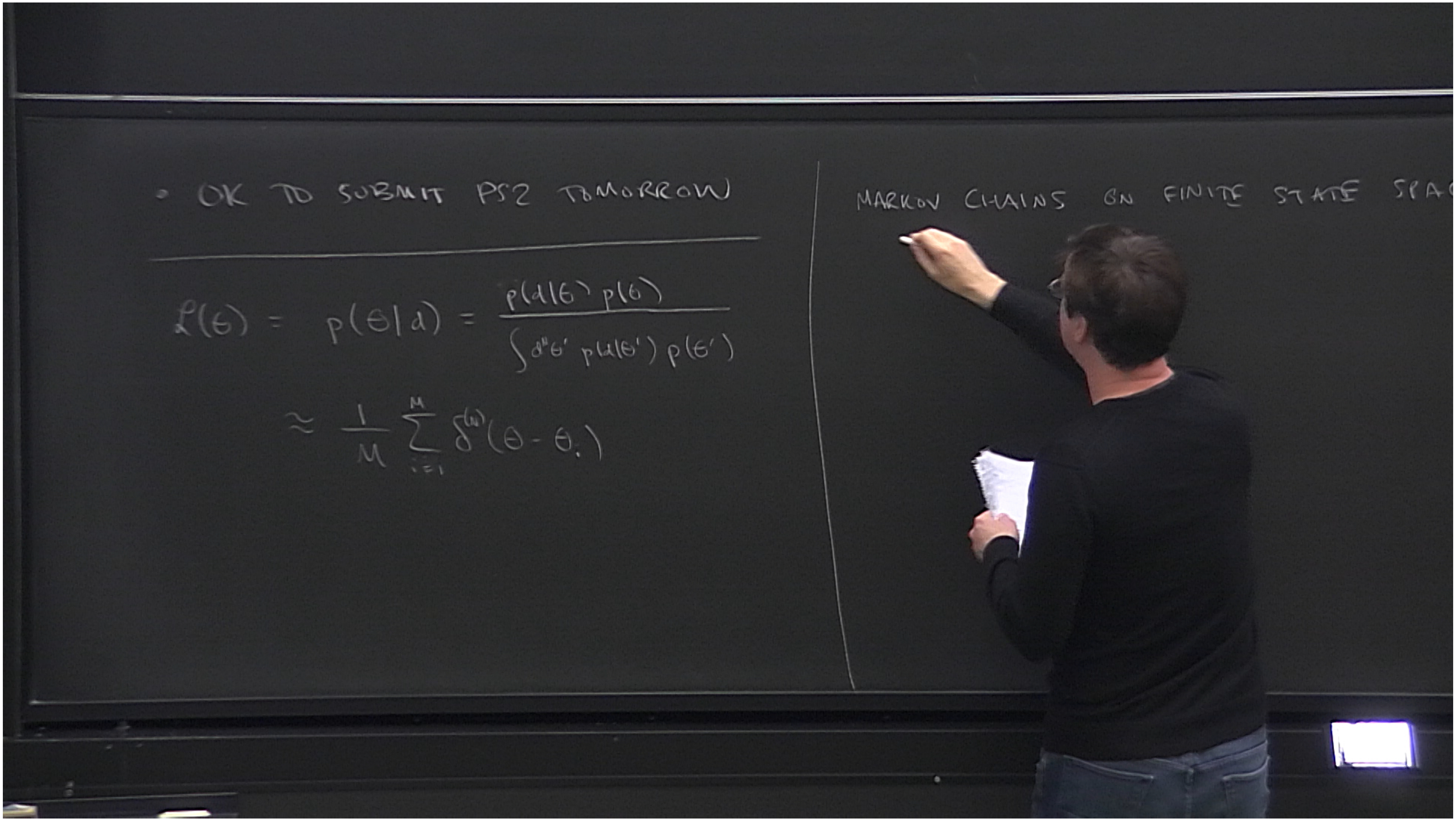
Title: PSI 2018/2019 - Explorations in Cosmology - Lecture 11

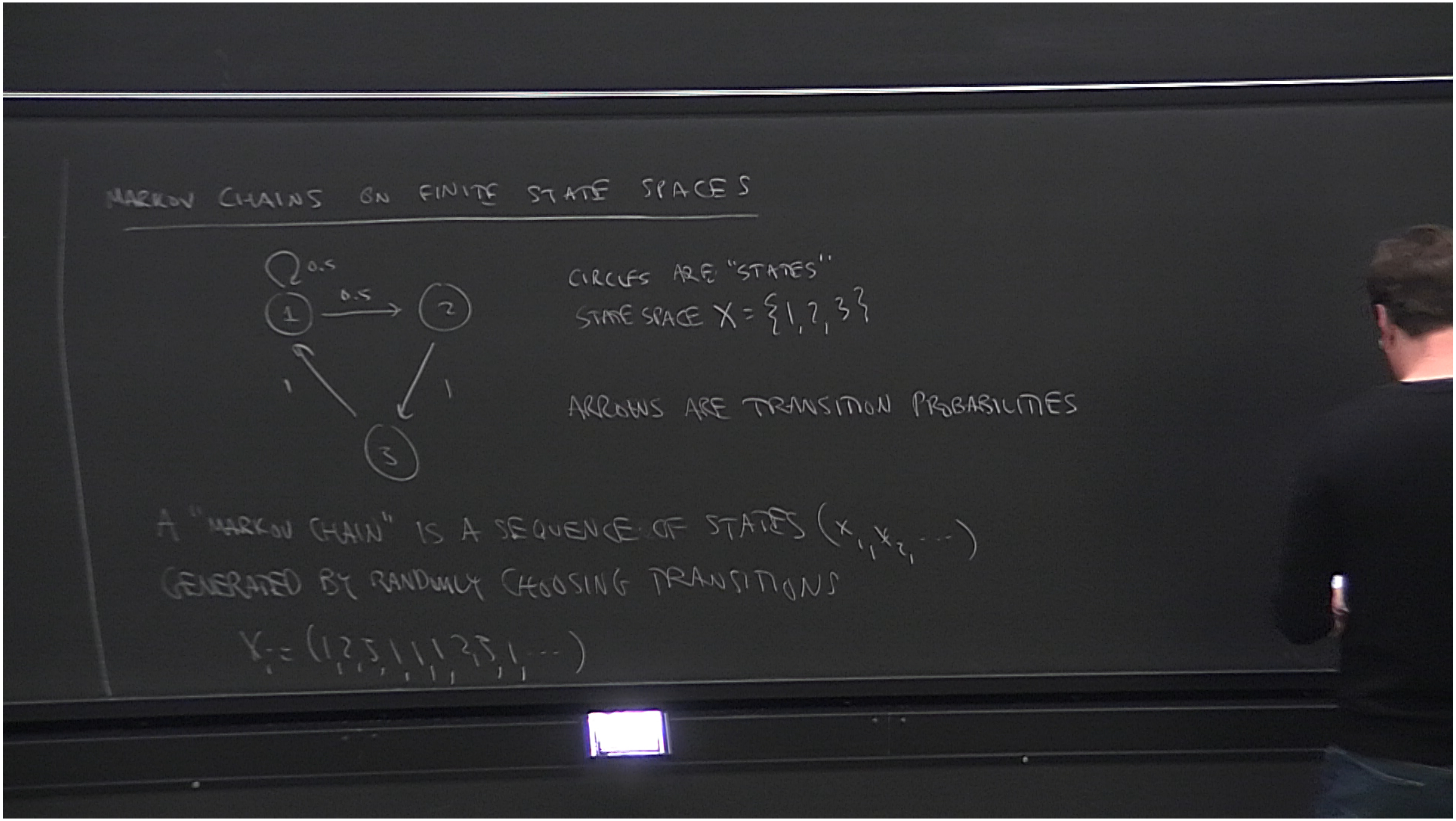
Speakers: Kendrick Smith

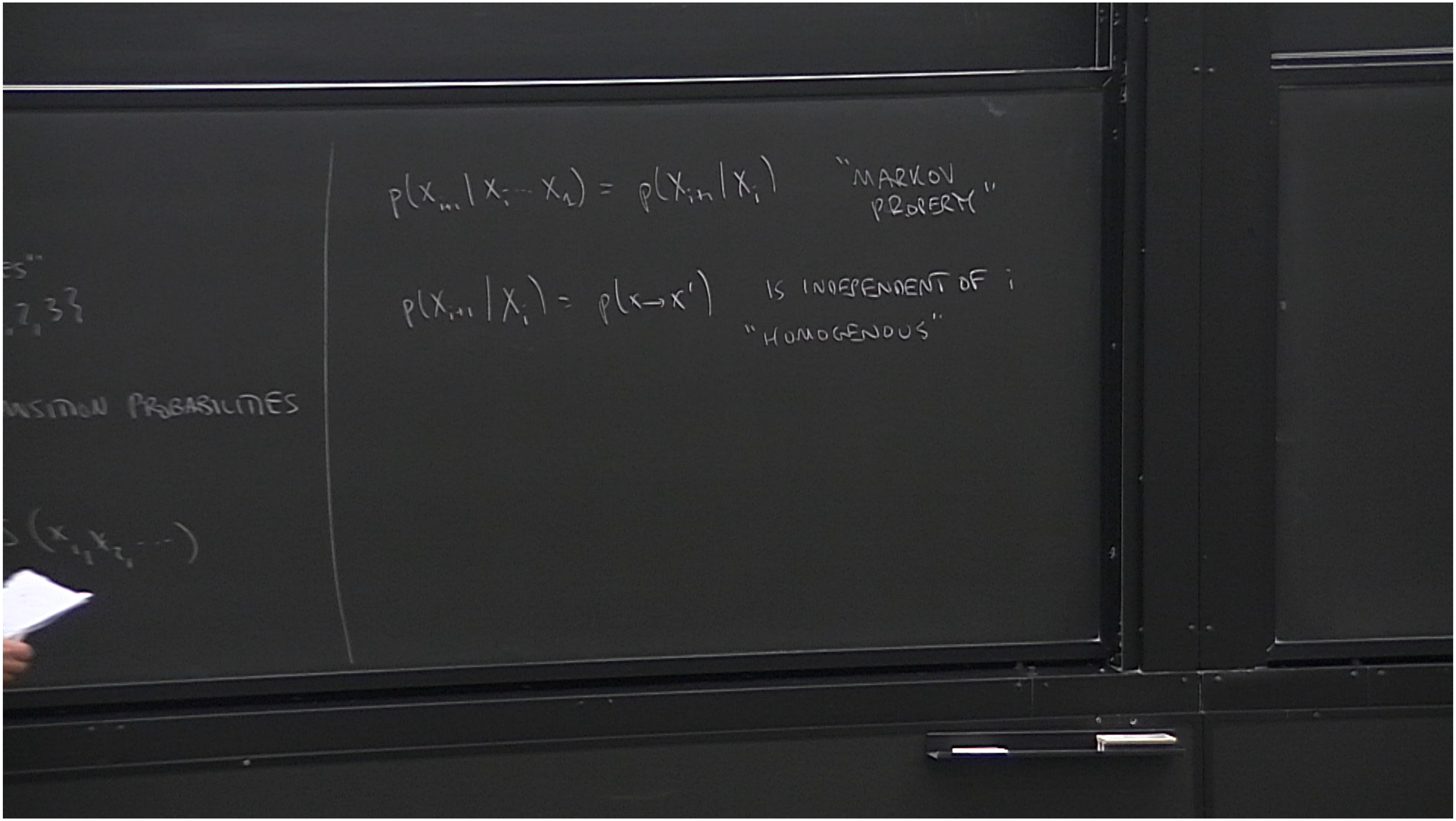
Collection: PSI 2018/2019 - Explorations in Cosmology (Smith)

Date: May 01, 2019 - 11:30 AM

URL: http://pirsa.org/19050005


• OK TO SUBMIT PS2 TOMORROW
$\mathcal{L}(\theta) = p(\theta|d) = \frac{p(d|\theta)\, p(\theta)}{\int d^n\theta'\, p(d|\theta')\, p(\theta')}$
$\approx \frac{1}{M} \sum_{i=1}^{M} \delta^{(n)}(\theta - \theta_i)$
MARKOV CHAINS ON FINITE STATE SPA


MARKOV CHAINS ON FINITE STATE SPACES
0.5
0.5
1
2
1
1
3
CIRCLES ARE "STATES"
STATE SPACE X = {1, 2, 3}
ARROWS ARE TRANSITION PROBABILITIES
A "MARKOV CHAIN" IS A SEQUENCE OF STATES (x1, x2, ...)
GENERATED BY RANDOMLY CHOOSING TRANSITIONS
Xi = (1, 2, 3, 1, 1, 1, 2, 3, 1, ...)


$p(x_{i+1} | x_i \cdots x_1) = p(x_{i+1} | x_i)$
"MARKOV PROPERTY"
$p(x_{i+1} | x_i) = p(x \to x')$
IS INDEPENDENT OF i
"HOMOGENOUS"
PROBABILITIES

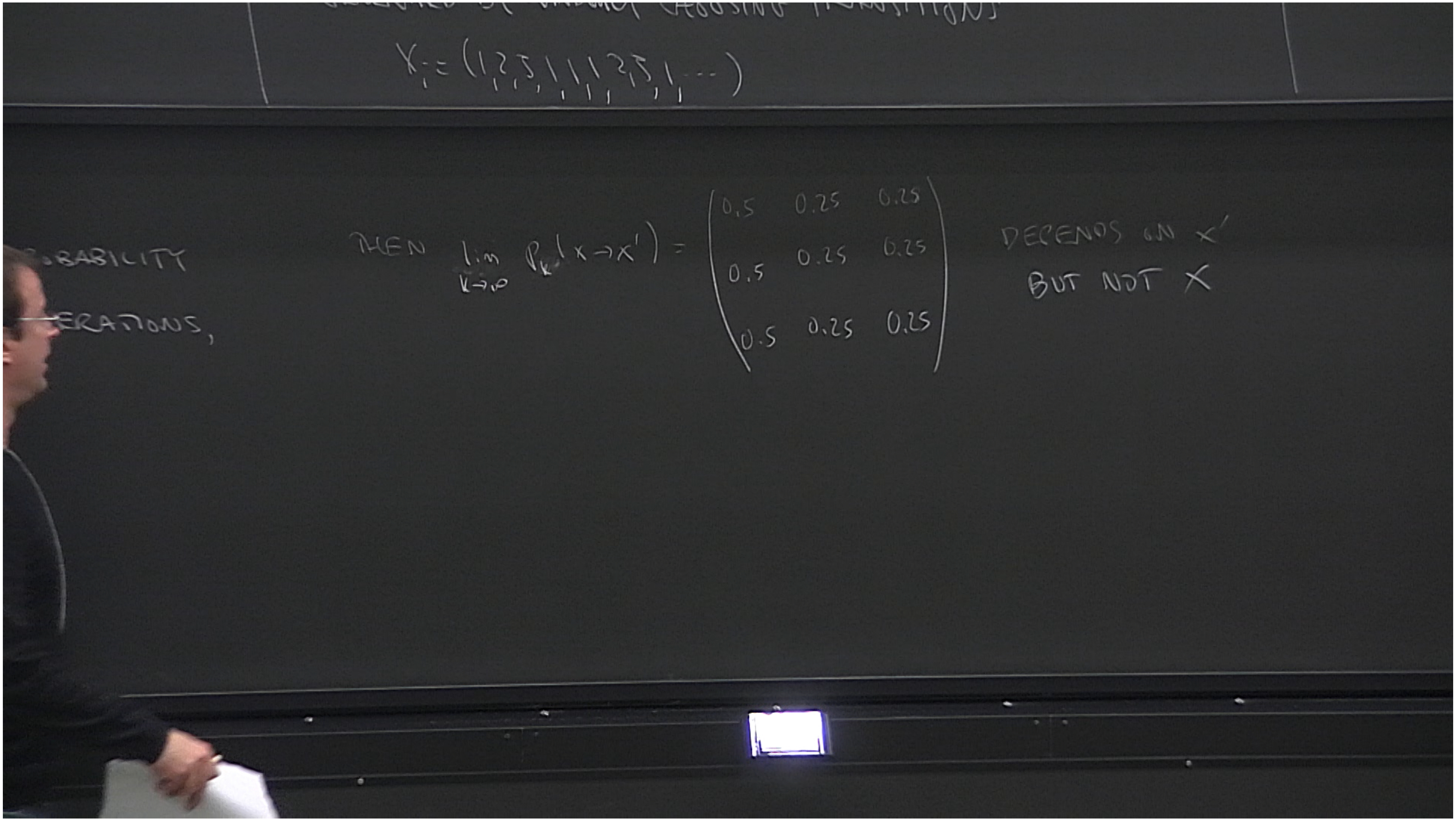
GENERATED BY RANDOMLY CHOOSING

$X_i = (1, 2, 3, 1, 1, 1, 2, 3, 1, \ldots)$

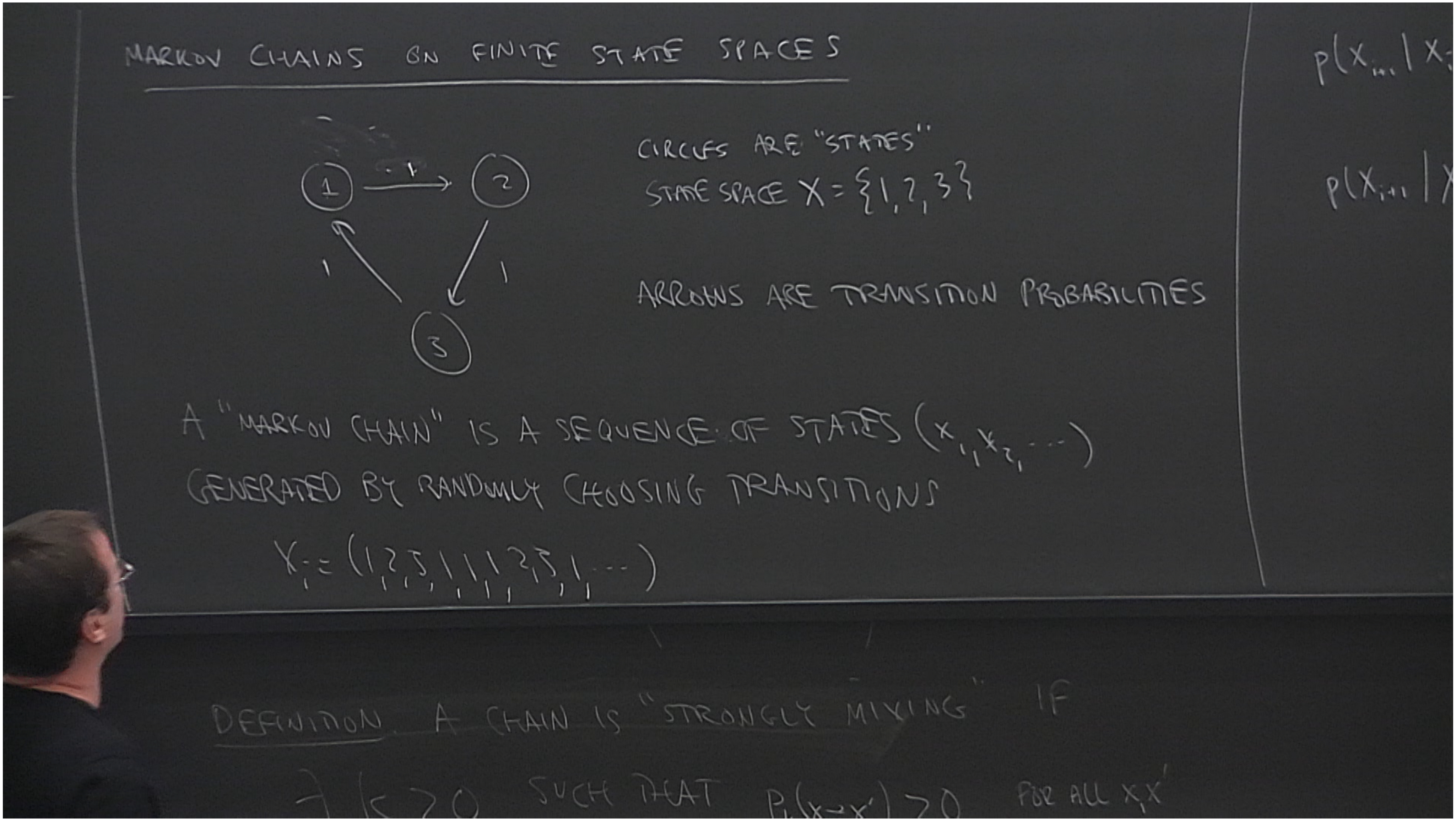
LET $P_k(x \to x')$ BE THE PROBABILITY OF ENDING AT x' AFTER k ITERATIONS, STARTING AT x.

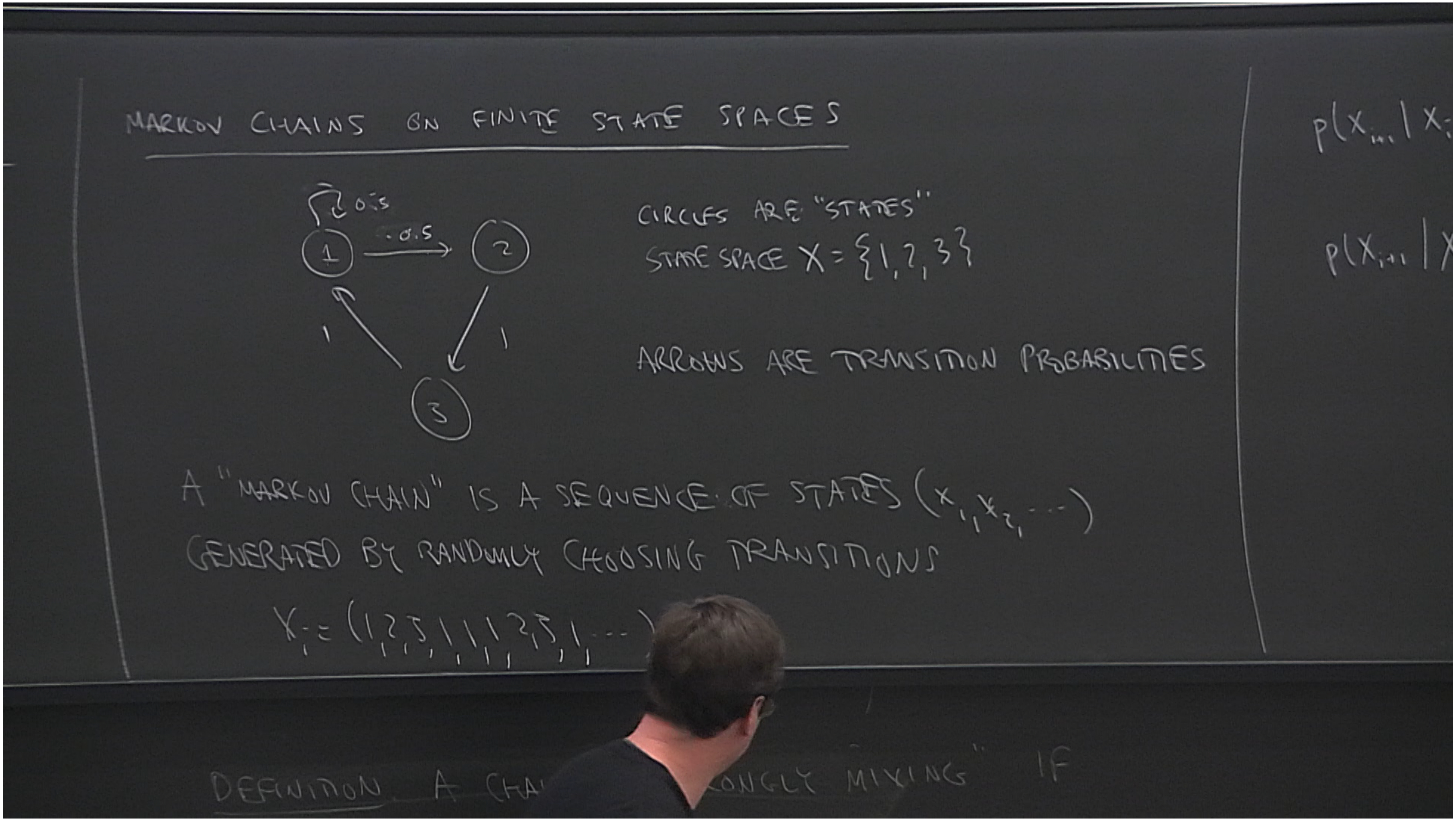
$$P_1(x \to x') = \begin{array}{cc} & x' \\ x & \begin{pmatrix} 0.5 & & \\ 0.5 & 1 & \\ & & 1 \end{pmatrix} \end{array}$$

THEN $\lim_{k \to \infty} P_k(x \to x') = \Big(\quad \Big)$



X_i = (1,2,3,1,1,1,2,3,1, ...)
OBABILITY
ERATIONS,
THEN lim_{k→∞} P_k(x→x') =
0.5 0.25 0.25
0.5 0.25 0.25
0.5 0.25 0.25
DEPENDS ON x'
BUT NOT X


MARKOV CHAINS ON FINITE STATE SPACES
1
2
3
CIRCLES ARE "STATES"
STATE SPACE X = {1, 2, 3}
ARROWS ARE TRANSITION PROBABILITIES
A "MARKOV CHAIN" IS A SEQUENCE OF STATES $(x_1, x_2, \ldots)$
GENERATED BY RANDOMLY CHOOSING TRANSITIONS
$X_i = (1, 2, 3, 1, 1, 1, 2, 3, 1, \ldots)$
DEFINITION. A CHAIN IS "STRONGLY MIXING" IF
$\exists\, k > 0$ SUCH THAT $P_k(x \to x') > 0$ FOR ALL x, x'


MARKOV CHAINS ON FINITE STATE SPACES
0.5
0.5
1
2
1
1
3
CIRCLES ARE "STATES"
STATE SPACE X = {1, 2, 3}
ARROWS ARE TRANSITION PROBABILITIES
A "MARKOV CHAIN" IS A SEQUENCE OF STATES (x1, x2, ...)
GENERATED BY RANDOMLY CHOOSING TRANSITIONS
Xi = (1, 2, 3, 1, 1, 1, 2, 3, 1, ...)
DEFINITION

A "MARKOV CHAIN" IS A SEQUENCE OF STATES $(x_1, x_2, \dots)$
GENERATED BY RANDOMLY CHOOSING TRANSITIONS

$x_i = (1, 2, 3, 1, 1, 1, 2, 3, 1, \dots)$

TY

NS,

THEN $\lim_{k \to \infty} P_k(x \to x') = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.25 \end{pmatrix}$ DEPENDS ON x' BUT NOT x

<u>DEFINITION</u>. A CHAIN IS "STRONGLY MIXING" IF

$\exists\, k > 0$ SUCH THAT $P_k(x \to x') > 0$ FOR ALL x, x'

FUNDAMENTAL THEOREM.

IF A CHAIN IS STRONGLY MIXING, THEN

$$\lim_{k \to \infty} p_k(x \to x') = p_\infty(x')$$

"LIMITING DISTRIBUTION" OF THE CHAIN

EXISTS AND IS INDEPENDENT OF x.

PROOF OMITTED

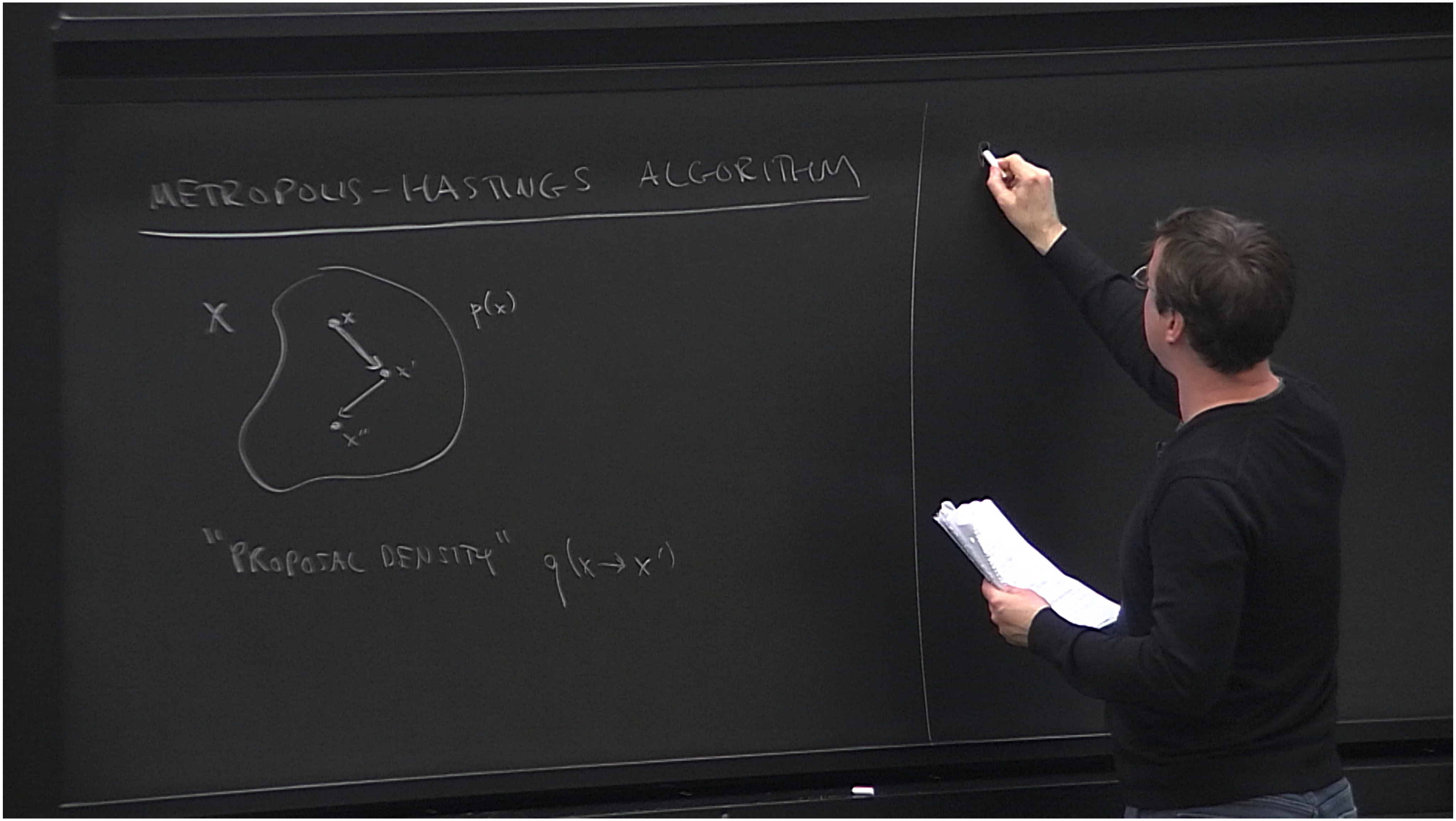
THE LIMITING DISTRIBUTION SATISFIES:

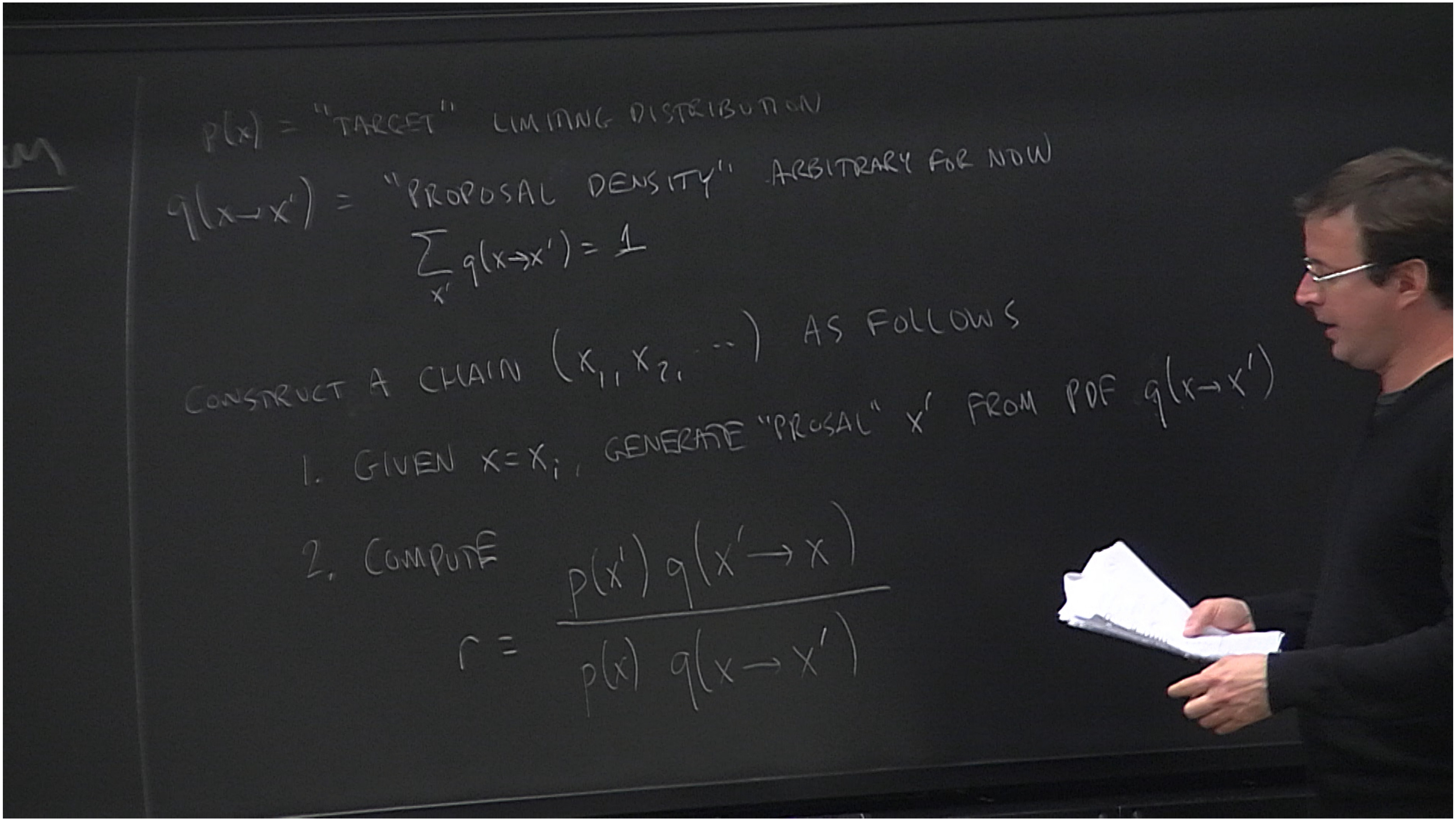
$$\sum_{x} p_\infty(x)\, p(x \to x') = p_\infty(x')$$

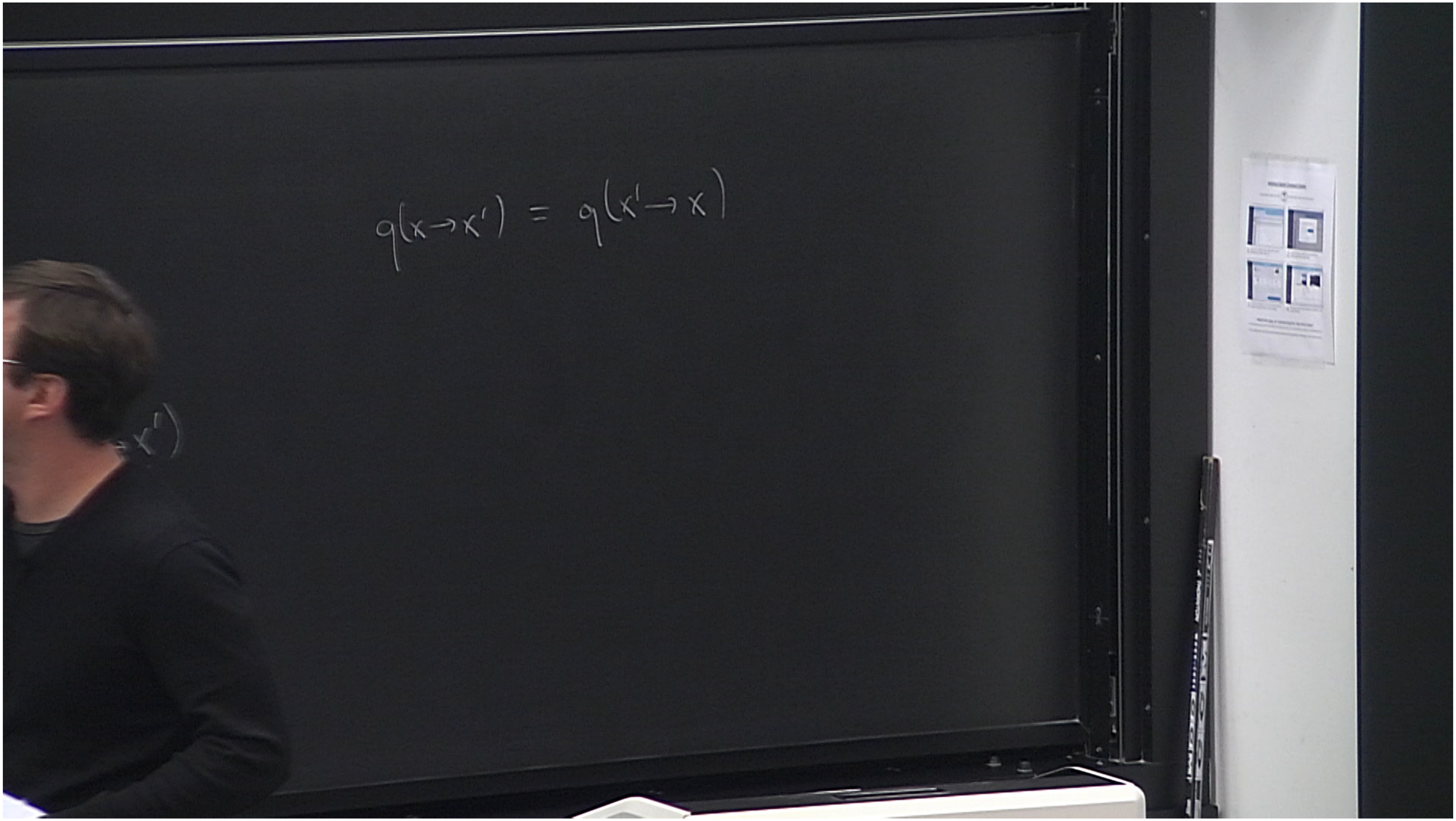
"DETAILED BALANCE CONDITION"

MCMC

IDEA: GIVEN A "TARGET" DISTRIBUTION $p(x)$, CONSTRUCT A SET OF TRANSITION PROBABILITIES $p(x \to x')$ SUCH THAT $p_\infty(x') = p(x)$


METROPOLIS-HASTINGS ALGORITHM
X
p(x)
x
x'
"x''"
"PROPOSAL DENSITY" q(x → x')


p(x) = "TARGET" LIMITING DISTRIBUTION
q(x→x') = "PROPOSAL DENSITY" ARBITRARY FOR NOW
$\sum_{x'} q(x \to x') = 1$
CONSTRUCT A CHAIN (x1, x2, ...) AS FOLLOWS
1. GIVEN x=xi, GENERATE "PROSAL" x' FROM PDF q(x→x')
2. COMPUTE
$r = \frac{p(x')\,q(x' \to x)}{p(x)\,q(x \to x')}$


$q(x \to x') = q(x' \to x)$

3. LET u BE A UNIFORM RANDOM ON $(0,1)$

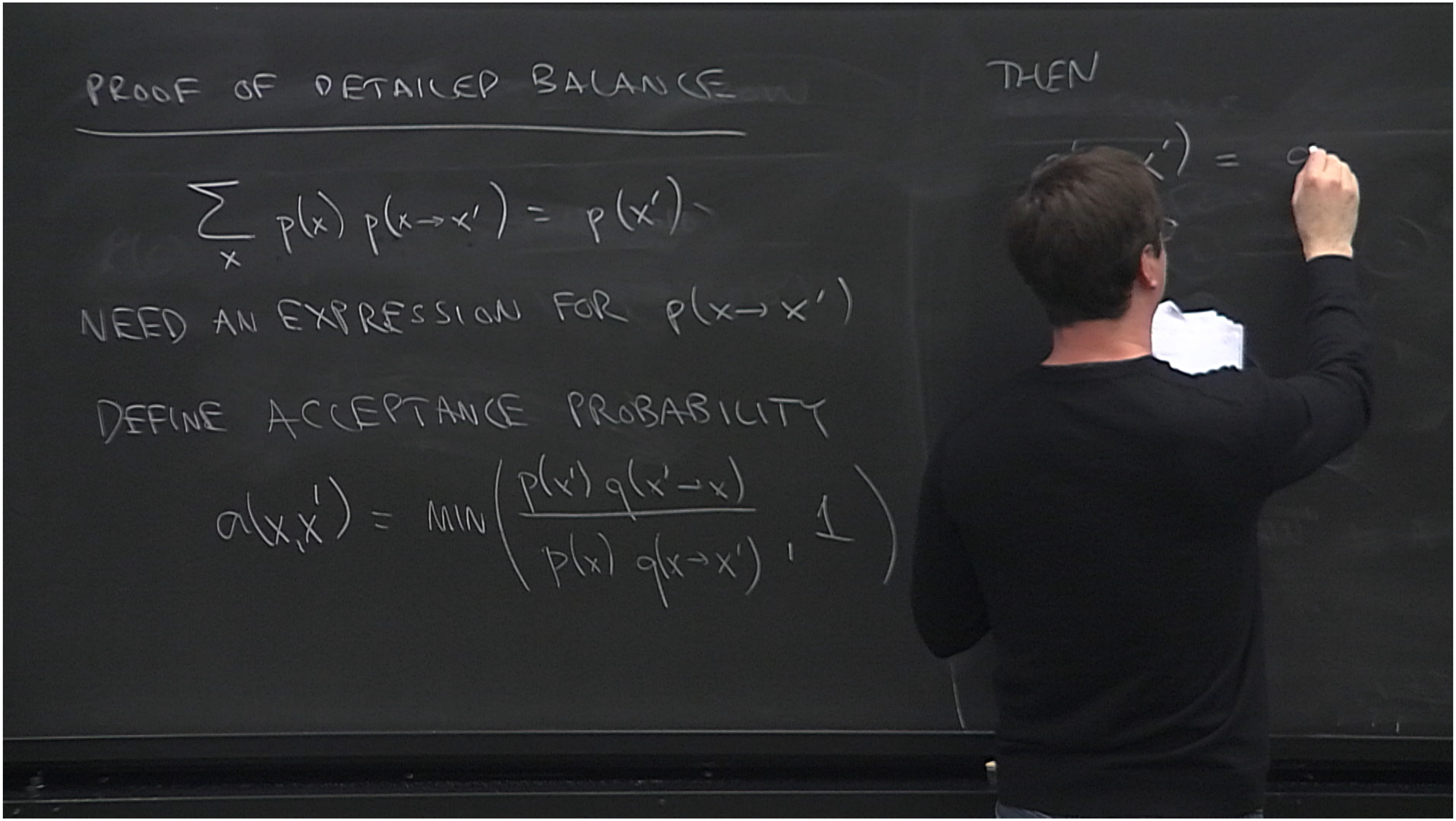
$$\text{LET } X_{i+1} = \begin{cases} X' & \text{IF } u < r \quad [\text{PROPOSAL "ACCEPTED"}] \\ X & \text{IF } u > r \quad [\text{PROPROSAL "REJECTED"}] \end{cases}$$

[IN PARTICULAR, ALWAYS ACCEPT IF $r > 1$]

PROOF OF DETAILED BALANCE

$$\sum_x p(x)\, p(x \to x') = p(x')$$

NEED AN EXPRESSION FOR $p(x \to x')$


PROOF OF DETAILED BALANCE
$\sum_x p(x)\, p(x \to x') = p(x')$
NEED AN EXPRESSION FOR $p(x \to x')$
DEFINE ACCEPTANCE PROBABILITY
$a(x,x') = \mathrm{MIN}\left(\frac{p(x')\, q(x' \to x)}{p(x)\, q(x \to x')}, 1 \right)$
THEN

THEN

$$p(x \to x') = q(x \to x')\, a(x, x') + \delta_{xx'} \sum_{x''} q(x \to x'') \left(1 - a(x, x'')\right)$$

$$= -\int_{-\infty}^{\infty} dt \int_{-\infty}^{t} dt' \, \hat{H}_I(t)\, \hat{H}_I(t')$$

$$\hat{\rho}_{d,0} = \rho^{(1)}_{d,T} + \rho^{(2)}_{d,t} + O(\lambda_u^3)$$

$\rho_{d,T} \sim \lambda_A$ NO SIG

THEN

$$p(x \to x') = q(x \to x')\, a(x, x') + \delta_{xx'} \sum_{x''} q(x \to x'') \left(1 - a(x, x'')\right)$$

NOW

$$\sum_x p(x)\, p(x \to x') = \sum_x p(x)\, q(x \to x') \min\left(\frac{p(x')\, q(x' \to x)}{p(x)\, q(x \to x')}, 1 \right)$$

$$+ \, p(x') \sum_{x''} q(x' \to x'') \left[1 - \min\left(\frac{p(x'')\, q(x'' \to x')}{p(x')\, q(x' \to x'')}, 1 \right) \right]$$

$$=$$

(0.5 0.25 0.25)

$$= \sum_{x} \mathrm{MIN}\Big(p(x')\, q(x' \to x),\ p(x)\, q(x \to x') \Big)$$

$$+\ p(x') \sum_{x''} q(x' \to x'')$$

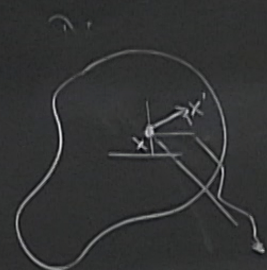
$$- \sum_{x''} \mathrm{MIN}\Big(p(x'')\, q(x'' \to x'),\ p(x')\, q(x' \to x'') \Big)$$

$$= p(x')\ \checkmark$$

$q(x \to x')$, 1)

PROBABILITY

ITERATIONS,



$$= \sum_{x} \mathrm{MIN}\left(p(x')\, q(x' \to x),\ p(x)\, q(x \to x') \right)$$

$$+\ p(x') \sum_{x''} q(x' \to x'')$$

$$-\ \sum_{x''} \mathrm{MIN}\left(p(x'')\, q(x'' \to x'),\ p(x')\, q(x' \to x'') \right)$$

$$= p(x') \quad \checkmark$$

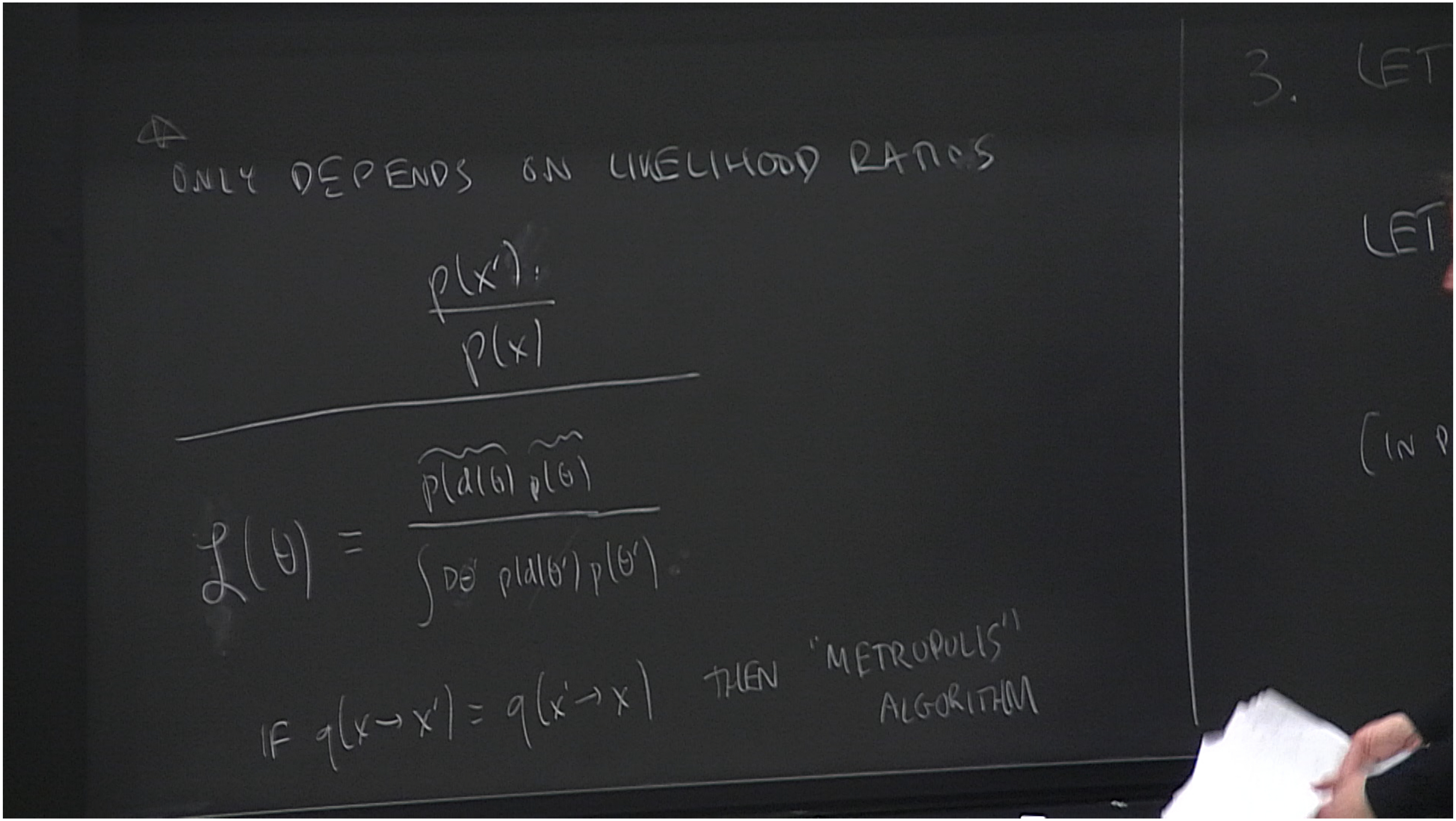
ONLY DEPENDS ON LIKELIHOOD RATIOS

$$\frac{p(x')}{p(x)}$$

3. LET

LET

(IN P



ONLY DEPENDS ON LIKELIHOOD RATIOS
$\dfrac{p(x')}{p(x)}$
$\mathcal{L}(\theta) = \dfrac{p(d|\theta)\, p(\theta)}{\int D\theta'\, p(d|\theta')\, p(\theta')}$
IF $q(x \to x') = q(x' \to x)$ THEN "METROPOLIS" ALGORITHM
3. LET
LET
(IN P

OFTEN $q(x \to x') \propto \exp\left[-\frac{1}{2}(x-x')^T C_q^{-1}(x-x')\right]$

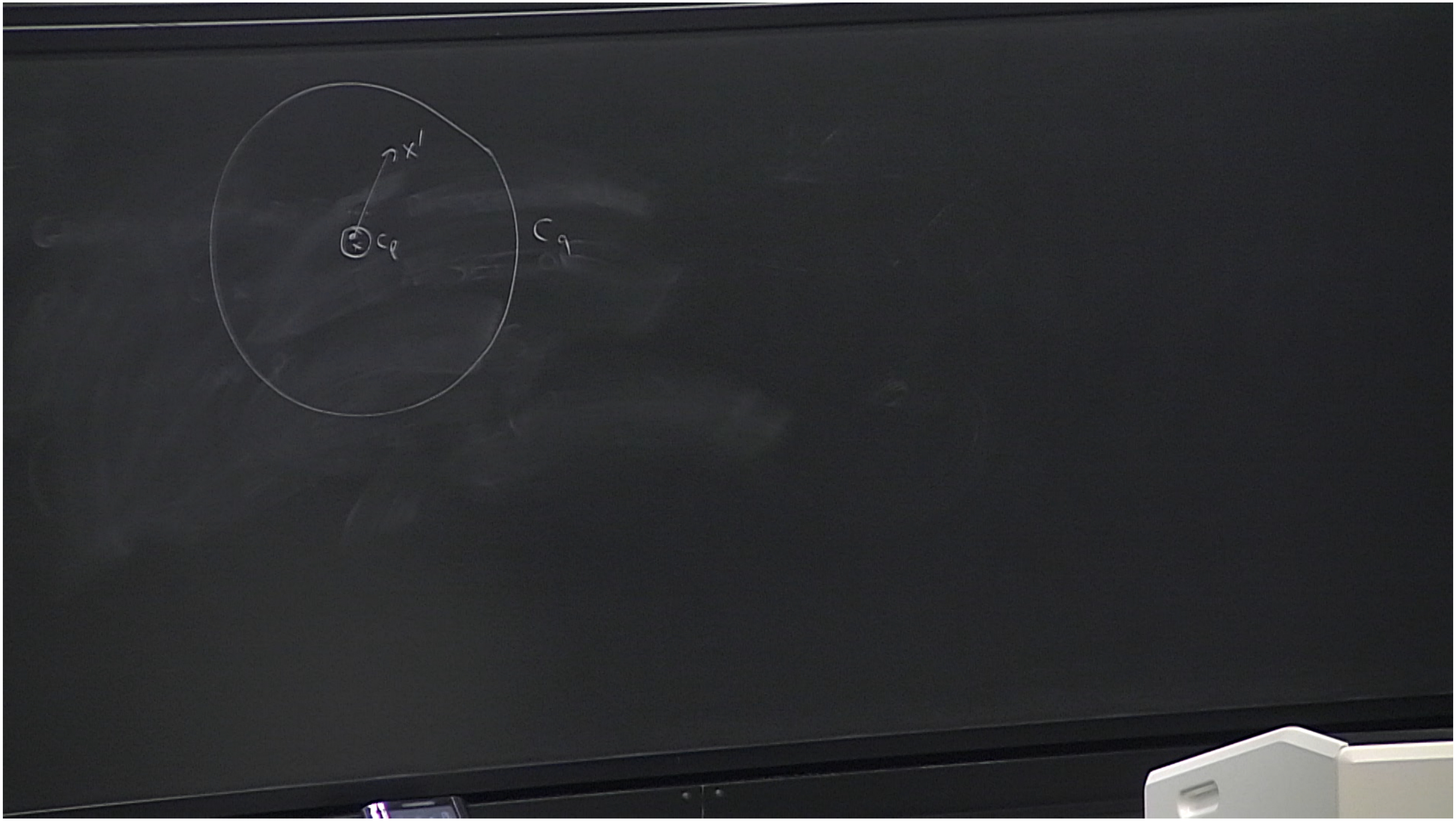
TOY EXAMPLE: $p(x) \propto \exp\left[-\frac{1}{2} x^T C_p^{-1} x\right]$

OFTEN $q(x \to x') \propto \exp\left[-\frac{1}{2}(x-x')^T C_q^{-1}(x-x')\right]$

TOY EXAMPLE: $p(x) \propto \exp\left[-\frac{1}{2} x^T C_p^{-1} x\right]$

EXAMPLE 1) $C_p = \sigma_p^2 \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$ $C_q = \sigma_q^2 \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$

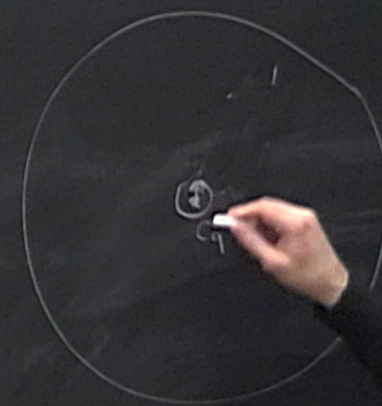
$\sigma_q \gg \sigma_p$ IS BAD: REJECTION RATE IS HIGH


x'
c_p
C_q

OFTEN $q(x \to x') \propto \exp\left[-\frac{1}{2}(x-x')^T C_q^{-1}(x-x')\right]$

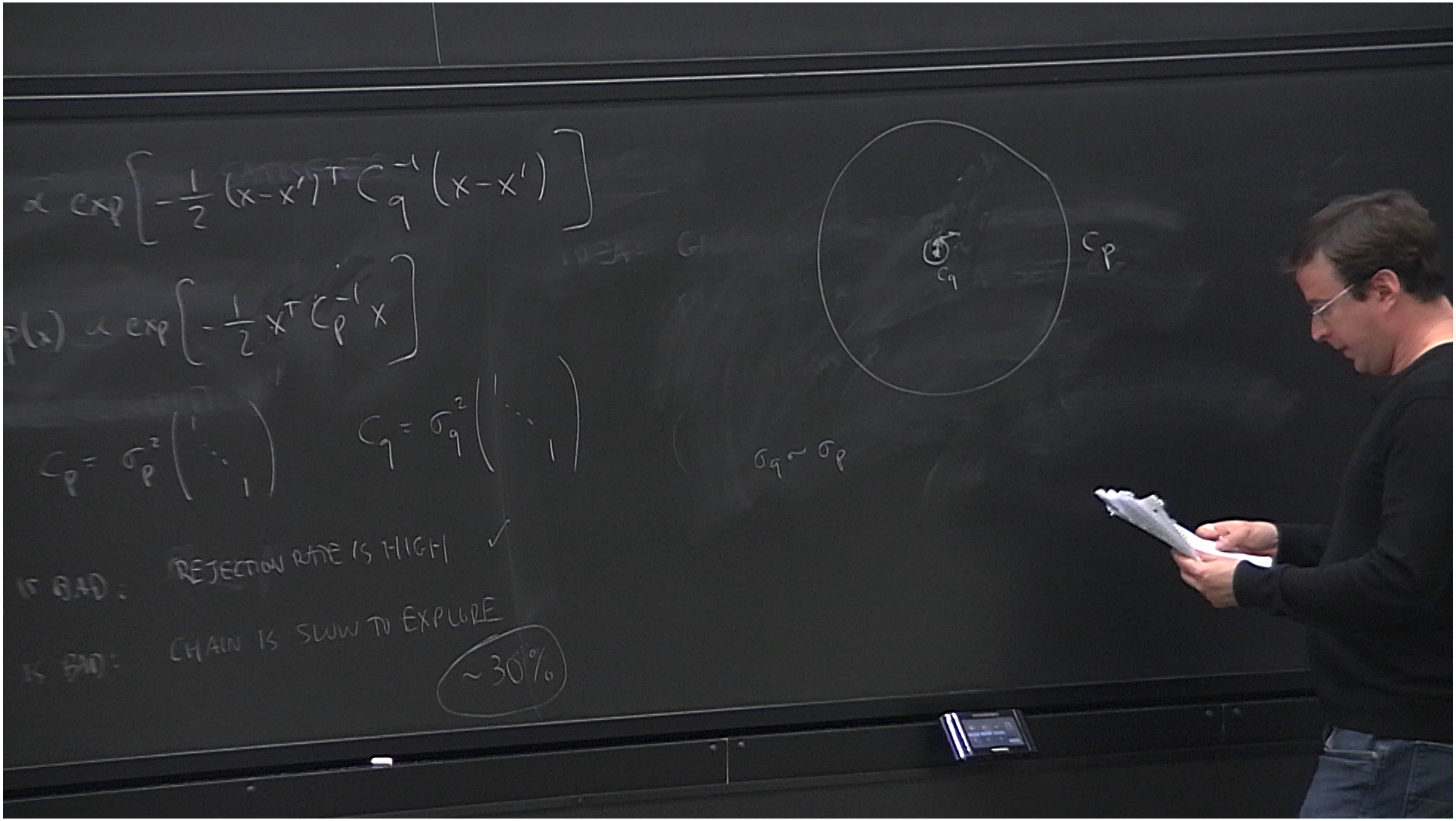
TOY EXAMPLE: $p(x) \propto \exp\left[-\frac{1}{2} x^T C_p^{-1} x\right]$

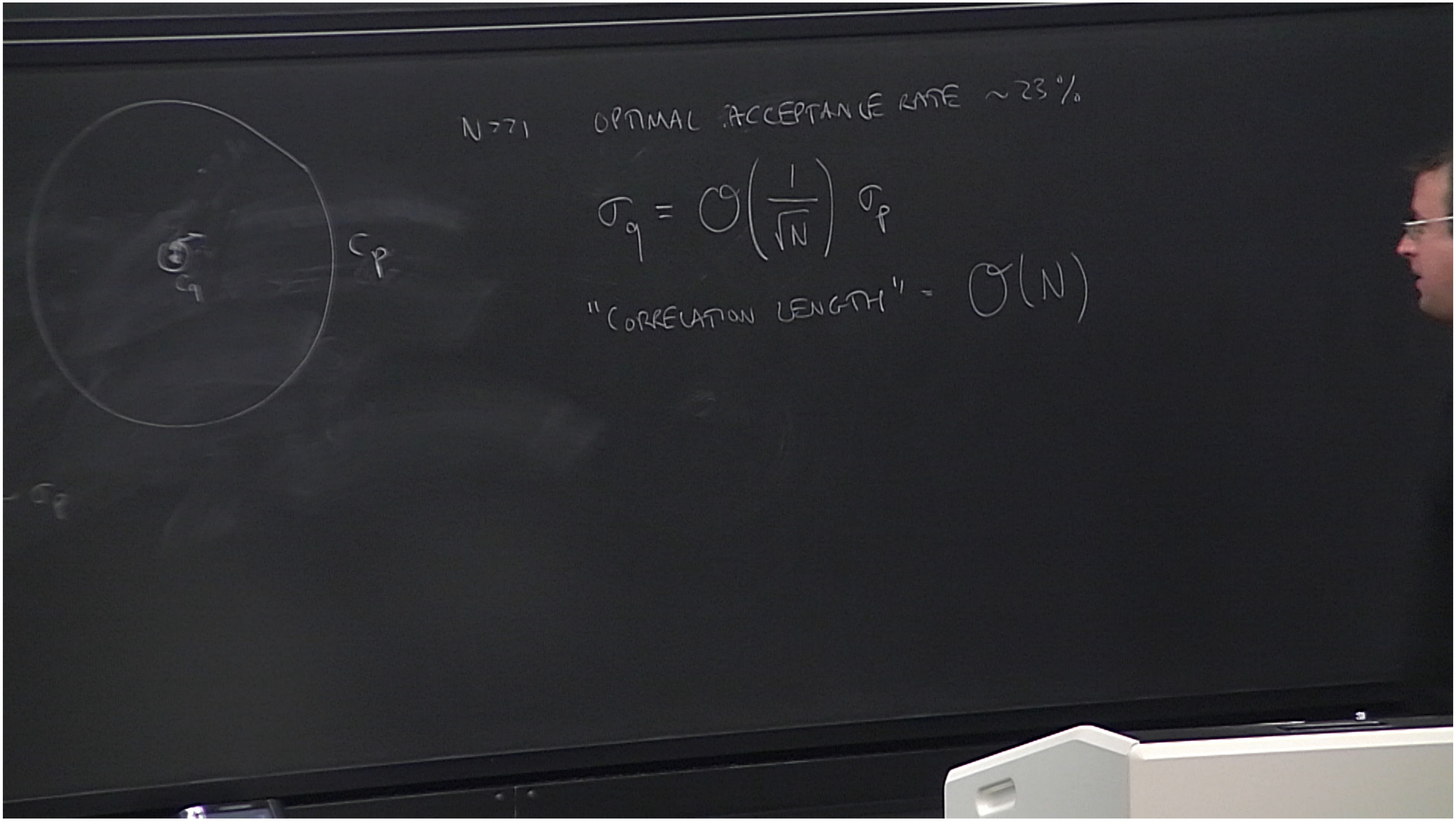
EXAMPLE 1) $C_p = \sigma_p^2 \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$ $C_q = \sigma_q^2 \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$

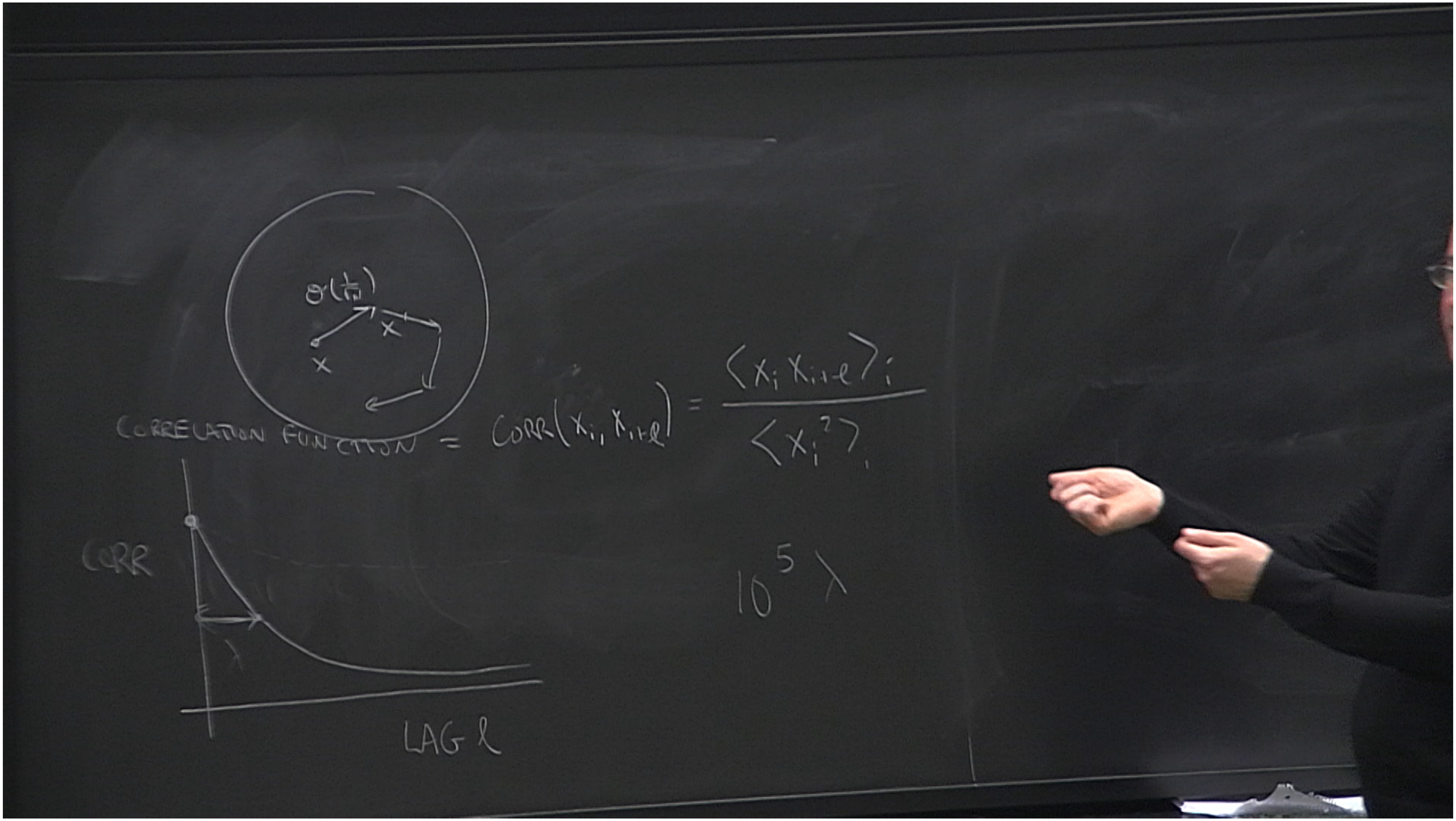


$\sigma_q \gg \sigma_p$ IS BAD: REJECTION RATE IS HIGH

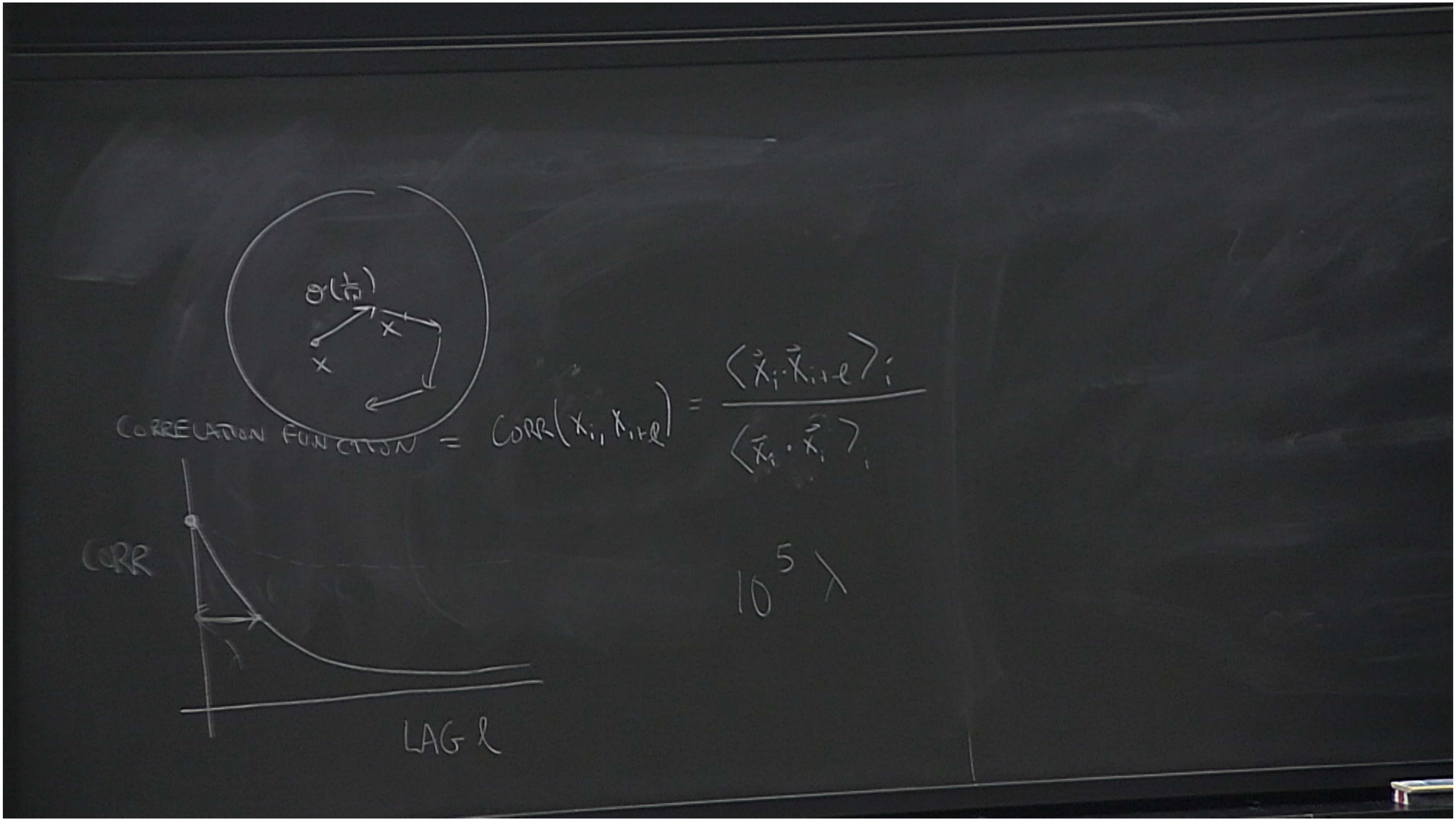
$\sigma_q \ll \sigma_p$ IS BAD: CHAIN IS SLOW TO EXPLORE

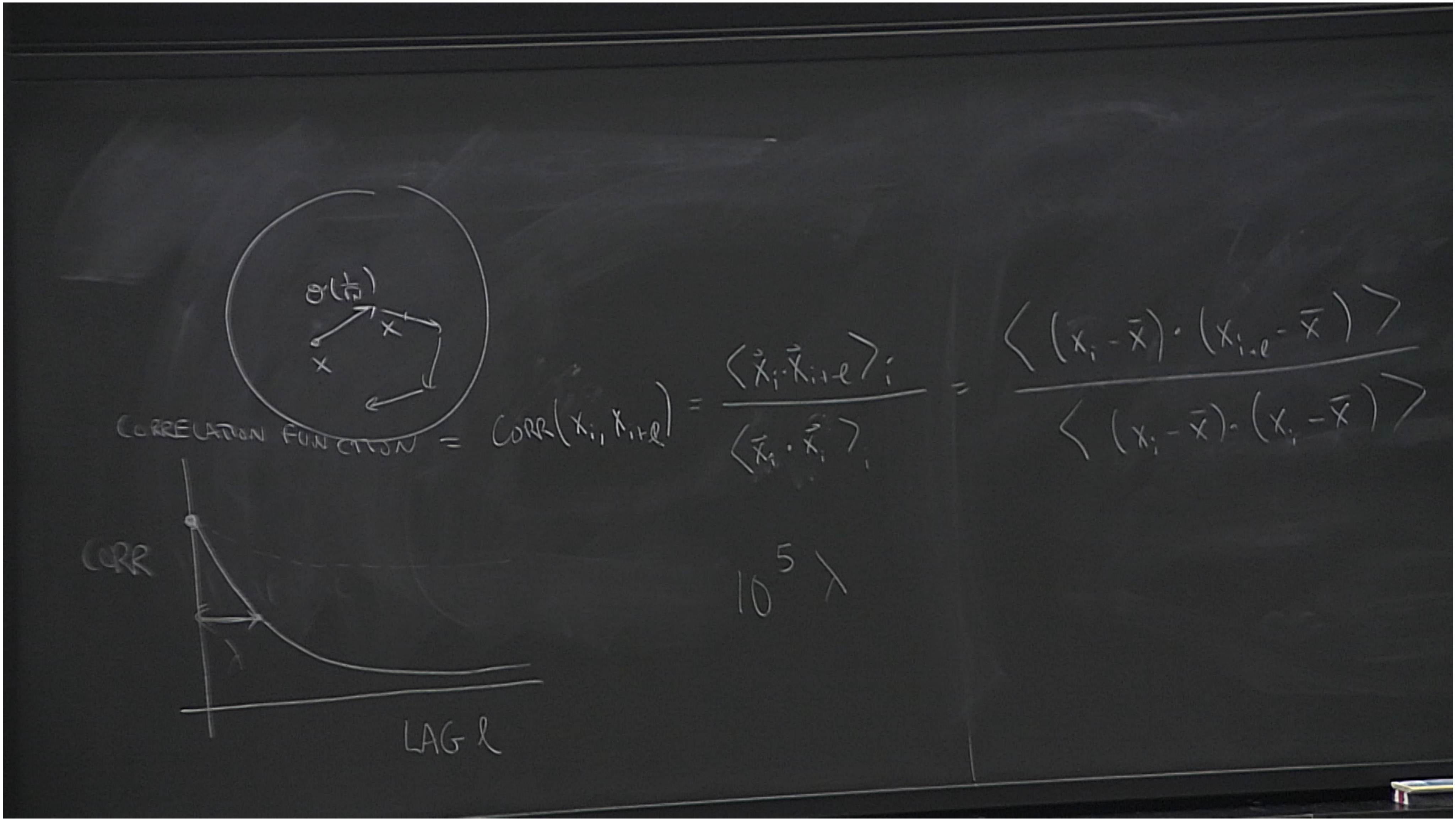

$\propto \exp\left[-\frac{1}{2}(x-x')^T C_q^{-1}(x-x')\right]$
$p(x) \propto \exp\left[-\frac{1}{2}x^T C_p^{-1} x\right]$
$C_p = \sigma_p^2 \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$
$C_q = \sigma_q^2 \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$
BAD: REJECTION RATE IS HIGH
BAD: CHAIN IS SLOW TO EXPLORE
~30%
C_p
C_q
$\sigma_q \sim \sigma_p$


N >> 1 OPTIMAL ACCEPTANCE RATE ~ 23%
$\sigma_q = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \sigma_p$
"CORRELATION LENGTH" = $\mathcal{O}(N)$
c_p
c_q



O(1/n)
CORRELATION FUNCTION = CORR(x_i, x_{i+l}) = ⟨x_i x_{i+l}⟩_i / ⟨x_i^2⟩_i
CORR
λ
LAG l
10^5 λ


CORRELATION FUNCTION = CORR(x_i, x_i+l) = ⟨x_i · x_i+l⟩_i / ⟨x_i · x_i⟩_i
CORR
LAG l
10^5 λ


$\theta(\frac{1}{n})$
CORRELATION FUNCTION = $\mathrm{CORR}(x_i, x_{i+l}) = \frac{\langle \vec{x}_i \cdot \vec{x}_{i+l} \rangle_i}{\langle \vec{x}_i \cdot \vec{x}_i \rangle_i} = \frac{\langle (x_i - \bar{x}) \cdot (x_{i+l} - \bar{x}) \rangle}{\langle (x_i - \bar{x}) \cdot (x_i - \bar{x}) \rangle}$
CORR
λ
LAG l
$10^5 \lambda$