Math 833 Branching graphs Week 4

We present a general formalism and then apply it to prove the De Finetti's theorem.

Setup: For each $n = 0, 1, 2, \ldots$ we are given a measurable space A_n (i.e. equipped with σ-algebra) and Markov kernel ("link") $L_{n+1}(x, B)$

element of A_{n+1}

measurable subset of A_n

[$L_{n+1}(x, B)$ is a measurable function of x and a probability measure in B]

$$A_0 \xleftarrow{L_1} A_1 \xleftarrow{L_2} A_2 \xleftarrow{L_3} A_3 \xleftarrow{L_4} \cdots$$

Like a Markov chain, but in inverse time.

Links are also called cotransition probabilities.

Definition: Coherent system of measures is a sequence $(M_0, M_1, M_2, \ldots)$, such that

- M_n is a probability measure on $\mathbb{A}_n$, $n = 0, 1, 2, \ldots$
- $M_{n+1} L_{n+1} = M_n$, $n = 0, 1, 2, \ldots$

$$\left[\, M_n(B) = \int_{\mathbb{A}_{n+1}} L_{n+1}(x, B)\, M_{n+1}(dx) \,\right]$$

Interpretation: M_n are fixed time distributions of a Markov chain with (co-)transitions L_n in inverse time

$$\xrightarrow{L_4} M_3 \xrightarrow{L_3} M_2 \xrightarrow{L_2} M_1 \xrightarrow{L_1} M_0$$

The initial condition of such a Markov chain is at $n = +\infty$

Essentially, this is our topic of interest

Example 1 $\quad A_n = \{ a_1 \le a_2 \le \ldots \le a_n \}$

$\qquad\qquad\qquad\qquad\quad \uparrow\mathbb{R} \qquad\qquad \uparrow\mathbb{R}$

L_n is a deterministic transition:

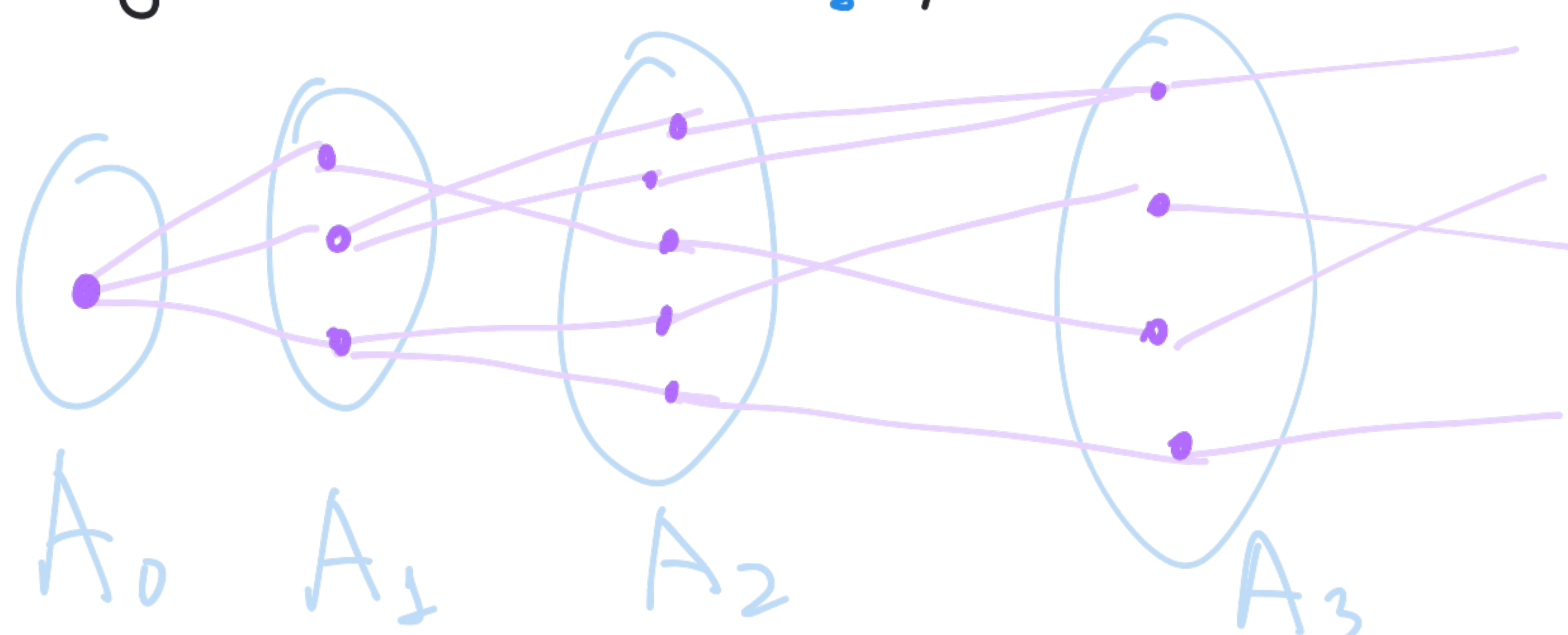
$$L_n\big((a_1, \ldots a_n) , B \big) = \mathbb{1}_{(b_1, \ldots b_{n-1}) \in B} \quad , \text{where}$$

$$\prod_{i=1}^{n-1} (z - b_i) = \frac{1}{n} \frac{\partial}{\partial z} \prod_{i=1}^{n} (z - a_i)$$

In words, L_n transitions from roots of $\Pi(z-a_i)$ to roots of its derivative.

The study of coherent systems with respect to such links was precisely the topic of Week 2.

Construction: Let A_n be finite or countable and suppose that we are given a graph with edges linking A_n to A_{n+1}, $n=0,1,2,\ldots$.



Assume that A_0 contains a unique element $\varnothing$ ("root")

For $\lambda \in A_n$ denote $\dim(\lambda)$ "dimension"

$$= \#\{\varnothing = \lambda^0 \to \lambda^1 \to \lambda^2 \to \cdots \to \lambda^n = \lambda\}$$

an edge in the graph $\quad$ A_1 $\quad$ A_2 $\quad$ (paths from $\varnothing$ to λ)

For graphs of representation-theoretic origin $\dim(\lambda)$ = dimension of representation λ

Definition: A graded graph $(A_0, A_1, A_2, \dots)$ as in the last slide is called a branching graph. It defines links and coherency relations through

$$L_{n+1}(\lambda \to \mu) = \begin{cases} \dfrac{\dim \mu}{\dim \lambda}, & \text{there is an edge } \mu - \lambda, \\ 0, & \text{otherwise.} \end{cases}$$

cotransitional probability from $\lambda \in A_{n+1}$ to $\mu \in A_n$

Since $\displaystyle\sum_{\substack{\mu \in A_n \\ (\text{linked to } \lambda)}} \frac{\dim \mu}{\dim \lambda} = 1$, $L_{n+1}(\lambda \to \cdot)$ is a probability measure

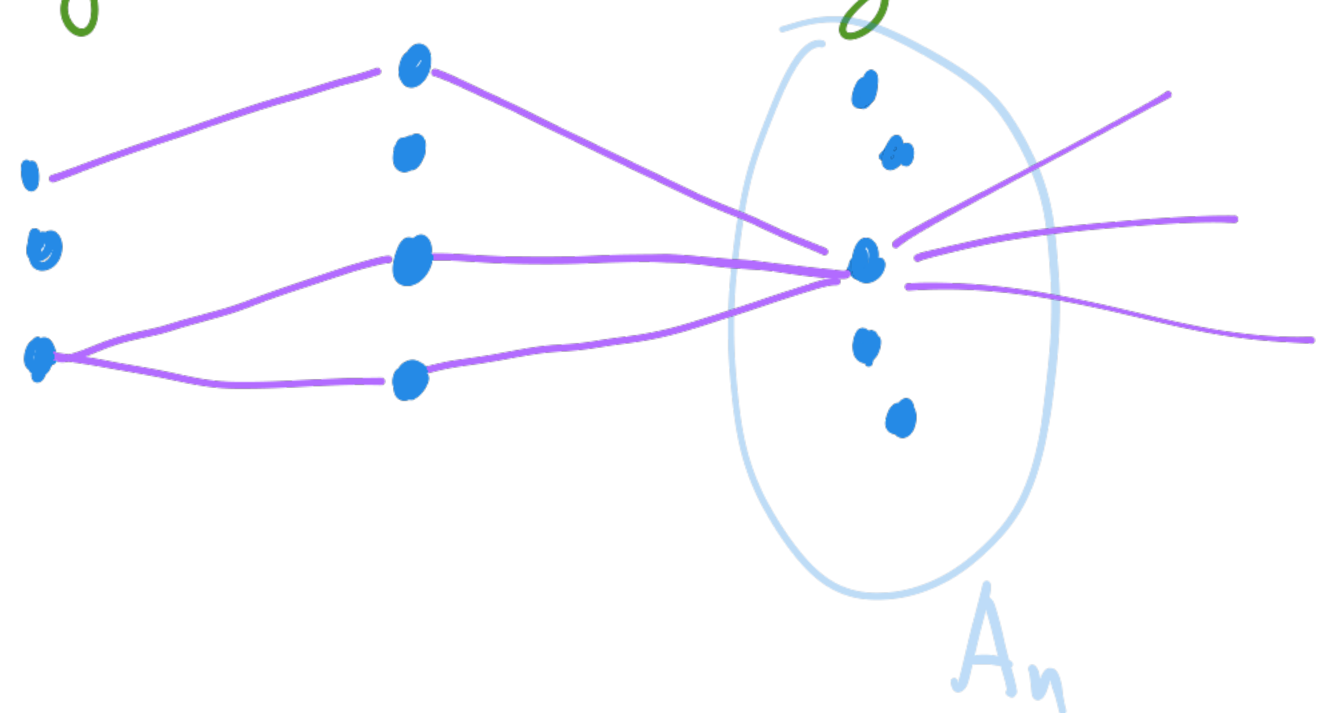
Theorem 1: Coherent systems of measures on A_n are in bijection with probability measures on space of infinite paths $\tau = (\underset{\emptyset}{\lambda^0} \to \underset{A_1}{\lambda^1} \to \underset{A_2}{\lambda^2} \to \dots)$ which are central

edge in graph

which means $\mathrm{Prob}(\lambda^0 = x_0, \lambda^1 = x_1, \dots, \lambda^n = x_n) = f(x_n)$, i.e. does not depend on $x_0, \dots, x_{n-1}$

Correspondence is by $M_n(x) = \mathrm{Prob}(\lambda^n = x)$

Meaning of centrality:



Conditional on $\lambda^n \in A_n$ all initial segments $\lambda^0 \to \lambda^1 \to \dots \to \lambda^{n-1}$ are equiprobable

---

Proof of Theorem 1:

$\boxed{\text{I}}$ Take a central measure p on paths and define $M_n(x) := \text{Prob}\,(\lambda^n = x)$, where $\lambda^0 \to \lambda^1 \to \lambda^2 \to \dots$ is a p-random path. Clearly for each n, $M_n(x)$ is a probability measure on A_n. Let us check the coherency relation.

$$M_n(x) = \sum_{\substack{x_0,\dots x_{n+1}\\ x_n = x}} \text{Prob}(\lambda^0 = x_0, \dots, \lambda^{n+1} = x_{n+1}) =$$

by centrality the probability depends only on x_{n+1}

$$= \sum_{\substack{x_0 \to x_1 \to \dots \to x_{n+1}\\ x_n = x}} \frac{M_{n+1}(x_{n+1})}{\dim(x_{n+1})}$$

summing over $x_0 \to \dots \to x_n$

$$= \sum_{x = x_n \to x_{n+1}} M_{n+1}(x_{n+1}) \cdot \frac{\dim(x)}{\dim(x_{n+1})}$$

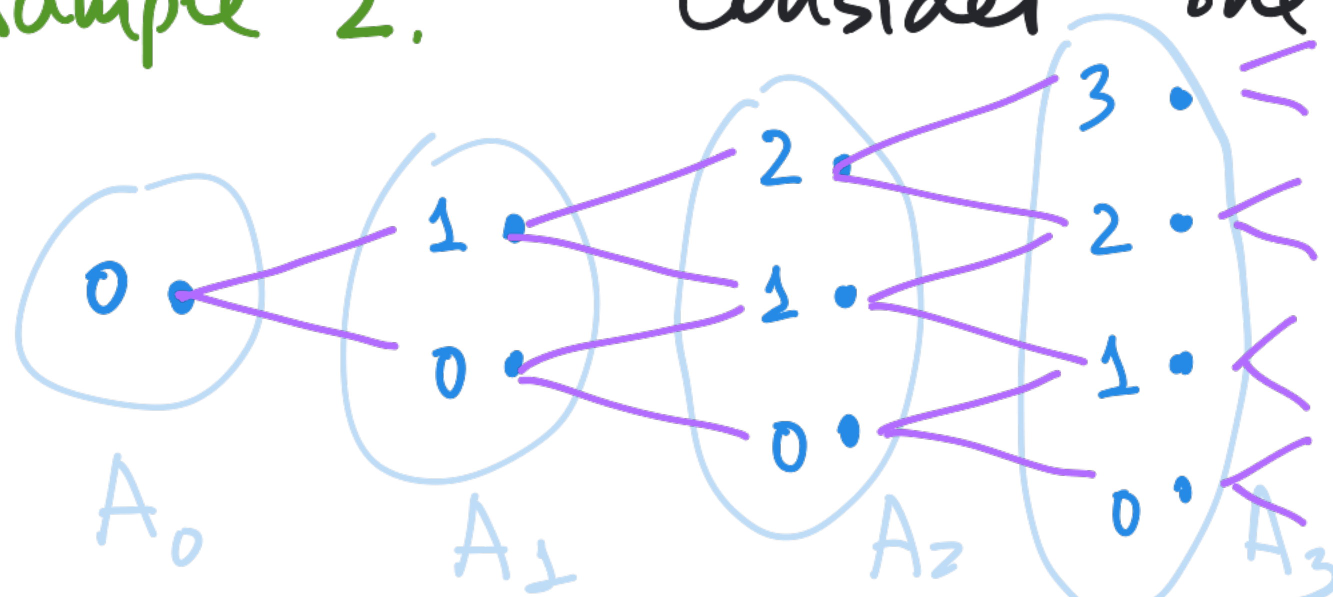
[as desired!].

$\boxed{\text{II.}}$ Given a coherent system $\{M_n\}_{n=0,1,2,\dots}$ define a random path $\lambda^0 \to \lambda^1 \to \lambda^2 \to \dots$ through finite-dimensional distributions

$$\text{Prob}(\lambda^0 = x_0, \dots, \lambda^n = x_n) := \frac{M_n(x_n)}{\dim(x_n)}$$

- The distributions are consistent over varying n by coherency relation. Hence, by Kolmogorov's consistency theorem they define a probability measure on infinite paths
- Centrality is immediate from definition. $\blacksquare$

Example 2. Consider the Pascal's graph



K on level n is linked to K and $K+1$ on level $n+1$

$[0 \leq K \leq n]$

Exercise 1: Compute the dimensions in this graph and show that the notion of coherent system matches Week 3, Slide 12, Theorem 4.

General problem: Given sets A_n and links $L_n(x, B)$, we would like to describe all coherent systems

In particular, in the branching graph setting we call this problem „identification of boundary of the graph".

Here is one way to construct a coherent system.

Notation: $L_{N\to K}(x, B) = L_N \circ L_{N-1} \circ \dots \circ L_K =$

$$= \int_{\lambda^{N-1}\in A_{N-1}, \dots, \lambda^{K+1}\in A_{K+1}} \cdots \int L_N(x, d\lambda^{N-1}) L_{N-1}(\lambda^{N-1}, d\lambda^{N-2}) \cdot \ldots \cdot L_K(\lambda^{K+1}, B)$$

$x \in A_N$, $B \subset A_K$ and $L_{N\to K}(x, B)$ is the composition

$$\begin{tikzcd} A_N \arrow[r, "L_N"] \arrow[rrr, bend right, dashed, "L_{N\to K}"'] & A_{N-1} \arrow[r, "L_{N-1}"] & \dots \arrow[r, "L_{K+1}"] & A_K \end{tikzcd}$$

and at most countable

Lemma: Assume that either A_n are discrete or each A_n is a topological space and links $L_n(x, A)$ are continuous functions of x when treated as x-dependent measures and with respect to the weak topology (= convergence in distribution) on measures.

Take a sequence $x_N \in A_N$, $N = 1, 2, 3, \dots$ such that for each K the measures $L_{N\to K}(x_N, \cdot)$ (weakly) converge to a measure $\mathcal{M}_K$.

Then $(\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2, \dots)$ is a coherent system

---

The set of all coherent systems obtainable in such a way is called **the Martin boundary**.

Compare with Week 2, Slide 14

Proof of Lemma:

$$\mu_{K+1} L_{K+1} = \left(\lim_{N\to\infty} L_{N\to K+1}(x_N, \cdot)\right) L_{K+1} \overset{\text{by continuity}}{=} \lim_{N\to\infty} \left(L_{N\to K+1}(x_N, \cdot)\, L_{K+1}\right)$$

$$= \lim_{N\to\infty} L_{N\to K}(x_N, \cdot) = \mu_K. \qquad \blacksquare$$

Definition: The set of all extreme points of the convex set of coherent systems is called the minimal boundary

---

Theorem 2: In the setting of branching graphs we always have the set-theoretic inclusion

minimal boundary $\subset$ Martin boundary

Theorem 3: In the setting of branching graphs the set of all coherent systems is always a simplex. In more detail:

- Equip coherent systems with topology and Borel σ-algebra induced from coordinate wise convergence
  $(M_0^N, M_1^N, \dots) \xrightarrow{N\to\infty} (\tilde{M}_0, \tilde{M}_1, \dots)$ iff $\forall K$ $M_K^N \to \tilde{M}_K$
  Then the minimal boundary Ex is a Borel subset
- For each coherent system $(M_0, M_1, \dots)$ there exists a unique probability measure ν on Ex such that

Lebesgue integral

$$M_K(B) = \int_{m\in Ex} M_K^{(m)}(B)\ \nu(dm)$$

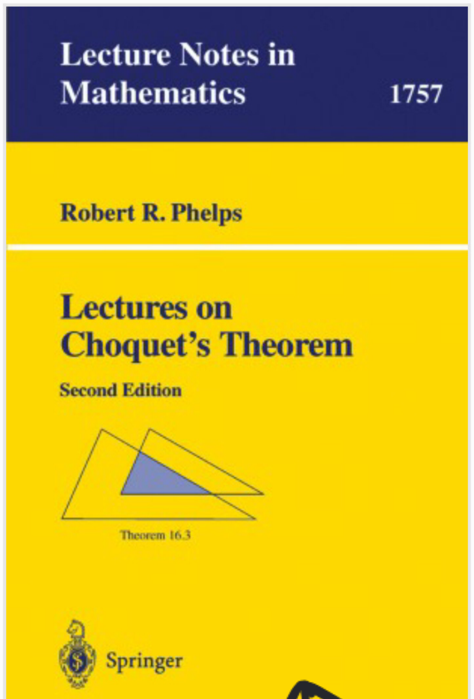
$B \subset A_K$ coherent system from minimal boundary corresponding to m

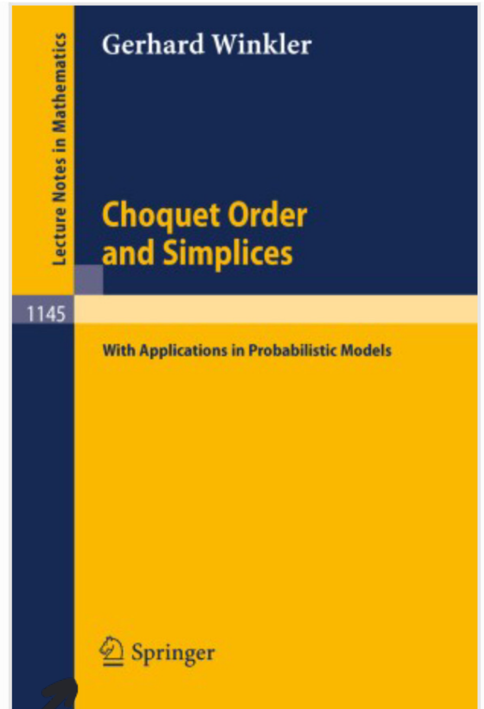
Remark 1: (A_n, L_n) being constructed from branching graphs is not important. Analogues of theorems 2 and 3 hold for any „nice" spaces and links.
[I never encountered in my career a setting where they fail]

Remark 2: In most natural examples minimal boundary = Martin boundary. However, there is no such general theorem and it is not hard to construct a counterexample, i.e. a graph for which this equality fails.

Theorem 2 and 3 are general convex analysis statements. We will not provide full proofs, instead only giving intuition and references to various treatments.

References for theorems 2 and 3





General simplices and extreme points

Dokl. Akad. Nauk SSSR, 1974, Volume 218, Number 4, Pages 749–752 (Mi dan38565)

This article is cited in **22** scientific papers (total in *23* papers)

MATHEMATICS

Description of invariant measures for the actions of some infinite-dimensional groups

A. M. Vershik

Leningrad State University

Vershik's ergodic method

In: Representation of Lie groups and related topics
A. Vershik and Zhelobenko, eds
(Advanced Studies in Contemp. Math. vol. 7)
Gordon & Breach 1990

Journal of Functional Analysis
Volume 205, Issue 2, 20 December 2003, Pages 464-524
ELSEVIER

The problem of harmonic analysis on the infinite-dimensional unitary group

Grigori Olshanski

Section 9: decomposition for countable A_n

7. Unitary representations of infinite-dimensional pairs (G, K) and the formalism of R. Howe

G. I. OL'SHANSKII

Section 22: Approximation in the context of representation theory

PARTIAL EXCHANGEABILITY AND SUFFICIENCY

BY

PERSI DIACONIS and DAVID FREEDMAN

TECHNICAL REPORT NO. 190
JULY 1982

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT MCS80-24649

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

General framework as an extension of De Finetti's theorem

Asymptotics of Jack polynomials as the number of variables goes to infinity

Andrei Okounkov, Grigori Olshanski

International Mathematics Research Notices, Volume 1998, Issue 13, 1998, Pages 641–682,

Section 6: approximation for countable A_n

arXiv.org > math > arXiv:1905.08684

Mathematics > Probability

[Submitted on 21 May 2019 (v1), last revised 20 Aug 2020 (this version, v2)]

The boundary of the orbital beta process

Theodoros Assiotis, Joseph Najnudel

Careful treatment for continuous A_n

Intuition 1 for theorems 2,3: Finite coherent systems $(M_0, M_1, \ldots, M_n)$ are in one-to-one correspondence with probability measures on A_n: this is M_n, which can be arbitrary. Probability measures on A_n form a simplex with extreme points — unit masses at $x \in A_n$

Hence, the space of all coherent measures is $n \to \infty$ limit of simplices.

Intuition 2 for theorems 2 and 3:

One can also interpret the space of infinite coherent systems as intersection of a decreasing sequence of the above finite n simplices. One could expect that extreme point of intersection is approximated by finite n extreme points. Compare with nested intervals in $\mathbb{R}$:

Endpoint of intersection = limit of endpoints.

Application: We finish the proof of De Finetti's theorem from last week. Week 3, Theorem 1 is a particular case of this week's Theorem 3. It remains to prove Week 3, Theorem 2. In present notations it says that:

(slide 8)

Claim: The minimal boundary of Pascal's graph is given by the set of measures parameterized by $p \in [0,1]$. The coherent system $M_n^{(p)}$ on $\{0,1,..n\}$ is given by

$$M_n^{(p)}(m) = p^m (1-p)^{n-m} \binom{n}{m}$$

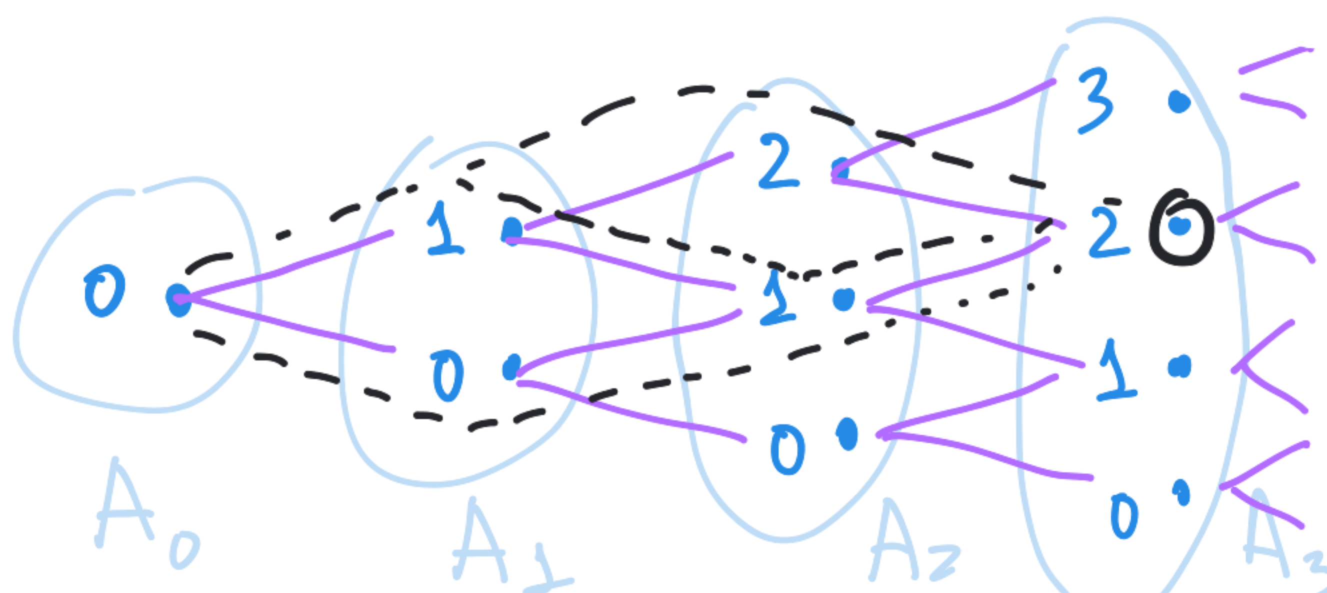
Next 3 slides prove the claim.

In order to prove the claim we compute the Martin boundary first. For that we take $x \in \{0,1,..,N\}$ on level N and compute $L_{N\to K}(x,m)$

Proposition: $$L_{N\to K}(x,m) = \frac{\binom{K}{m}\cdot\binom{N-K}{x-m}}{\binom{N}{x}}, \quad m \le x$$

[0, if $m > x$].

Proof Using identification between coherent systems and measures on paths, we need to study the following problem:



Consider the uniform measure on paths from $0 \in A_0$ to $x \in A_N$. What is the probability that random path passes through $m \in A_K$? There are $\binom{N}{x}$ paths $0 \in A_0 \to x \in A_N$, $\binom{K}{m}$ paths $0 \in A_0 \to m \in A_K$ and $\binom{N-K}{x-m}$ paths $m \in A_K \to x \in A_N$ $\blacksquare$

Question: How should x vary with $N \to \infty$, so that $L_{N \to K}(x, m)$ has $N \to \infty$ limit for each fixed $0 \le m \le K$?

Answer:
$$L_{N \to K}(x, m) = \frac{x!\,(N-x)!\,(N-K)!}{N!\,(x-m)!\,(N-K-(x-m))!} \cdot \binom{K}{m} =$$

$$= \frac{x(x-1)\ldots(x-m+1) \cdot (N-x)(N-x-1)\ldots(N-x-(K-m)+1)}{N(N-1)\ldots(N-K)} \cdot \binom{K}{m}$$

does not change with $N \to \infty$

$$\sim \left(\frac{x}{N}\right)\left(\frac{x}{N} - \frac{1}{N}\right) \cdot \ldots \cdot \left(\frac{x}{N} - \frac{(m-1)}{N}\right) \cdot \left(\frac{N-x}{N}\right)\left(\frac{N-x}{N} - \frac{1}{N}\right) \cdot \ldots \cdot \left(\frac{N-x}{N} - \frac{(K-m)}{N}\right) \cdot \binom{K}{m}$$

these all become negligible as $N \to \infty$

Conclusion: $L_{N \to K}(x, m)$ converges as $N \to \infty$ if and only if $x = x(N)$ is such that $\lim_{N \to \infty} \frac{x}{N} = p \in [0,1]$. In this case

$$\lim_{N \to \infty} L_{N \to K}(x, m) = p^m (1-p)^{K-m} \cdot \binom{K}{m}$$ as in the claim of slide 16.

We have shown that the Martin boundary of the Pascal's graph is given by the measures of the claim, corresponding to the sequences of i.i.d. Bernoulli random variables. By Theorem 2 the minimal boundary is a subset. Let us show that, in fact, minimal boundary = Martin boundary in this case.

We argue by contradiction. If we are wrong, then

$$\exists p \in [0,1], \text{ such that } B_p = \int_{q \in [0,1] \setminus \{p\}} B_q \, \nu(dq) \qquad (*)$$

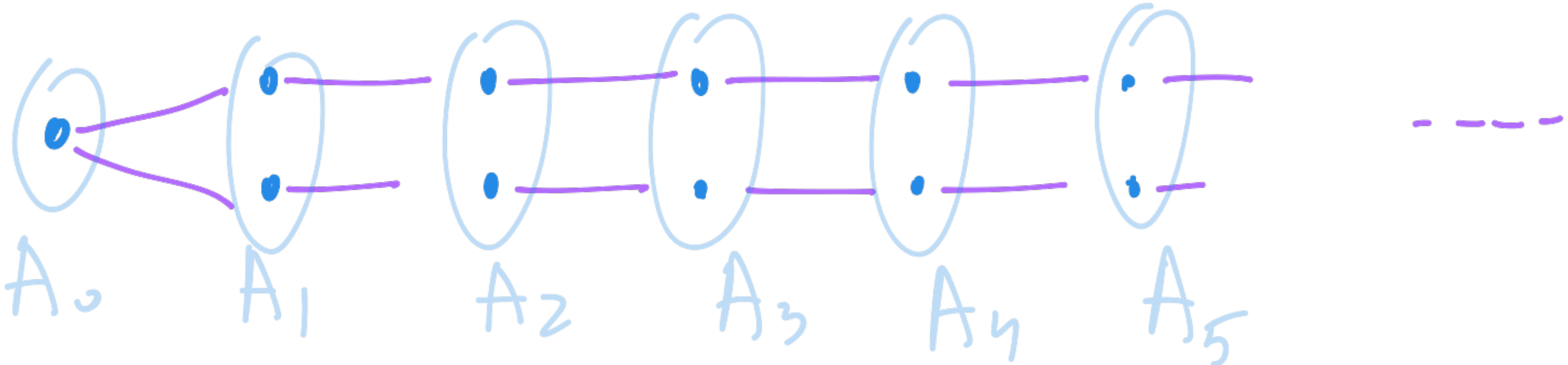
law of i.i.d. Bernoulli-p sequence of 0,1

law of i.i.d. Bernoulli-q sequence of 0,1

measure on minimal boundary from Theorem 1.

But for B_p-distributed $(\xi_1, \xi_2, \dots)$ $\lim_{n\to\infty} \frac{\xi_1 + \dots + \xi_n}{n} = p$ by strong LLN. Which fails by the same LLN for the right-hand side of $(*)$. Contradiction. ■

Exercise 2. Consider the following graph



Find its minimal boundary.