Do these processes have interesting ("non-trivial") stationary distributions ?

Note, for example:

- for voter model, $s_0 \in S$, the constant state $\bar{\xi}$, $\bar{\xi}(x) = s_0 \quad \forall x \in \mathbb{Z}^d$ is "absorbing": if $\xi_0 = \bar{\xi}$ then $\xi_t = \bar{\xi} \quad \forall t \geq 0$

  ($\delta_{\bar{\xi}}$ is stationary distrib.)

  "consensus"
  "everybody holds the same opinion"

- for contact process the constant state $\bar{\xi}$, $\bar{\xi}(x) = 0 \quad \forall x \in \mathbb{Z}^d$ is "absorbing"

  ($\delta_{\bar{\xi}}$ stationary distrib.) "no living plants anywhere"

More interesting stationary distributions may exist, and they may govern the long time behavior of the systems. For example:

Theorem For the voter model with $S = \{0,1\}$ we have

- if $d \leq 2$ then for any ξ_0 and any $x, y \in \mathbb{Z}^d$ we have $\mathbb{P}[\xi_t(x) \neq \xi_t(y)] \xrightarrow[t\to\infty]{} 0$.
- if $d > 2$ and $\xi^\theta = (\xi_t^\theta)_{t\geq 0}$ denotes the process started from the distribution $\xi_0^\theta \sim \text{Bernoulli}(\theta)^{\otimes \mathbb{Z}^d}$, $\theta \in (0,1)$, we have $\xi_t^\theta \xrightarrow[t\to\infty]{w} \xi_\infty^\theta$, where the law of ξ_∞^θ is translation invariant and stationary and $\forall x \in \mathbb{Z}^d$ and not a convex combination of delta measures on constant configurations.

Theorem Fix $d \in \mathbb{N}$ and consider the contact process. Then we have: $\exists$ nontrivial translation inv. stationary dist. $\iff \frac{\delta}{\lambda} < r_c$.

CONSTRUCTION OF INTERACTING PARTICLE SYSTEMS

We now prove the existence of processes $(\xi_t)_{t\geq 0}$ with given flip rates $c_\alpha(x,\xi) = g_\alpha(\xi(x+z_1),\dots,\xi(x+z_n))$.

Again, since $\max\limits_{\alpha,s_1,\dots,s_n\in S} g_\alpha(s_1,\dots,s_n)$ is finite, performing a linear time change if necessary we may assume

$$c_\alpha(x,\xi) \leq 1 \qquad \forall x\in\mathbb{Z}^d,\ \xi\in S^{\mathbb{Z}^d},\ \alpha\in S.$$

Construction

- Input: Take for each $x\in\mathbb{Z}^d$, $\alpha\in S$ an indep. unit rate Poisson process with arrival times $(T_n^{x;\alpha})_{n\in\mathbb{N}}$ and independent i.i.d. $(U_n^{x;\alpha})_{n\in\mathbb{N}}$ $U_n^{x;\alpha}\sim \mathrm{Unif}([0,1])$.
- Prescription Start from a given $\xi_0\in S^{\mathbb{Z}^d}$.

  At time $t=T_n^{x;\alpha}$ flip value at x to α if $U_n^{x;\alpha}\leq c_\alpha(x,\xi_{t-})$.

Remark: Infinitely many Poisson processes $\longrightarrow$ no first arrival (in an arbitrarily small time interval there are infinitely many arrivals). Thus it is not obvious that the prescription is well-defined. We will prove well-definedness by a non-percolation argument due to Harris.

Consider the following percolation variant on $\mathbb{Z}^d$ with bonds $\{x,y\}$ s.t. $y-x \in \mathcal{N}^*$ where $\mathcal{N} = \{z_1, \dots, z_n\}$ and $\mathcal{N}^* = \{+z_1, -z_1, +z_2, -z_2, \dots, +z_n, -z_n\}$.

Fix $t_0 > 0$ (small). Declare the bond $\{x,y\}$ open iff for some $\alpha \in S$ either $T_1^{x;\alpha} \leq t_0$ or $T_1^{y;\alpha} \leq t_0$

(An open bond between x and y indicates that to determine the state of x at time t_0 we may have needed the knowledge of the state at y, or vice versa.)

Define connected components as usual.

Theorem For $t_0 > 0$ small enough, almost surely all components of the above percolation model are finite

Corollary The prescription above gives rise to a well-defined Markov process $(\xi_t)_{t \geq 0}$ with flip rates $c_a(x, \xi)$.

Proof of Corollary: Up to time t_0, the state of any $x \in \mathbb{Z}^d$ can be determined from the finitely many arrival times $T_n^{y;\alpha} \leq t_0$ with y in the finite component of x, and the corresponding uniform random variables $U_n^{y;\alpha}$. To continue up to $2t_0$ we repeat the procedure from the random state ξ_{t_0}, and iterate. $\square$

Proof of Theorem:

We prove that the component of 0 is a.s. finite. By translation invariance it follows that the component of any given $x \in \mathbb{Z}^d$ is a.s. finite, and by countable (sub)additivity it follows that a.s. all components are finite.

We observe that if the component of 0 were not finite, then for any $\ell \in \mathbb{N}$ there would exist open paths of length ℓ starting at 0, i.e. distinct sites $x_1, x_2, \dots, x_\ell$ s.t. $\{0, x_1\}, \{x_1, x_2\}, \dots, \{x_{\ell-1}, x_\ell\}$ are all open.

Note the following:

- the number of paths $0 = x_0, x_1, \dots, x_\ell$ s.t. $\{x_{j-1}, x_j\} \in \mathcal{N}^*$ is at most N^ℓ where $N = \#\mathcal{N}^*$
- the probability that a bond $\{x_{j-1}, x_j\}$ is open is at most $(1 - \exp(-2\kappa t_0))$ where $\kappa = \# S$ (this is the probability that at least one of the 2κ independent Poisson processes $(T_n^{\alpha; x_j}), (T_n^{\alpha; x_{j-1}})$ has an arrival before t_0)
- if $x, y, z, w \in \mathbb{Z}^d$ are distinct, then the openness of $\{x,y\}$ and $\{z,w\}$ are independent (they are determined by independent Poisson processes)

Using the above, we estimate that the probability of existence of an open path of length ℓ is at most $N^\ell \cdot (1 - e^{-2\kappa t_0})^{\lfloor \ell/2 \rfloor}$. Choose $t_0 > 0$ small enough so that $(1 - e^{-2\kappa t_0}) < \frac{1}{4N^2}$. Then as $\ell \to \infty$ the probability tends to zero, showing that almost surely the component of 0 is finite. $\square$

The associated semigroup

If $f: S^{\mathbb{Z}^d} \to \mathbb{R}$ is bounded, we can define, for any $t \geq 0$, $T_t f : S^{\mathbb{Z}^d} \to \mathbb{R}$ by

$$(T_t f)(\xi_0) = \mathbb{E}_{\xi_0}[f(\xi_t)].$$

Again, by Markov property, $T_t \circ T_s = T_{t+s}$, i.e., $(T_t)_{t \geq 0}$ is a semigroup of operators.

Proposition $(T_t)_{t\geq 0}$ is a Feller semigroup, i.e.,
if $f: S^{\mathbb{Z}^d} \to \mathbb{R}$ is continuous (w.r.t. product topology, c.f. metric on countable product of finite sets)
then also $T_t f$ is continuous

Proof: First of all, $S^{\mathbb{Z}^d}$ is compact, so a continuous f is bounded and $T_t f$ is well-defined. Let $(\xi_0^{(n)})_{n\in\mathbb{N}}$ be a sequence $\xi_0^{(n)} \to \xi_0$. Recall that this means that for any $x \in \mathbb{Z}^d$, for n large enough, $\xi_0^{(n)}(x) = \xi_0(x)$. Consider the processes $(\xi_t^{(n)})_{t\geq 0}$ started from $\xi_0^{(n)}$ for $t \leq t_0$ first. We can use the same input in the construction, i.e. the same $(T_n^{x;\kappa})$ and $(U_n^{x;\kappa})$.
Fixing $x \in \mathbb{Z}^d$, the component of x is finite, and therefore for n large enough $\xi_0^{(n)}(y) = \xi_0(y)$ for all y in the component of x. But then by construction $\xi_t^{(n)}(x) = \xi_t(x)$ for $t \leq t_0$.
We see that for $t \leq t_0$ $\xi_t^{(n)} \xrightarrow[n\to\infty]{} \xi_t$. By dominated convergence (recall: f is continuous and thus bounded) we get $(T_t f)(\xi_0^{(n)}) = \mathbb{E}[f(\xi_t^{(n)})] \longrightarrow \mathbb{E}[f(\xi_t)] = (T_t f)(\xi_0)$.
This shows $T_t f$ is continuous if $t \leq t_0$.
By semigroup property $T_{t+s} f = T_t(T_s f)$, this holds for all $t \geq 0$. $\square$

INTERACTING PARTICLE SYSTEMS, CONTINUED

VOTER MODEL CLUSTERING AND COEXISTENCE

Recall the voter model:

- S finite set of "opinions" $(\# S \geq 2)$
- process $(\xi_t)_{t \geq 0}$ $\xi_t \in S^{\mathbb{Z}^d}$ with flip rates

$$\frac{1}{s} \mathbb{P}\left[\xi_{t+s}(x) = \alpha \mid \xi_t = \xi \right] \longrightarrow n_\alpha^{(x)}(\xi) \qquad \begin{pmatrix} \alpha \in S \\ \xi \in S^{\mathbb{Z}^d} \\ \xi(x) \neq \alpha \end{pmatrix}$$

$$:= \# \left\{ y \in \mathbb{Z}^d \mid \|y-x\| = 1, \ \xi(y) = \alpha \right\}$$

= "number of neighbors of x with opinion α"

Trivial stationary measures : "consensus"

$\alpha_0 \in S$, constant config. $\bar{\xi}^{(\alpha_0)}(x) = \alpha_0 \quad \forall x \in \mathbb{Z}^d$

Delta measure on constant configuration $\delta_{\bar{\xi}^{(\alpha_0)}}$ is trivially stationary.

Any convex combination $\mu = \sum_{\alpha_0 \in S} p_{\alpha_0} \cdot \delta_{\bar{\xi}^{(\alpha_0)}}$ $\left(\sum p_{\alpha_0} = 1 \right)$ is of course also stationary.

These μ are called trivial, other stationary measures non-trivial.

Q: Are there "interesting" (non-trivial) stationary measures?

A:
- Yes, if $d > 2$.
- No, if $d \leq 2$.

this dichotomy is the main goal of this lecture

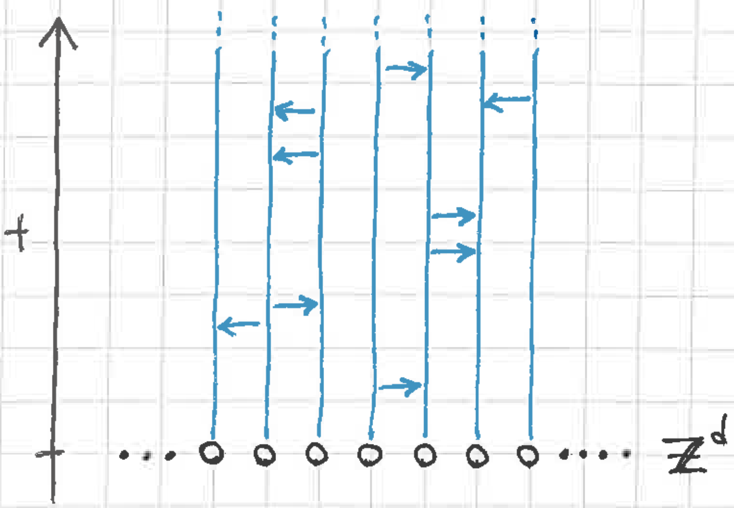
For simplicity of notation, we take $S = \{0, 1\}$, but the results and proofs carry through with minor modifications to the case of general finite S.

ANOTHER CONSTRUCTION AND DUALITY

We give an alternative construction of the voter model process $(\xi_t)_{t\geq 0}$, which is quite similar in spirit to the construction of interacting particle systems in the last lecture. The alternative construction makes a dual process of coalescent random walks more transparent.

For every pair (x,y) of sites $x,y \in \mathbb{Z}^d$ that are neighbors, $\|x-y\| = 1$, take an independent unit rate Poisson process with arrival times denoted $(T_n^{x\to y})_{n\in\mathbb{N}}$.

Graphically, we draw a time axis above each site $x \in \mathbb{Z}^d$ and draw an arrow from y to x at the arrival times $(T_n^{x\to y})_{n\in\mathbb{N}}$.



The process is started from some given configuration $\xi_0 \in \{0,1\}^{\mathbb{Z}^d}$.

At times $T_n^{x\to y}$, the voter at x adopts the opinion held by voter at y at that time.

The well-definedness of this prescription follows from an easier dual description, but could also be proven similarly to last lecture.

The construction indeed gives rise to the correct flip rates:

$$\mathbb{P}\left[\xi_{t+s}(x) = \alpha \,\middle|\, \xi_t = \xi\right] = \mathbb{P}\left[\text{for some } y, \text{ on } (t,t+s] \text{ there is an arrival } T_n^{x\leftarrow y} \text{ and no other arrivals to } x \text{ and its nbrs, and } \xi(y)=\alpha\right] + o(s)$$

$$= \sum_{y:\|y-x\|=1} \mathbb{P}\left[\text{on } (t,t+s] \text{ there is } T_n^{x\to y} \text{ and no others, and } \xi(y)=\alpha\right] + o(s)$$

$\leq$ probability of several arrivals on $(t,t+s)$

$$= s \cdot \sum_{y:\|y-x\|=1} \mathbb{1}_{\xi(y)=\alpha} + o(s) = s\cdot n_\alpha^{(x)}(\xi) + o(s).$$

The dual process

Note that for any $t \geq 0$ and $x \in \mathbb{Z}^d$, we can descend the time axis until an arrow points out of it to some neighbor of x, then keep descending and following arrows. This process tracks where the opinion held by voter x at time t originates. Denote this process by $(D_s^{x;t})_{s\in[0,t]}$ so that $s \in [0,t]$ denotes the amount we have descended on the time axis.

Then,
$$\xi_t(x) = \xi_{t-s}(D_s^{x;t}) \qquad \text{by construction.}$$

From the construction it is clear that $(D_s^{x;t})_{s\in[0,t]}$ is a continuous time random walk on $\mathbb{Z}^d$, started from $D_0^{x;t} = x$, and with jump rate 1 to each of the neighbors. In other words, its law is that of a discrete time simple random walk $(S_n)_{n\in\mathbb{Z}_{\geq 0}}$ on $\mathbb{Z}^d$, time-changed by a Poisson process $(P_s)_{s\in[0,t]}$ of intensity $2d$ (number of neighbors of a site), shifted by x
$$D_s^{x;t} \overset{(d)}{=} S_{P_s} + x .$$

The processes $D_s^{x;t}$ for different $x \in \mathbb{Z}^d$ are not independent. Once $D_{s_0}^{x;t} = D_{s_0}^{y;t}$ for some $s_0 \geq 0$, the arrows followed by the two are the same, so $D_s^{x;t} = D_s^{y;t}$ for all $s \geq s_0$. Until the first such $s_0 = s_0^{x,y;t}$, the processes are determined by independent Poisson processes (either $(T_n^{x\to y})$ for different x and y, or the arrivals on disjoint time intervals of $(T_n^{x\to y'})$), and thus they are independent, $(D_s^{x;t})_{s\in[0,s_0)} \perp\!\!\!\perp (D_s^{y;t})_{s\in[0,s_0)}$.
The collection of these is called coalescent random walks.

The dichotomy of existence/non-existence of non-trivial stationary measures for the voter model is closely related to transience/recurrence of the random walks.

It natural to extend the coalescent random walks to all times $s \geq 0$, i.e. to processes $(D_s^x)_{s\geq 0}$ for $x \in \mathbb{Z}^d$

These can be constructed by following Poissonian arrows between neighboring sites not only for $s \in [0,t]$ but for $s \in [0,\infty)$. We no longer use the same arrows as the voter model construction, so the process is defined on a different probability space, but we only care about its law. Then for any $t \geq 0$

$$(\xi_t(x))_{x\in\mathbb{Z}^d} \overset{(d)}{=} (\xi_0(D_t^x))_{x\in\mathbb{Z}^d}$$

is an equality of laws (Compare with: $\xi_t(x) = \xi_0(D_t^{x;t})$ a.s.)

<u>Theorem 1:</u> Let $d \leq 2$. For any initial configuration $\xi_0 \in \{0,1\}^{\mathbb{Z}^d}$ and any $x,y \in \mathbb{Z}^d$ we have

"clustering": $\mathbb{P}[\xi_t(x) \neq \xi_t(y)] \longrightarrow 0$ as $t \to \infty$.

<u>Proof</u>: Note that for whichever initial configuration ξ_0, we have $\mathbb{P}[\xi_t(x) \neq \xi_t(y)] \leq \mathbb{P}[D_t^x \neq D_t^y]$.

Then note that $D_t = D_t^x - D_t^y$ is a continuous time random walk until it hits zero, and then it stays at zero. The jump rates of D_t to each of the neighbors are 2 (the intensities of the Poisson processes governing D_t^x and D_t^y add up).

Before hitting zero: $D_t \overset{(d)}{=} S_{P_t} + x - y$, where $(P_t)_{t\geq 0}$ is a Poisson process with intensity $4d$, and $(S_n)_{n\in\mathbb{N}}$ is a standard discrete time random walk on $\mathbb{Z}^d$.

For $d \leq 2$ (S_n) is recurrent, and since $P_t \longrightarrow \infty$ a.s. as $t \to \infty$, we get that almost surely, $D_t = 0$ for some $t \geq 0$ (and then stays at 0).

Thus $\mathbb{P}[\xi_t(x) \neq \xi_t(y)] \leq \mathbb{P}[D_t \neq 0] \xrightarrow[t\to\infty]{} 0$. $\square$

Corollary: If $d = 2$, then any stationary measure is trivial.

To show the existence of non-trivial stationary measures in $d > 2$, we start the process from the product measure

$$\mu_\theta = \left(\text{Bernoulli}(\theta)\right)^{\otimes \mathbb{Z}^d} \qquad \theta \in [0,1].$$

i.e. $(\xi_0(x))_{x \in \mathbb{Z}^d}$ are i.i.d. with $\mathbb{P}[\xi_0(x) = 1] = \theta$
$\mathbb{P}[\xi_0(x) = 0] = 1 - \theta$

Theorem 2: Let $d > 2$ and $\xi_0 \sim \mu_\theta$ for $\theta \in (0,1)$.

Then, as $t \to \infty$, we have $\xi_t \xrightarrow{w} \xi^{(\theta)}$, where $\xi^{(\theta)}$ follows a non-trivial stationary distribution.

Proof: Recall that on $\{0,1\}^{\mathbb{Z}^d}$ to prove weak convergence, it is sufficient to show that probabilities of cylinder events converge, and by inclusion-exclusion (as in thermodynamical limit of Ising model) it is in fact sufficient to show that the limits

$$\lim_{t \to \infty} \mathbb{P}[\xi_t(x) = 1 \quad \forall x \in A]$$

exist for all finite subsets $A \subset \mathbb{Z}^d$.

Let $A \subset \mathbb{Z}^d$ be finite, and consider D_t^x for $x \in A$. Denote by $N_t^A = \#\{D_t^x \mid x \in A\}$ the number of distinct positions of these coalescent random walks at time t. Clearly $t \mapsto N_t^A$ is decreasing because of the coalescence, and thus $N_t^A \searrow N_\infty^A$. By duality

$$\mathbb{P}[\xi_t(x) = 1 \quad \forall x \in A] = \mathbb{P}[\xi_0(D_t^x) = 1 \quad \forall x \in A] = \mathbb{E}[\theta^{N_t^A}].$$

But $\theta^{N_t^A} \nearrow \theta^{N_\infty^A}$ so by monotone convergence these probabilities have limits. Thus ξ_t converges weakly as $t \to \infty$ to some $\xi^{(\theta)}$, with the property

$$\mathbb{P}[\xi^{(\theta)}(x) = 1 \quad \forall x \in A] = \mathbb{E}[\theta^{N_\infty^A}].$$

If $A = \{x\}$ then $N_t^A = 1$ $\forall t \geq 0$, so
$$\mathbb{P}[\xi^{(\theta)}(x) = 1] = \theta.$$
The only trivial stationary measure with this property is $\mu = \theta \cdot \delta_{\bar{\xi}(1)} + (1-\theta) \cdot \delta_{\bar{\xi}(0)}$. We will show that if $d > 2$, then the law of $\xi^{(\theta)}$ is not μ.

Let $x, y \in \mathbb{Z}^d$ be distinct, $x \neq y$.
Then $\mu[\xi(x) = 1, \xi(y) = 1] = \theta$.
The random walk $D_t = D_t^x - D_t^y$ in dimension $d > 2$ is transient, so there is a positive probability $c_{x,y} > 0$ that $D_t \neq 0$ for all $t \geq 0$.
On this event, if $A = \{x, y\}$, we have $N_t^A = 2$ $\forall t \geq 0$.
Therefore $\mathbb{P}[\xi^{(\theta)}(x) = 1, \xi^{(\theta)}(y) = 1] = \mathbb{E}[\theta^{N_t^A}]$
$$= c_{x,y} \cdot \theta^2 + (1 - c_{x,y})\,\theta \neq \theta.$$
This shows that the distribution of $\xi^{(\theta)}$ is not trivial if $d > 2$.
Stationarity of the distribution of $\xi^{(\theta)}$ follows by a standard argument (see below).
$\square$

<u>Theorem</u> For $(\xi_t)_{t \geq 0}$ a Markov process on a metric space $\mathcal{X}$, whose Markov semigroup $(T_t)_{t \geq 0}$ is Feller (i.e. $f: \mathcal{X} \to \mathbb{R}$ continuous $\Longrightarrow$ $T_t f: \mathcal{X} \to \mathbb{R}$ continuous) if $\xi_t \xrightarrow{w} \xi_\infty$ as $t \to \infty$, then the law of ξ_∞ is a stationary distribution.

<u>Proof</u>: Recall that a proba measure ν on $\mathcal{X}$ is determined by the integrals $\int f(\xi)\,d\nu(\xi)$ of bounded contin. functions $f: \mathcal{X} \to \mathbb{R}$.
For ν a proba measure, define νT_s by
$$\int f(\xi)\, d(\nu T_s)(\xi) = \int (T_s f)(\xi)\, d\nu(\xi) = \int \mathbb{E}_{\xi_0 = \xi}[f(\xi_s)]\, d\nu(\xi).$$
for all bounded continuous f — continuous, since (T_t) is Feller

The rightmost expression shows that νT_s is the law of ξ_s if $\xi_0 \sim \nu$.

Suppose now $(\xi_t)_{t\geq 0}$ is such that $\xi_t \xrightarrow{w} \xi_\infty \sim \mu$ and denote by μ_t the law of ξ_t, so $\mu_t \xrightarrow{w} \mu$.

We want to show that $\mu_t T_s \longrightarrow \mu T_s$ as $t \to \infty$.

This follows, since for all bounded continuous f

$$\lim_{t\to\infty} \int f(\xi)\, d(\mu_t T_s)(\xi) = \lim_{t\to\infty} \int (T_s f)(\xi)\, d\mu_t(\xi)$$

Feller $\Rightarrow T_s f$ is bounded continuous & $\mu_t \xrightarrow{w} \mu$ $\rightsquigarrow$

$$= \int (T_s f)(\xi)\, d\mu(\xi) = \int f(\xi)\, d(\mu T_s)(\xi).$$

But $\mu_t T_s = \mu_{t+s}$ by Markov property. Therefore

$$\begin{array}{ccc} \mu_t T_s & = & \mu_{t+s} \\ w \big\downarrow t\to\infty & & w \big\downarrow t\to\infty \\ \mu T_s & = & \mu \end{array}$$

This shows that μ is stationary. $\square$