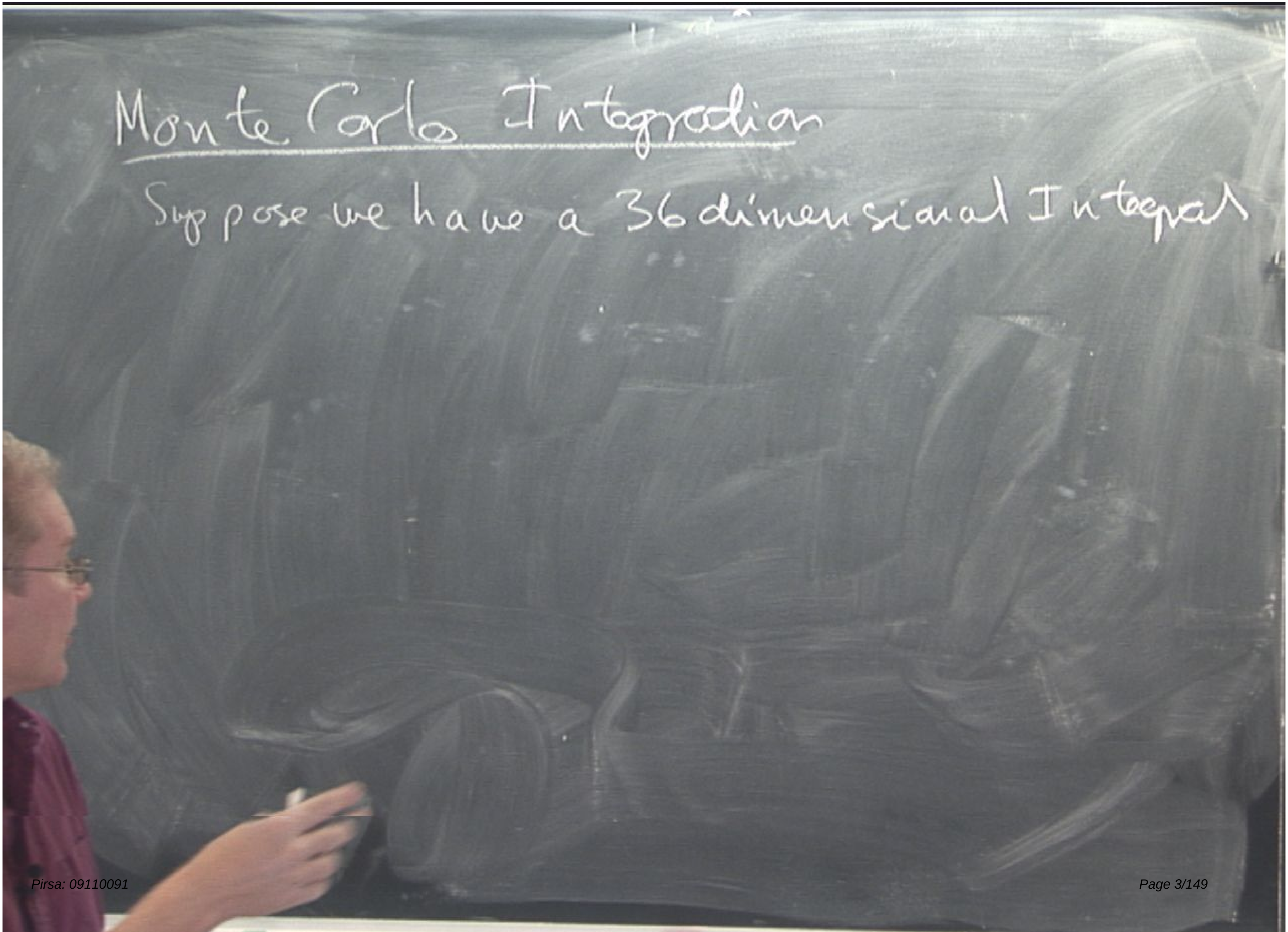
Title: Scientific Computation (PHYS 608) - Lecture 14

Date: Nov 12, 2009 10:30 AM

URL: http://pirsa.org/09110091

Abstract:


Monte Carlo Integration



Monte Carlo Integration
Suppose we have a 36 dimensional Integral

Monte Carlo Integration

Suppose we have a 36 dimensional Integral

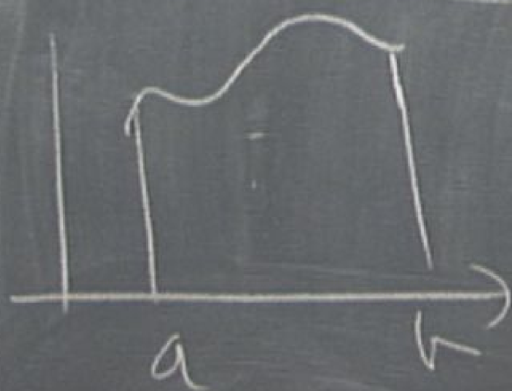
1000 per axis

Monte Carlo Integration

Suppose we have a 36 dimensional Integral

1000 per axis

Mean Value Theorem



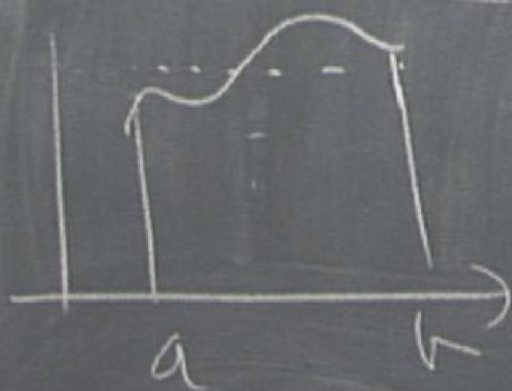
$$I = \int_a^b f(x)\,dx$$

Monte Carlo Integration

Suppose we have a 36 dimensional Integral

1000 per axis

Mean Value Theorem



$$I = \int_a^b f(x)\,dx$$

$$= (b-a)\langle f \rangle$$

So, if we can estimate $\langle f \rangle$ then

$$I \simeq (b-a) \langle f \rangle$$

Simplest approach

So, if we can estimate $\langle f \rangle$ then

$$I \simeq (b-a) \langle f \rangle$$

Simplest approach

$$\langle f \rangle = \sum_{i=1}^{M} f(x_i)$$

x_i are uniform (a,b)

So, if we can estimate $\langle f \rangle$ then

$$I \simeq (b-a) \langle f \rangle$$

Simplest approach

$$\langle f \rangle = \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

x_i as uniform (a,b)

$$\sigma(\langle f \rangle) = \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{N-1}}$$

general expression

$$\int f \, dV$$

So, if we can estimate $\langle f \rangle$ then

$$I \simeq (b-a)\langle f \rangle$$

Simplest approach

$$\langle f \rangle = \frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

x_i as uniform (a,b)

$$\sigma(\langle f \rangle) = \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{N-1}}$$

general expression

$$\int f\,dV = V\langle f \rangle \pm V\sigma(\langle f \rangle)$$

So, if we can estimate $\langle f \rangle$ then

$$I \simeq (b-a) \langle f \rangle$$

Simplest approach

$$\langle f \rangle = \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

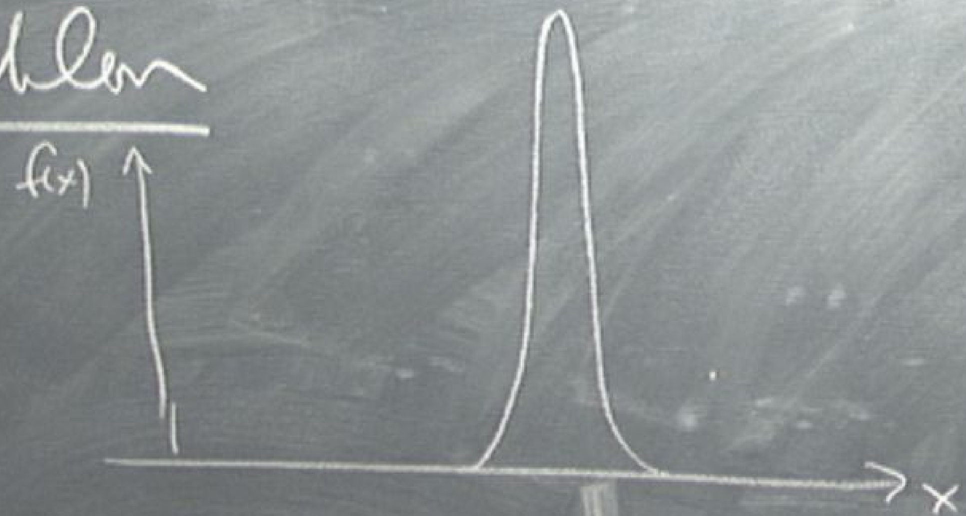
x_i ap uniform (a,b)

$$\sigma(\langle f \rangle) = \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{N-1}}$$

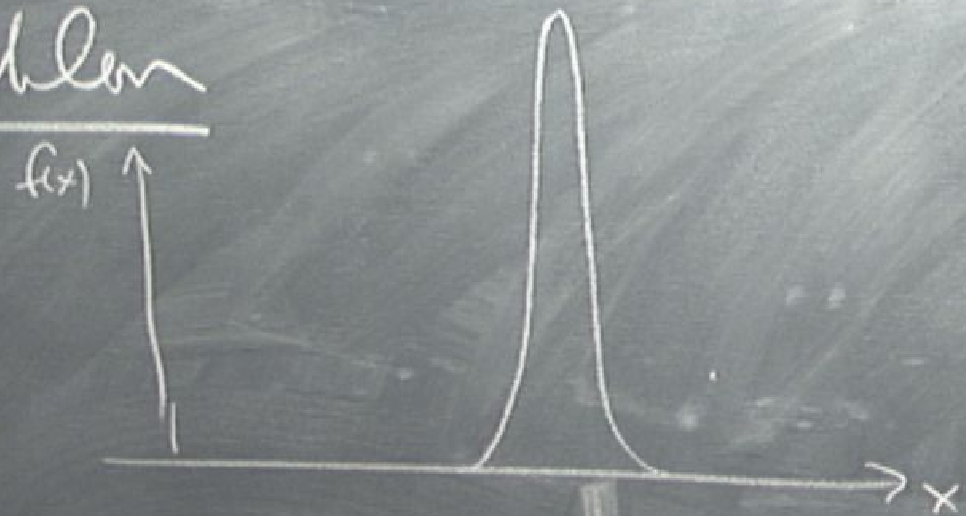
general expression

$$\int f \, dV = V \langle f \rangle \pm V \sigma(\langle f \rangle)$$

Problem

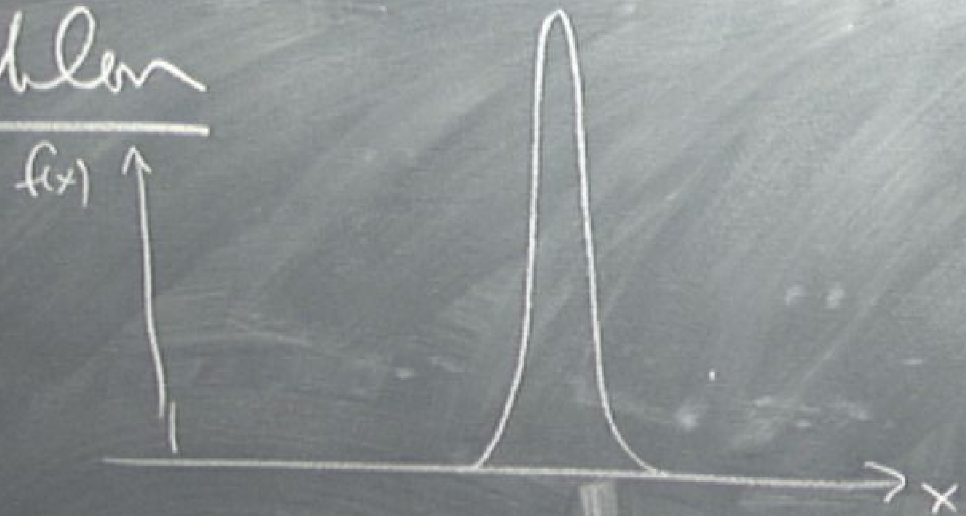


Problem



Importance Sampling : Choose the x_i with a non-uniform distribution

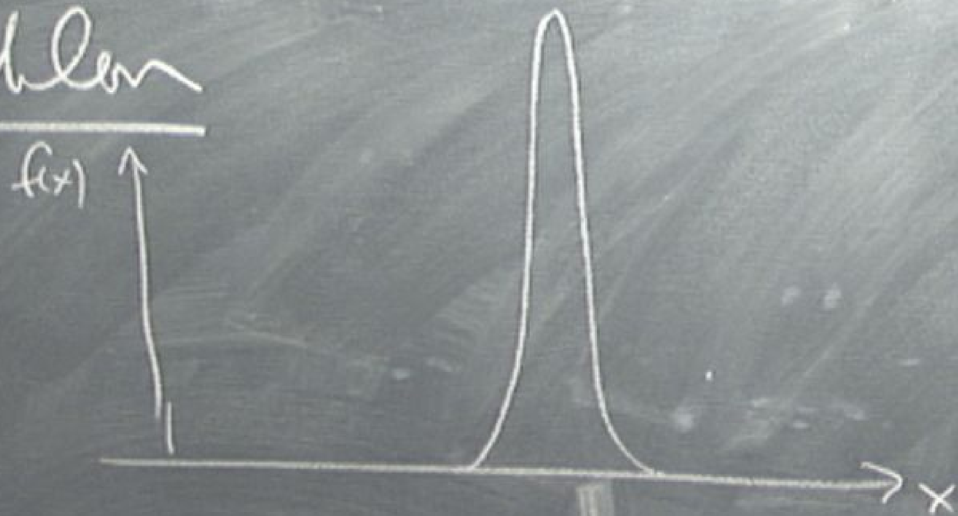
Problem



Importance Sampling: Choose the x_i with a non-uniform distribution

If

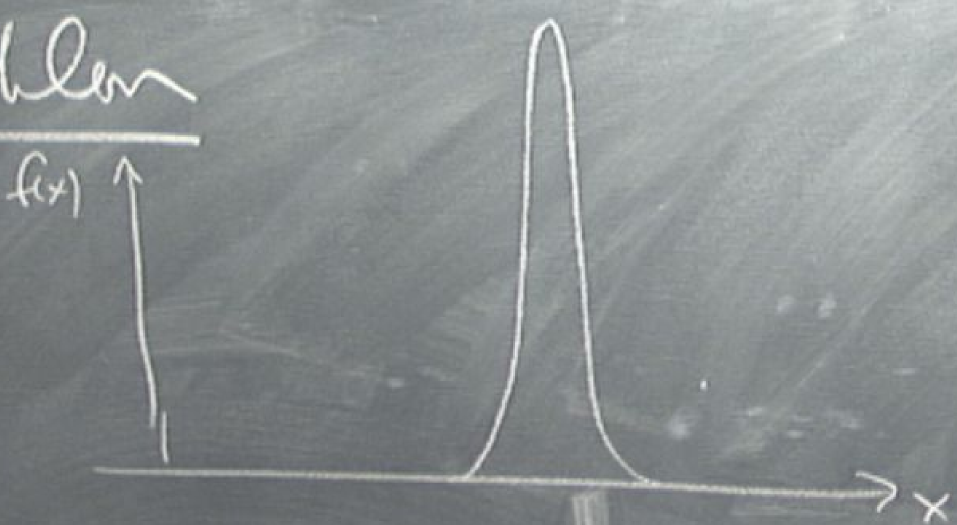
Problem



Importance Sampling : Choose the x_i with a non-uniform distribution

If $f = hg$

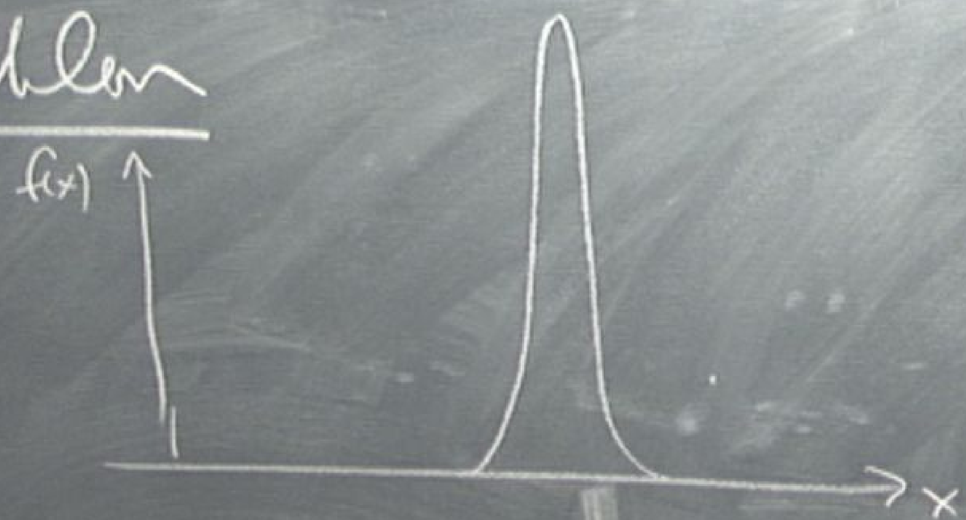
Problem



Importance Sampling: Choose the x_i with a non-uniform distribution as an probability density on $(b-a)$

$f = hg$ $\cdot$ $g(x)$

Problem

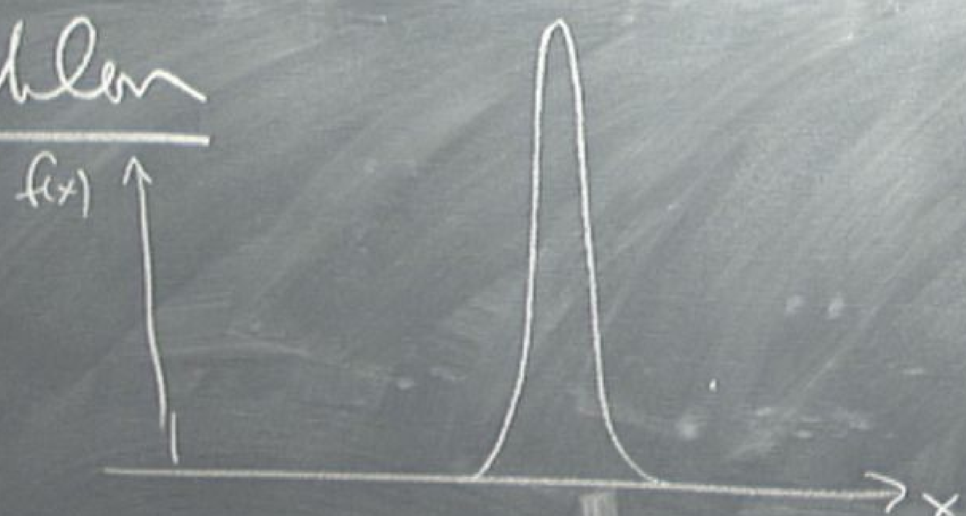


Importance Sampling: Choose the x_i with a non-uniform distribution $g(x)$ as an probability density on $(b-a)$

If $f = hg$

then $\int_a^b f(x)\,dx = \int_a^b \frac{f}{g}\, g\,dx$

Problem

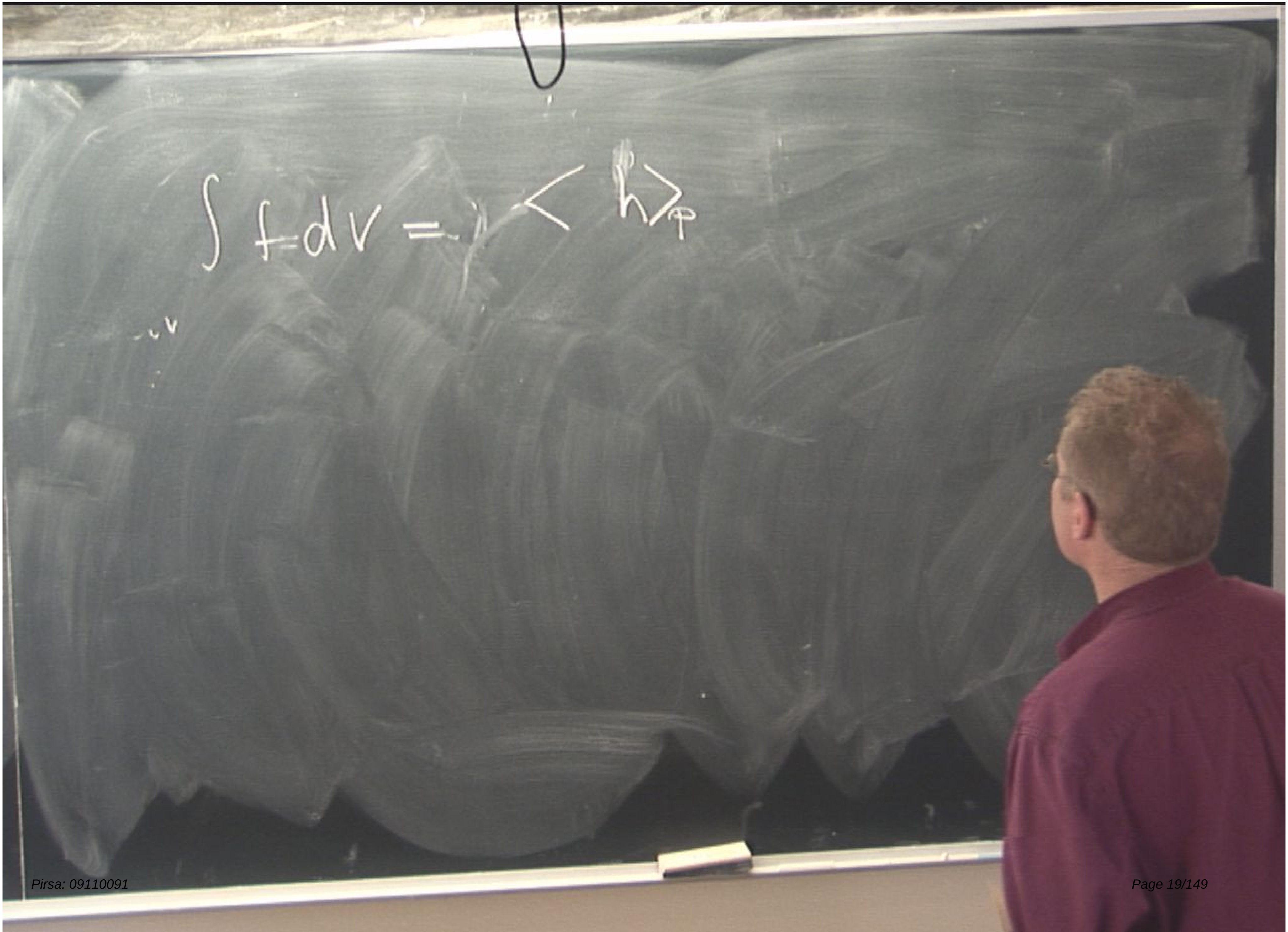


Importance Sampling: Choose the x_i with a non-uniform distribution $g(x)$ as an probability density

If $f = hg$

then $\int_a^b f(x)\,dx = \int_a^b \frac{f}{g}\, g\,dx = \int_a^b h \; g(x)\,dx$

an $(b-a)$


$\int f\,dv = \langle h \rangle_{\varphi}$

$$\int f \, dV = \langle h \rangle_g \pm \sqrt{\frac{\langle h^2 \rangle_g - \langle h \rangle_g^2}{N-1}}$$

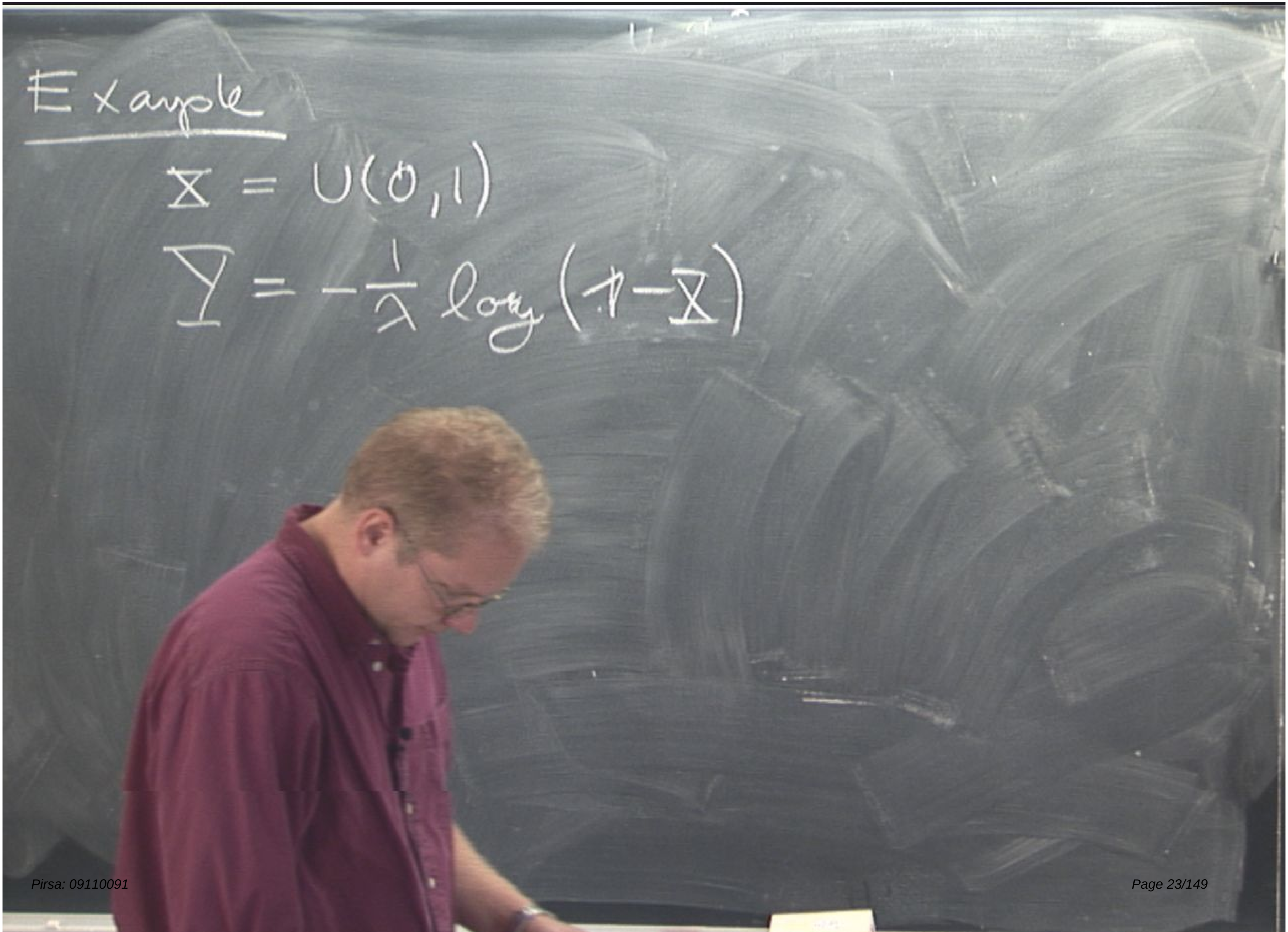
$$\int f\, dv = \langle h \rangle_g \pm \sqrt{\frac{\langle h^2 \rangle_g - \langle h \rangle_g^2}{N-1}}$$

$\langle \ \rangle_g$ average with respect to g

$$\int f\, dV = \langle h \rangle_g \pm \sqrt{\frac{\langle h^2\rangle_g - \langle h\rangle_g^2}{N-1}}$$

$\langle \; \rangle_g$ average with respect to g

We need to generate points with a desired distribution

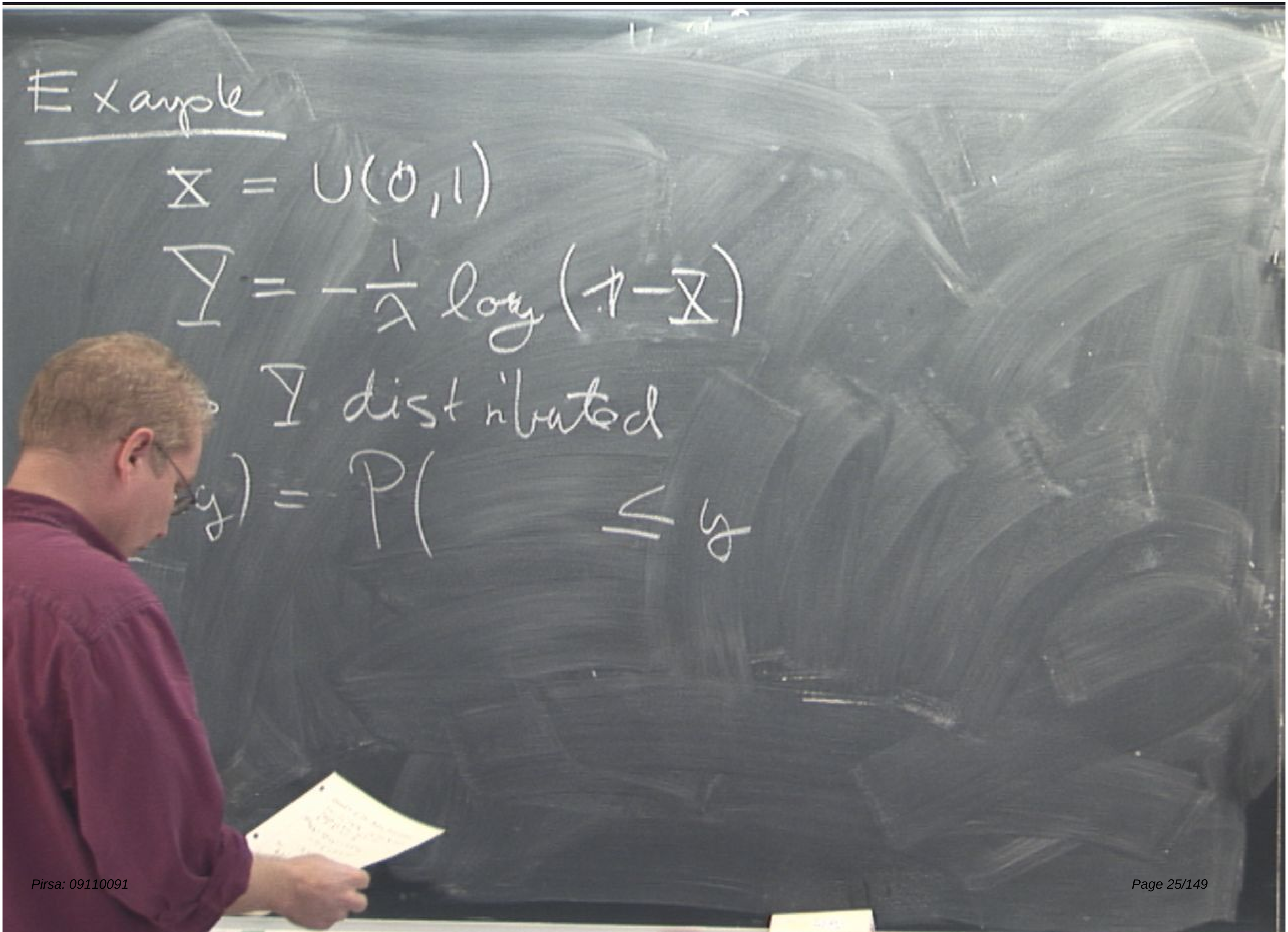

Example
X = U(0,1)
Y = -1/λ log(1-X)

Example

$$X = U(0,1)$$

$$Y = -\frac{1}{\lambda}\log(1-X)$$

How is Y distributed


Example
X = U(0,1)
Y = − 1/λ log(1−X)
Y distributed
y) = P(≤ y

Example

$$\mathbb{X} = U(0,1)$$

$$\mathbb{Y} = -\frac{1}{\lambda} \log(1-\mathbb{X})$$

How is $\mathbb{Y}$ distributed

$$F_{\mathbb{Y}}(y) = P(\mathbb{Y} \leq y)$$

$$= P\left(-\frac{1}{\lambda}\log(1-\mathbb{X}) \leq y\right)$$

Example

$$X = U(0,1)$$

$$Y = -\frac{1}{\lambda}\log(1-X)$$

How is Y distributed

$$F_Y(y) = P(Y \leq y)$$

$$= P\left(-\frac{1}{\lambda}\log(1-X) \leq y\right)$$

$$= P(\log(1-X) \geq -\lambda y)$$

Example

$$X = U(0,1)$$

$$Y = -\frac{1}{\lambda}\log(1-X)$$

How is Y distributed

$$F_Y(y) = P(Y \le y)$$

$$= P\left(-\frac{1}{\lambda}\log(1-X) \le y\right)$$

$$= P(\log(1-X) \geq -\lambda y)$$

$$= P(1-X \geq e^{-\lambda y})$$

Example

$$X = U(0,1)$$

$$Y = -\frac{1}{\lambda}\log(1-X)$$

How is Y distributed

$$F_Y(y) = P(Y \leq y)$$

$$= P\left(-\frac{1}{\lambda}\log(1-X) \leq y\right)$$

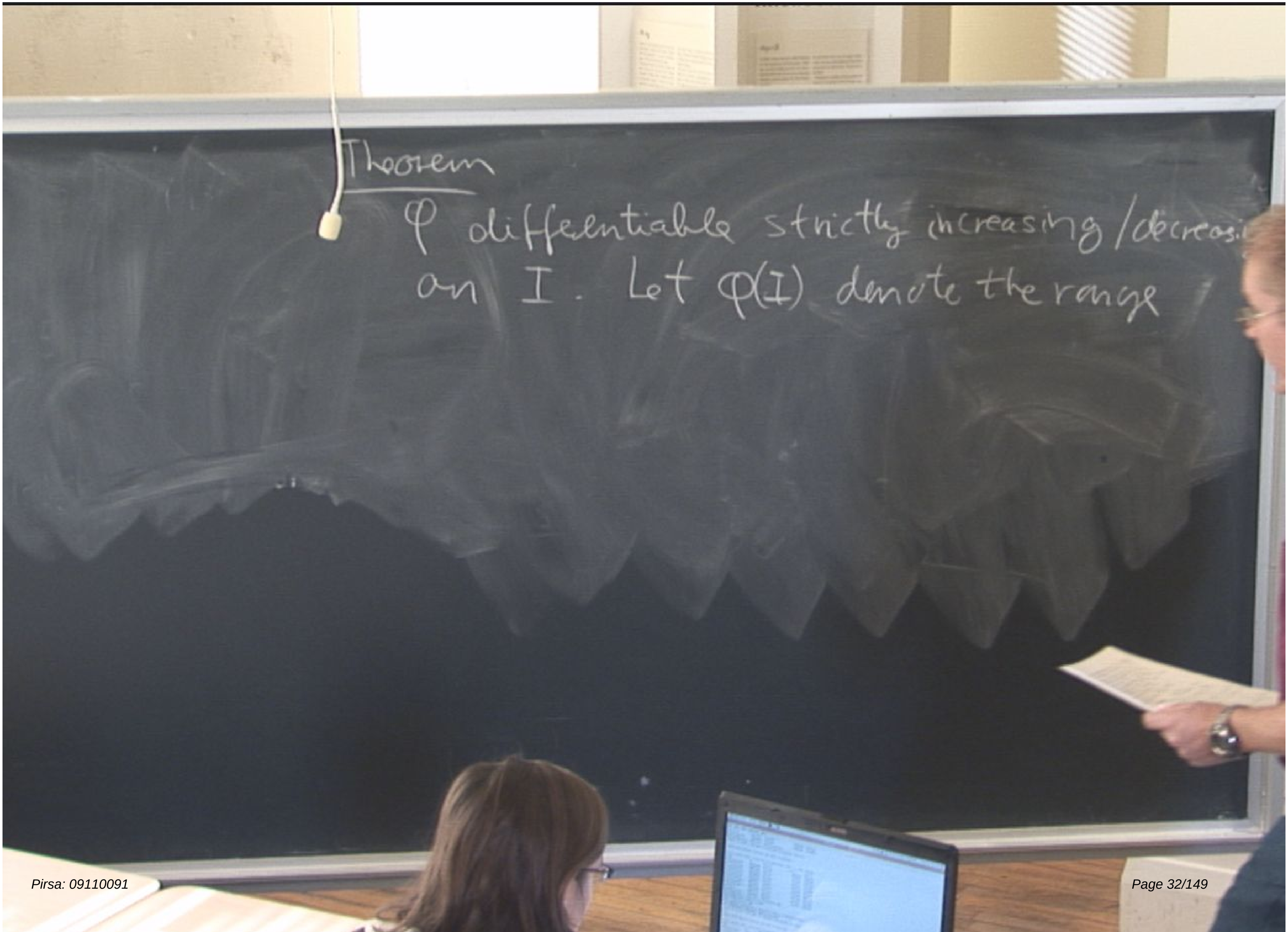
$$= P(\log(1-X) \geq -\lambda y)$$

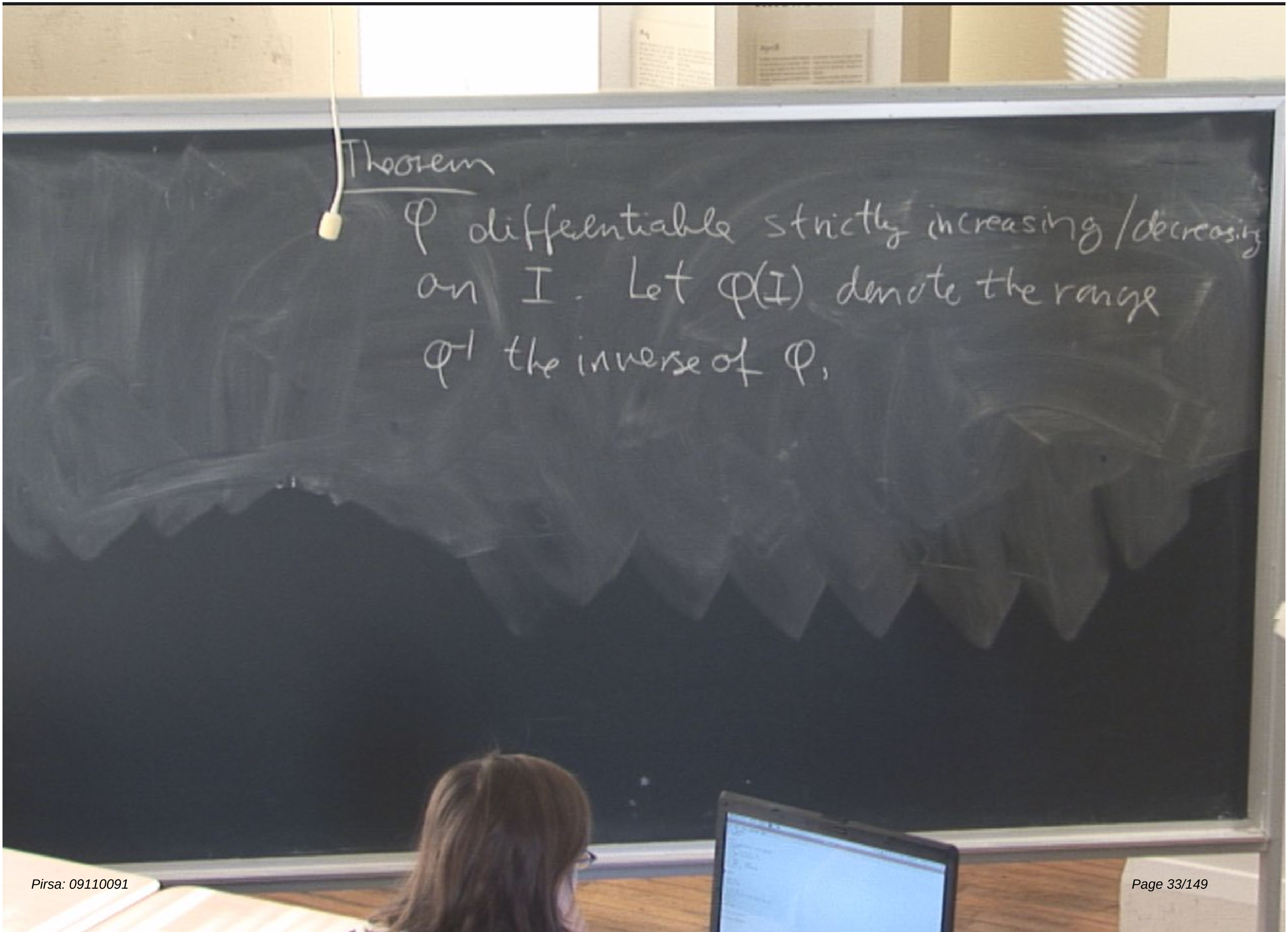
$$= P(1-X \geq e^{-\lambda y}) = P(X \leq 1-e^{-\lambda y})$$

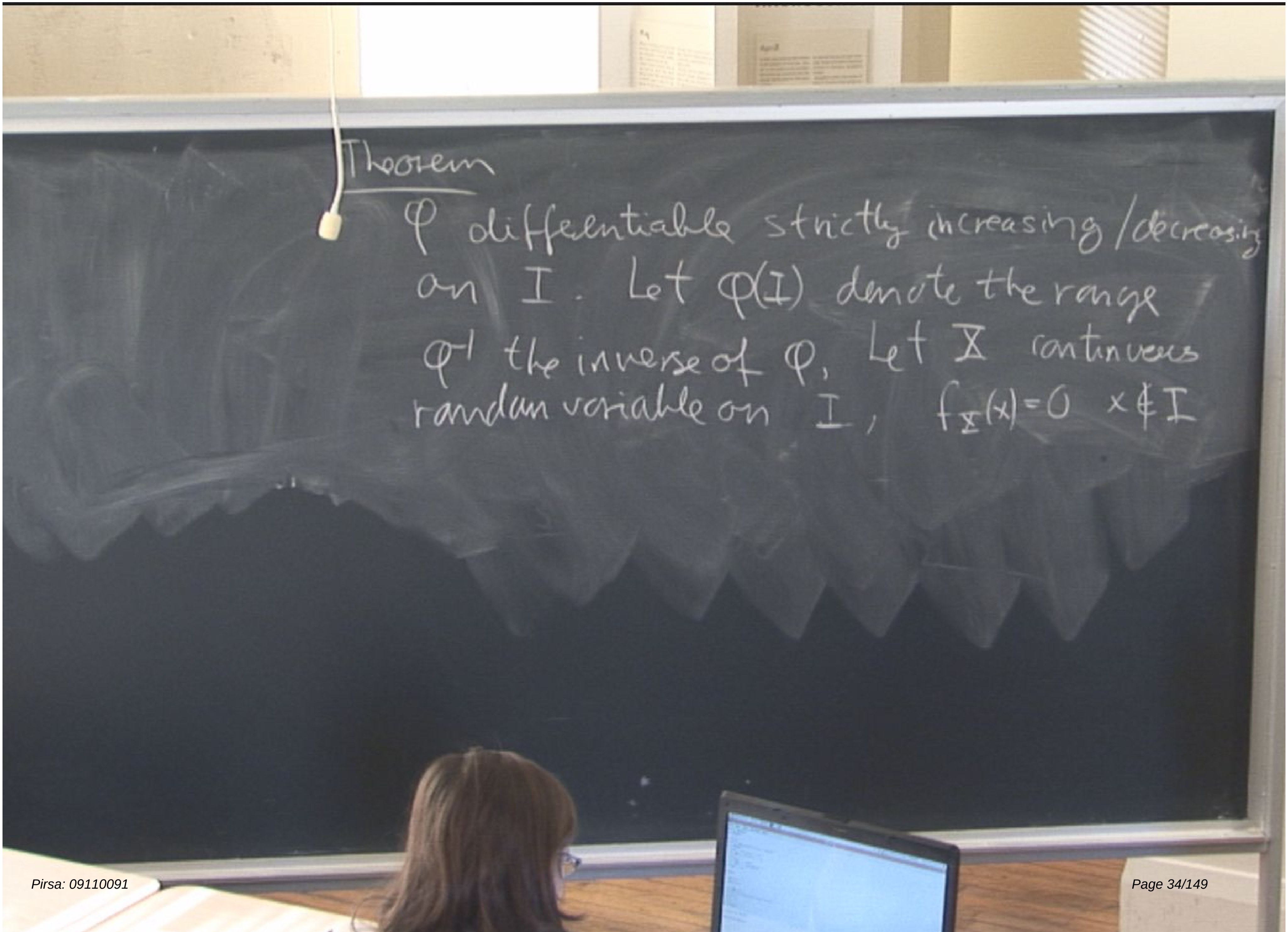
$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y} & y \geqslant 0 \\ 0 & y < 0 \end{cases}$$

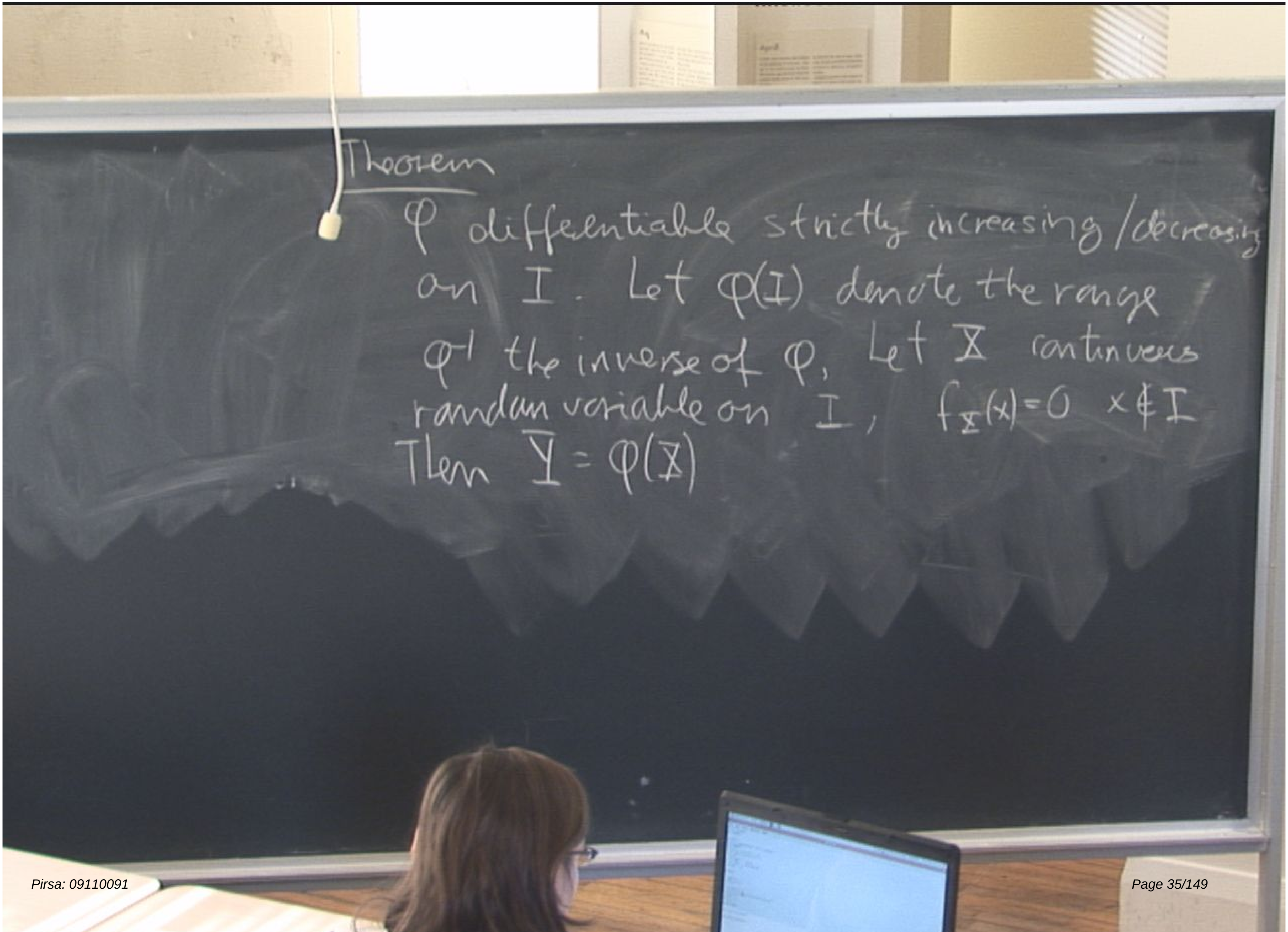
specify this and then calculate

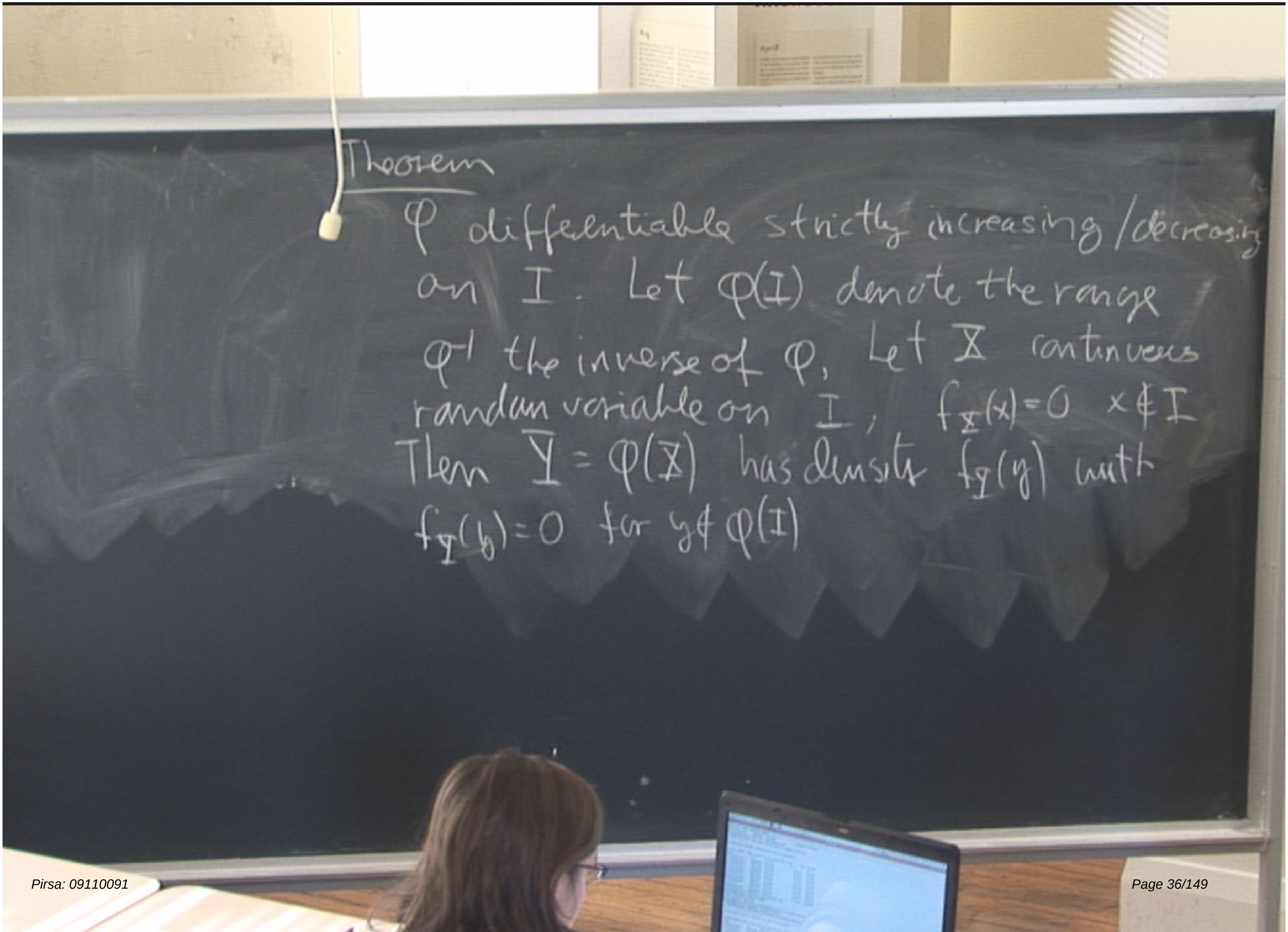

Theorem
φ differentiable strictly increasing / decreasi
on I. Let φ(I) denote the range

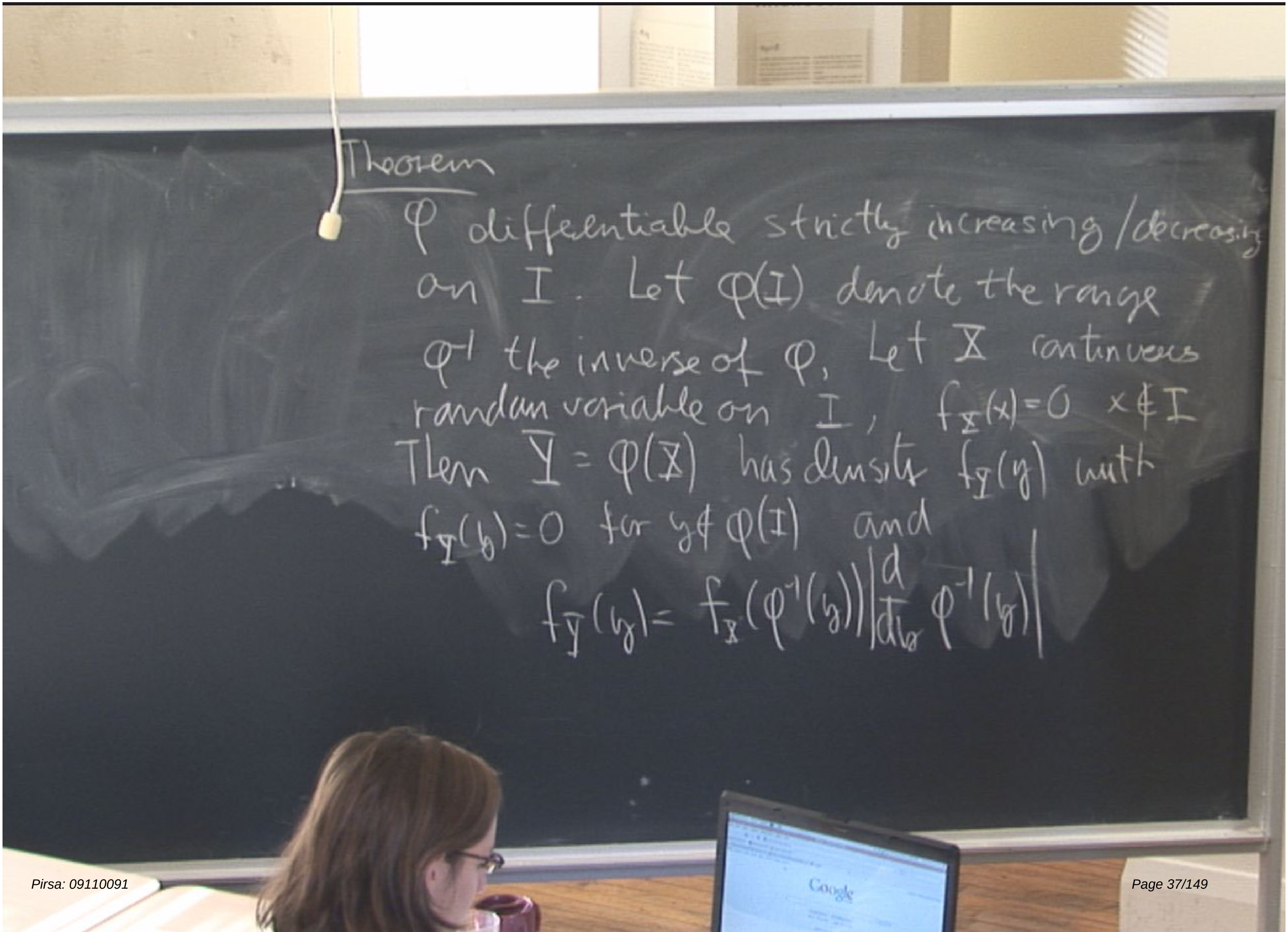

Theorem
φ differentiable strictly increasing/decreasing
on I. Let φ(I) denote the range
φ⁻¹ the inverse of φ.



Theorem
φ differentiable strictly increasing/decreasing
on I. Let $\varphi(I)$ denote the range
φ^{-1} the inverse of φ, Let X continuous
random variable on I, $f_X(x) = 0$ $x \notin I$


Theorem
φ differentiable strictly increasing/decreasing
on I. Let φ(I) denote the range
φ⁻¹ the inverse of φ, Let X continuous
random variable on I, f_X(x) = 0 x ∉ I
Then Y = φ(X)


Theorem
φ differentiable strictly increasing / decreasing
on I. Let $\varphi(I)$ denote the range
φ^{-1} the inverse of φ. Let X continuous
random variable on I, $f_X(x) = 0 \quad x \notin I$
Then $Y = \varphi(X)$ has density $f_Y(y)$ with
$f_Y(y) = 0$ for $y \notin \varphi(I)$


Theorem
φ differentiable strictly increasing / decreasing
on I. Let $\varphi(I)$ denote the range
φ^{-1} the inverse of φ, Let X continuous
random variable on I, $f_X(x) = 0 \quad x \notin I$
Then $Y = \varphi(X)$ has density $f_Y(y)$ with
$f_Y(y) = 0$ for $y \notin \varphi(I)$ and
$f_Y(y) = f_X(\varphi^{-1}(y)) \left| \frac{d}{dy} \varphi^{-1}(y) \right|$

Theorem

φ differentiable strictly increasing/decreasing on I. Let $\varphi(I)$ denote the range, φ^{-1} the inverse of φ. Let X continuous random variable on I, $f_X(x) = 0 \quad x \notin I$

Then $Y = \varphi(X)$ has density $f_Y(y)$ with

$f_Y(y) = 0$ for $y \notin \varphi(I)$ and

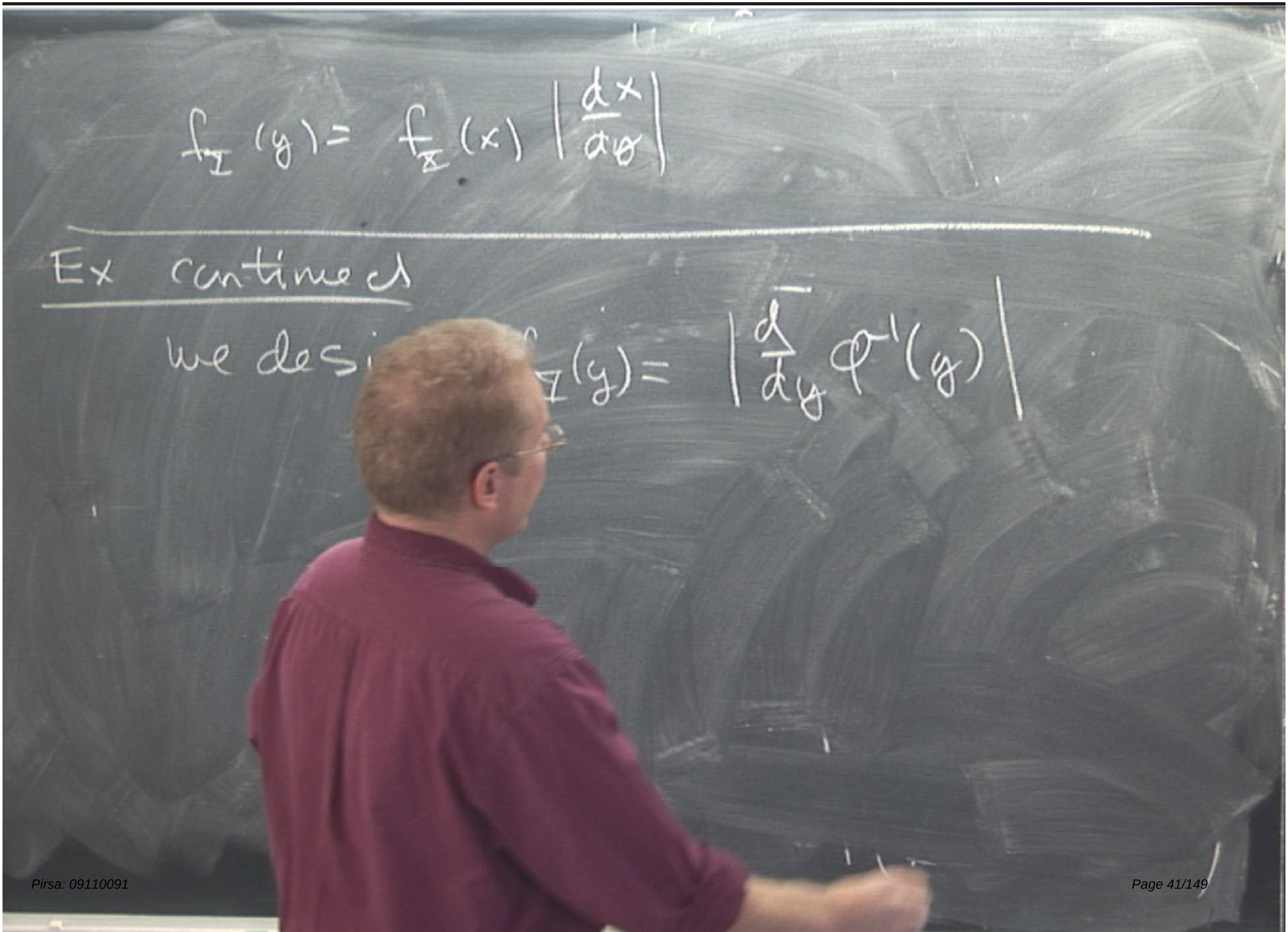
$x = \varphi^{-1}(y)$

$$\boxed{f_Y(y) = f_X(\varphi^{-1}(y)) \left| \frac{d}{dy} \varphi^{-1}(y) \right|}$$

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

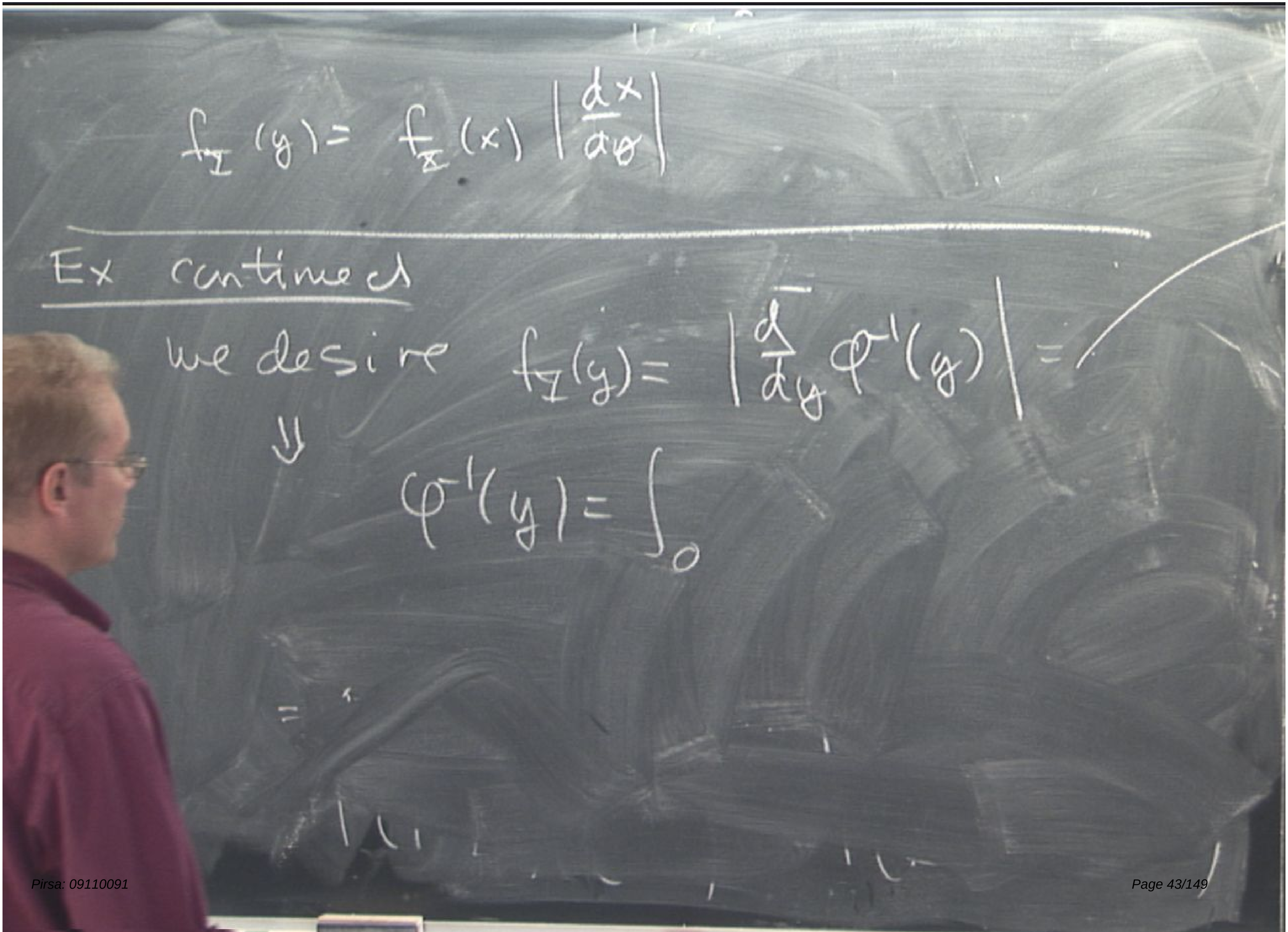
Ex continued


$f_Y(y) = f_X(x) \left|\frac{dx}{dy}\right|$
Ex continued
we desi
$f_Y(y) = \left|\frac{d}{dy}\varphi^{-1}(y)\right|$

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right|$$

Ex continued

we desire $f_Y(y) = \left| \frac{d}{dy} \varphi^{-1}(y) \right| =$


$f_Y(y) = f_X(x) \left|\frac{dx}{dy}\right|$
Ex continued
we desire $f_Y(y) = \left|\frac{d}{dy}\varphi^{-1}(y)\right| =$
$\Downarrow$
$\varphi^{-1}(y) = \int_0$
$=$

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

Ex continued

we desire $f_Y(y) = \left|\frac{d}{dy}\varphi^{-1}(y)\right| =$

$\Downarrow$

$$\varphi^{-1}(y) = \int_0^y \lambda e^{-\lambda z}\, dz$$

$=$

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

Ex continued

we desire $f_Y(y) = \left|\frac{d}{dy}\varphi^{-1}(y)\right| =$

$\Downarrow$

$$\varphi^{-1}(y) = \int_0^y \lambda e^{-\lambda z}\, dz = 1-e^{-}$$

=

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

Ex continued

we desire $f_Y(y) = \left|\frac{d}{dy}\varphi^{-1}(y)\right| = $

$\Downarrow$

$$\varphi^{-1}(y) = \int_0^y \lambda e^{-\lambda z}\, dz = 1-e^{\ldots}$$

$\Downarrow$

$$X = 1-e^{-\lambda y}$$

Solve for $y(x)$

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

Ex continued

we desire $f_Y(y) = \left|\frac{d}{dy}\varphi^{-1}(y)\right| =$

$\Downarrow$

$$\varphi^{-1}(y) = \int_0^y \lambda e^{-\lambda z}\, dz =$$

$\Downarrow$

$$x = 1 - e^{-\lambda y}$$

Solve for $y(x) = -\frac{1}{\lambda}\log(1-x)$

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right|$$

Ex continued

we desire $f_Y(y) = \left| \frac{d}{dy} \varphi^{-1}(y) \right| =$

$\Downarrow$

$$\varphi^{-1}(y) = \int_0^y \lambda e^{-\lambda z} dz = 1 - e^{-\lambda y}$$

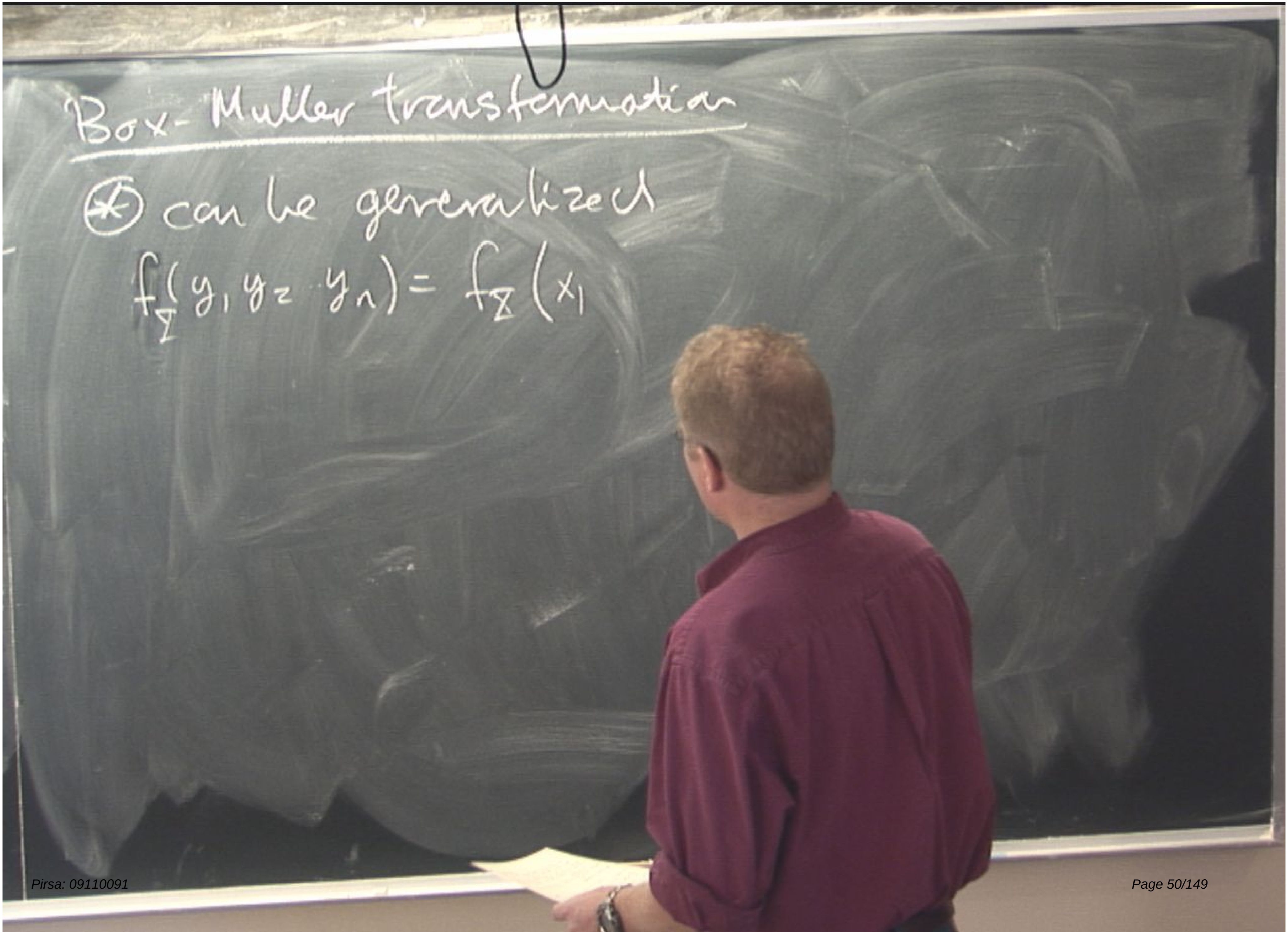
$\Downarrow$

$$X = 1 - e^{-\lambda y}$$

Solve for $y(X) = -\frac{1}{\lambda} \log(1-X)$

Box-Muller transformation

(*) can be generalized


Box-Muller transformation
⊛ can be generalized
$f_Y(y_1, y_2 \cdots y_n) = f_X(x_1$

Box-Muller transformation

(*) can be generalized

$$f_{\underline{Y}}(y_1, y_2 \cdots y_n)\, dy_1\, dy_2 \cdots dy_n = f_{\underline{X}}(x_1 \cdots x_n)$$

Box-Muller transformation

(*) can be generalized

$$f_{\underline{Y}}(y_1, y_2 \cdots y_n)\, dy_1\, dy_2 \cdots dy_n = f_{\underline{X}}(x_1 \cdots x_n) \left| \frac{\partial(x_1, x_2 \cdots x_n)}{\partial(y_1 \cdots y_n)} \right| dy_1 \cdots dy_n$$

Box-Muller transformation

(*) can be generalized

$$f_Y(y_1, y_2 \cdots y_n)\, dy_1\, dy_2 \cdots dy_n = f_X(x_1 \cdots x_n) \left| \frac{\partial(x_1, x_2 \cdots x_n)}{\partial(y_1 \cdots y_n)} \right| dy_1 \cdots dy_n$$

Example 2 $\quad x_1\ x_2 \quad U(0,1)$

Box-Muller transformation

(*) can be generalized

$$f_{\underline{Y}}(y_1, y_2 \cdots y_n)\, dy_1\, dy_2 \cdots dy_n = f_{\underline{X}}(x_1 \cdots x_n) \left| \frac{\partial(x_1, x_2 \cdots x_n)}{\partial(y_1 \cdots y_n)} \right| dy_1 \cdots dy_n$$

Example 2 $\quad x_1 \; x_2 \quad U(0,1)$

$$y_1 = \sqrt{-2\ln x_1}\, \cos(2\pi x_2)$$

$$y_2 = \sqrt{-2\ln x_1}\, \sin(2\pi x_2)$$

Box-Muller transformation

(*) can be generalized

$$f_{\underline{Y}}(y_1, y_2 \cdots y_n)\, dy_1\, dy_2 \cdots dy_n = f_{\underline{X}}(x_1 \cdots x_n) \left| \frac{\partial(x_1 x_2 \cdots x_n)}{\partial(y_1 \cdots y_n)} \right| dy_1 \cdots dy_n$$

Example 2 $\quad X_1 \; X_2 \quad U(0,1)$

$$\left.\begin{aligned} y_1 &= \sqrt{-2\ln x_1}\,\cos(2\pi x_2) \\ y_2 &= \sqrt{-2\ln x_1}\,\sin(2\pi x_2) \end{aligned}\right) \quad \begin{aligned} x_1 &= e^{-\frac{1}{2}(y_1^2 + y_2^2)} \\ x_2 &= \end{aligned}$$

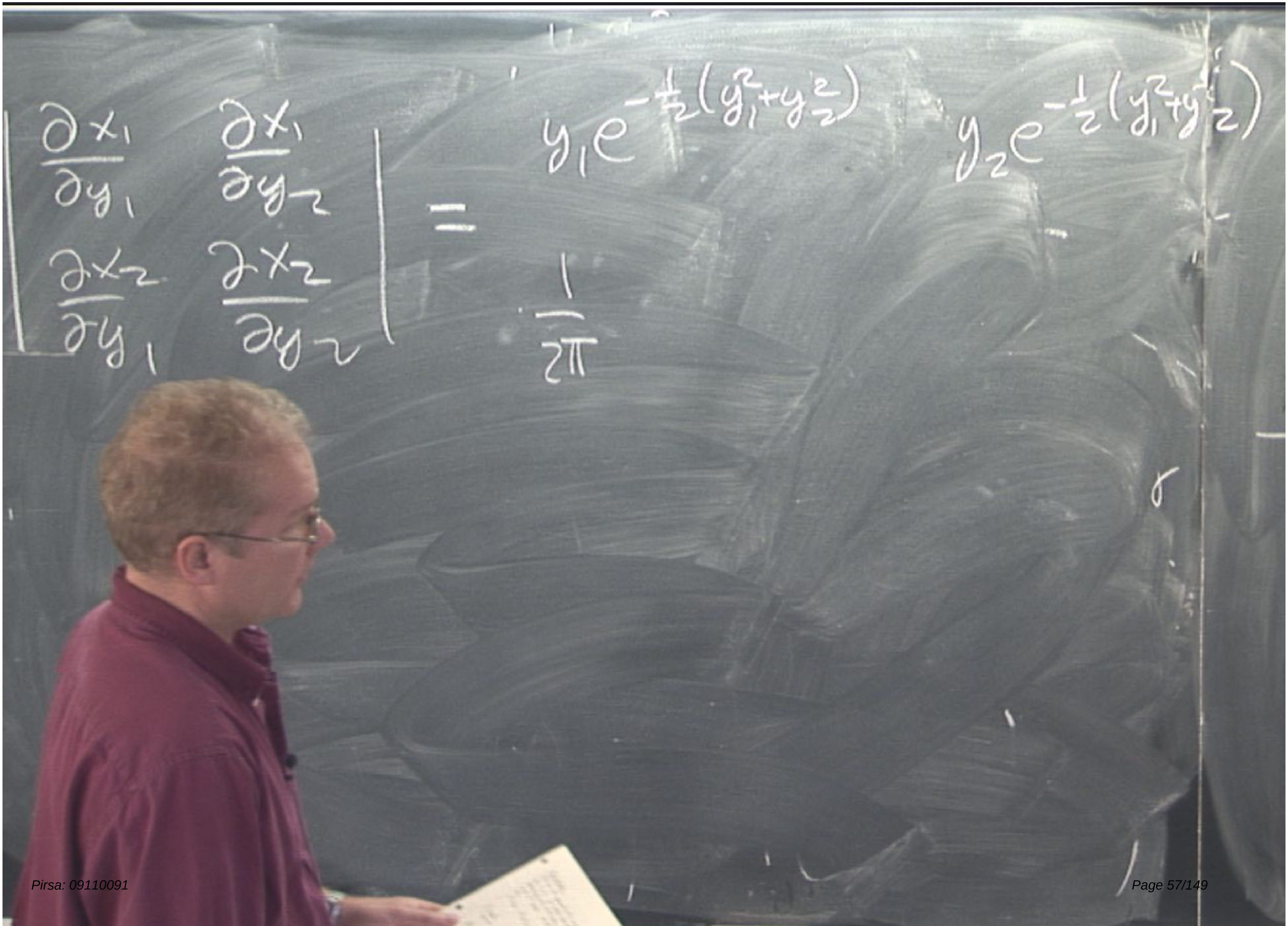
Box-Muller transformation

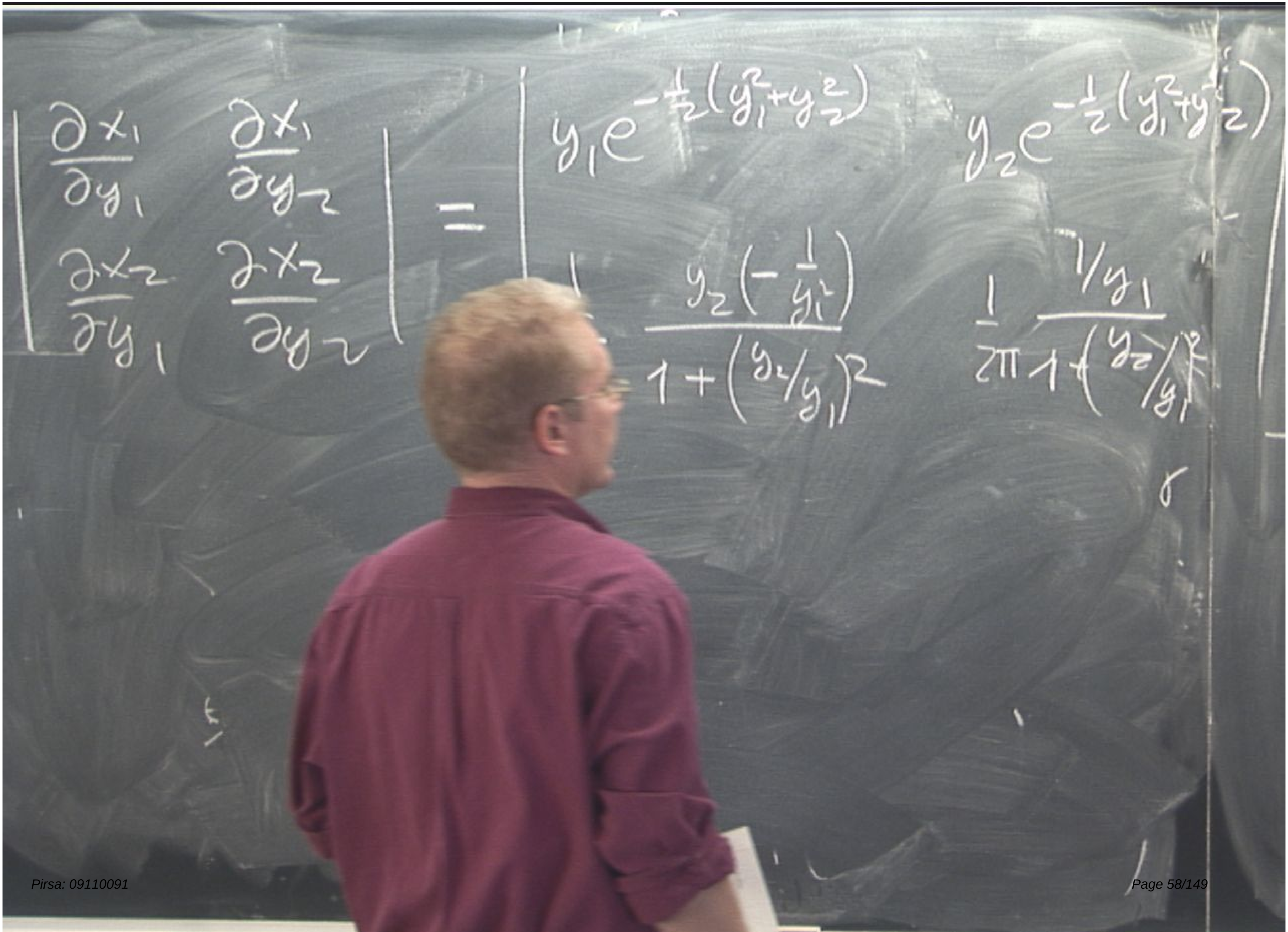
(*) can be generalized

$$f_{\underline{Y}}(y_1, y_2 \cdots y_n)\, dy_1\, dy_2 \cdots dy_n = f_{\underline{X}}(x_1 \cdots x_n) \left| \frac{\partial(x_1 x_2 \cdots x_n)}{\partial(y_1 \cdots y_n)} \right| dy_1 \cdots dy_n$$

Example 2 $\quad x_1\ x_2 \quad U(0,1)$

$$\left.\begin{aligned} y_1 &= \sqrt{-2\ln x_1}\,\cos(2\pi x_2) \\ y_2 &= \sqrt{-2\ln x_1}\,\sin(2\pi x_2) \end{aligned}\right\} \qquad \begin{aligned} x_1 &= e^{-\frac{1}{2}(y_1^2+y_2^2)} \\ x_2 &= \frac{1}{2\pi}\tan^{-1}\frac{y_2}{y_1} \end{aligned}$$


$\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} =$
$y_1 e^{-\frac{1}{2}(y_1^2+y_2^2)}$
$y_2 e^{-\frac{1}{2}(y_1^2+y_2^2)}$
$\frac{1}{2\pi}$



$\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} =$
$y_1 e^{-\frac{1}{2}(y_1^2+y_2^2)}$
$y_2 e^{-\frac{1}{2}(y_1^2+y_2^2)}$
$\frac{y_2\left(-\frac{1}{y_1^2}\right)}{1+\left(y_2/y_1\right)^2}$
$\frac{1}{2\pi}\frac{1/y_1}{1+\left(y_2/y_1\right)^2}$

$$\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_1 e^{-\frac{1}{2}(y_1^2+y_2^2)} & y_2 e^{-\frac{1}{2}(y_1^2+y_2^2)} \\ \frac{1}{2\pi}\frac{y_2\left(-\frac{1}{y_1^2}\right)}{1+\left(y_2/y_1\right)^2} & \frac{1}{2\pi}\frac{1/y_1}{1+\left(y_2/y_1\right)^2} \end{vmatrix}$$

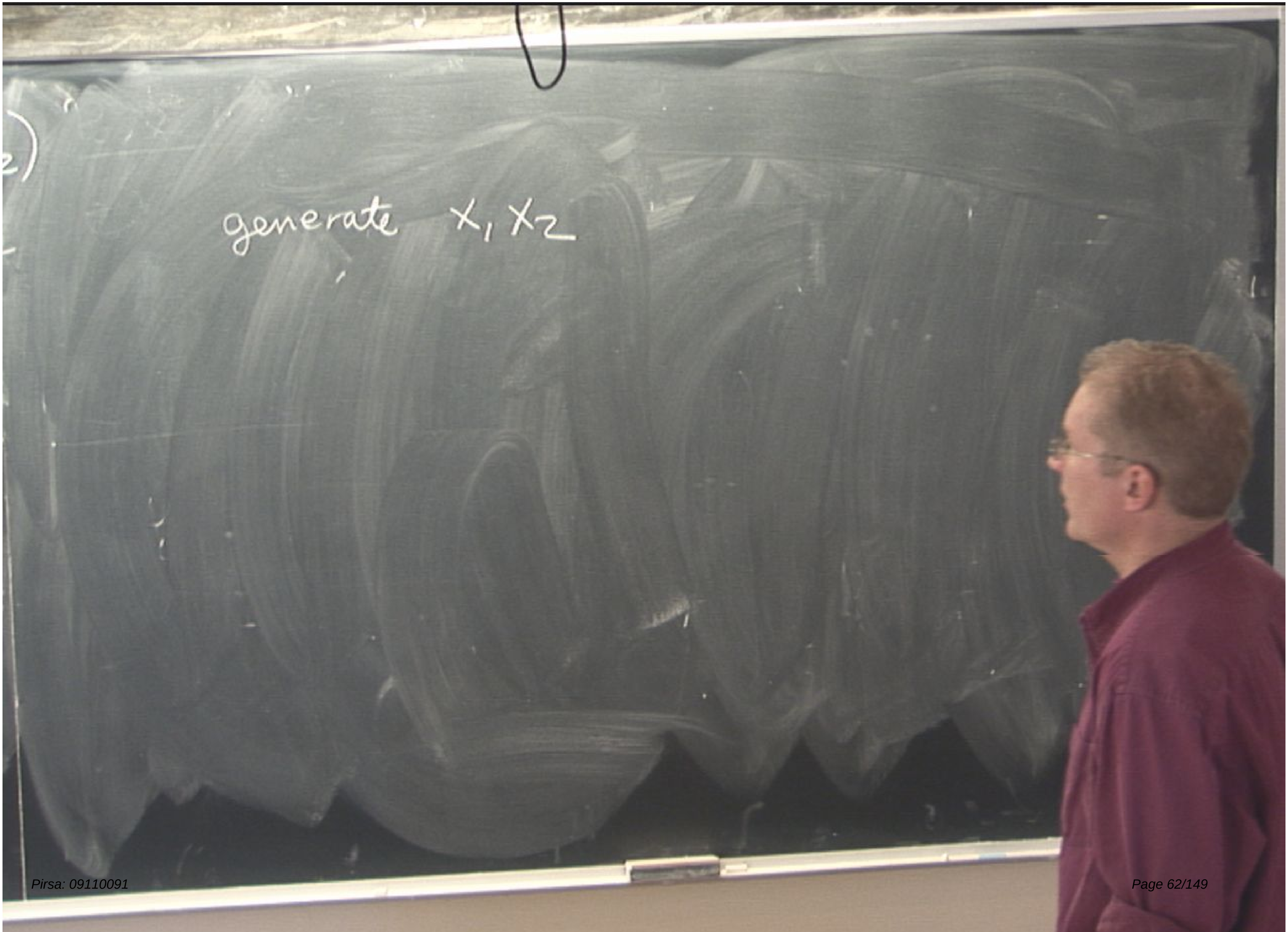
$$\begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_1 e^{-\frac{1}{2}(y_1^2+y_2^2)} & y_2 e^{-\frac{1}{2}(y_1^2+y_2^2)} \\ \dfrac{1}{2\pi}\dfrac{y_2\left(-\frac{1}{y_1^2}\right)}{1+\left(y_2/y_1\right)^2} & \dfrac{1}{2\pi}\dfrac{1/y_1}{1+\left(y_2/y_1\right)^2} \end{vmatrix}$$

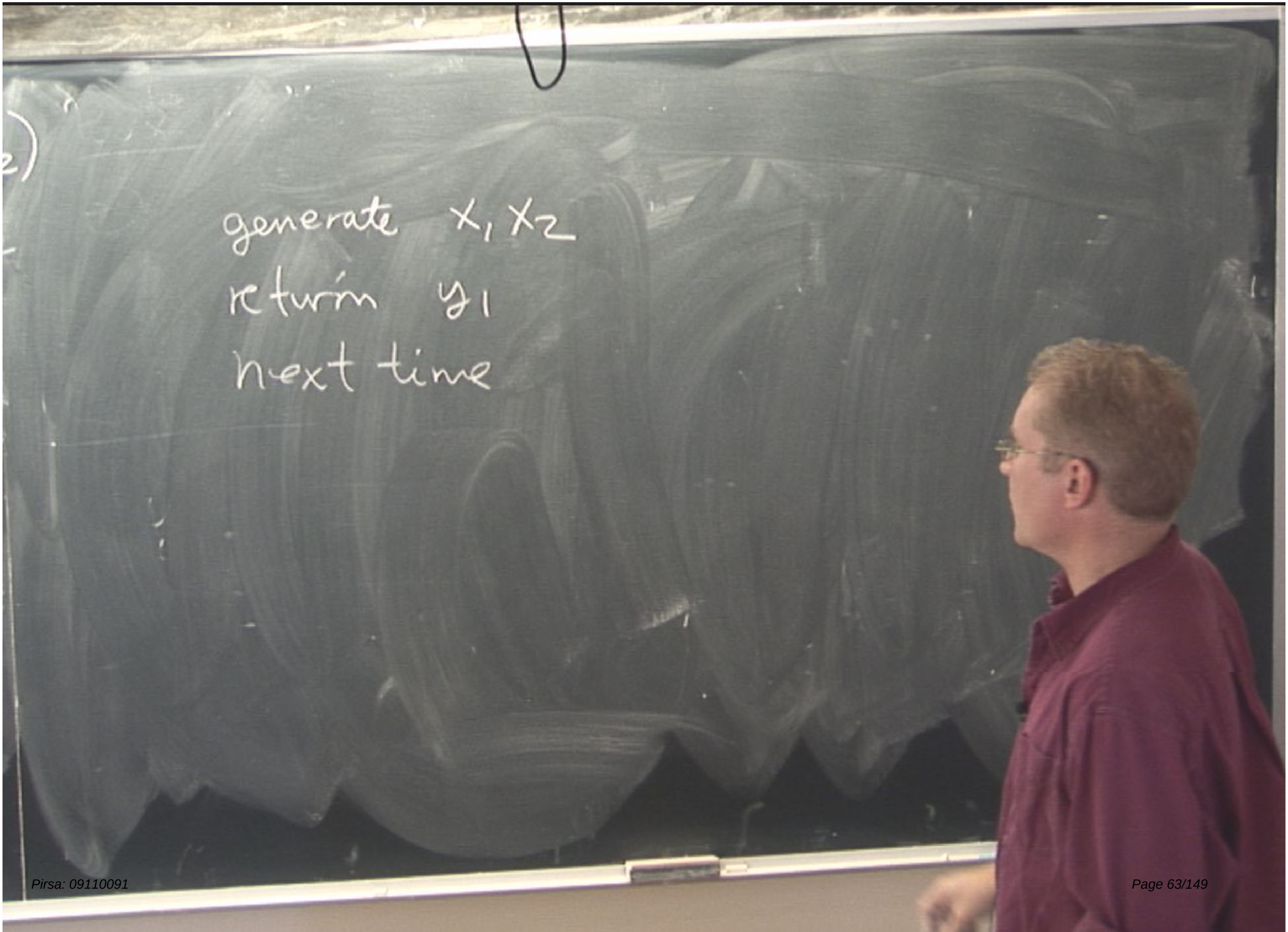
$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y_1^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y_2^2}$$

$$\begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_1 e^{-\frac{1}{2}(y_1^2+y_2^2)} & y_2 e^{-\frac{1}{2}(y_1^2+y_2^2)} \\ \dfrac{1}{2\pi} \dfrac{y_2\left(-\frac{1}{y_1^2}\right)}{1+\left(y_2/y_1\right)^2} & \dfrac{1}{2\pi} \dfrac{1/y_1}{1+\left(y_2/y_1\right)^2} \end{vmatrix}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y_1^2} \, \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y_2^2}$$

$\Rightarrow$ y_1 y_2 are independent.


generate x_1, x_2


generate x_1, x_2
return y_1
next time

generate x_1, x_2

return y_1

next time y_2

⋮

generate new x_1, x_2

generate x_1, x_2

return y_1

next time y_2

⋮

generate new x_1 x_2

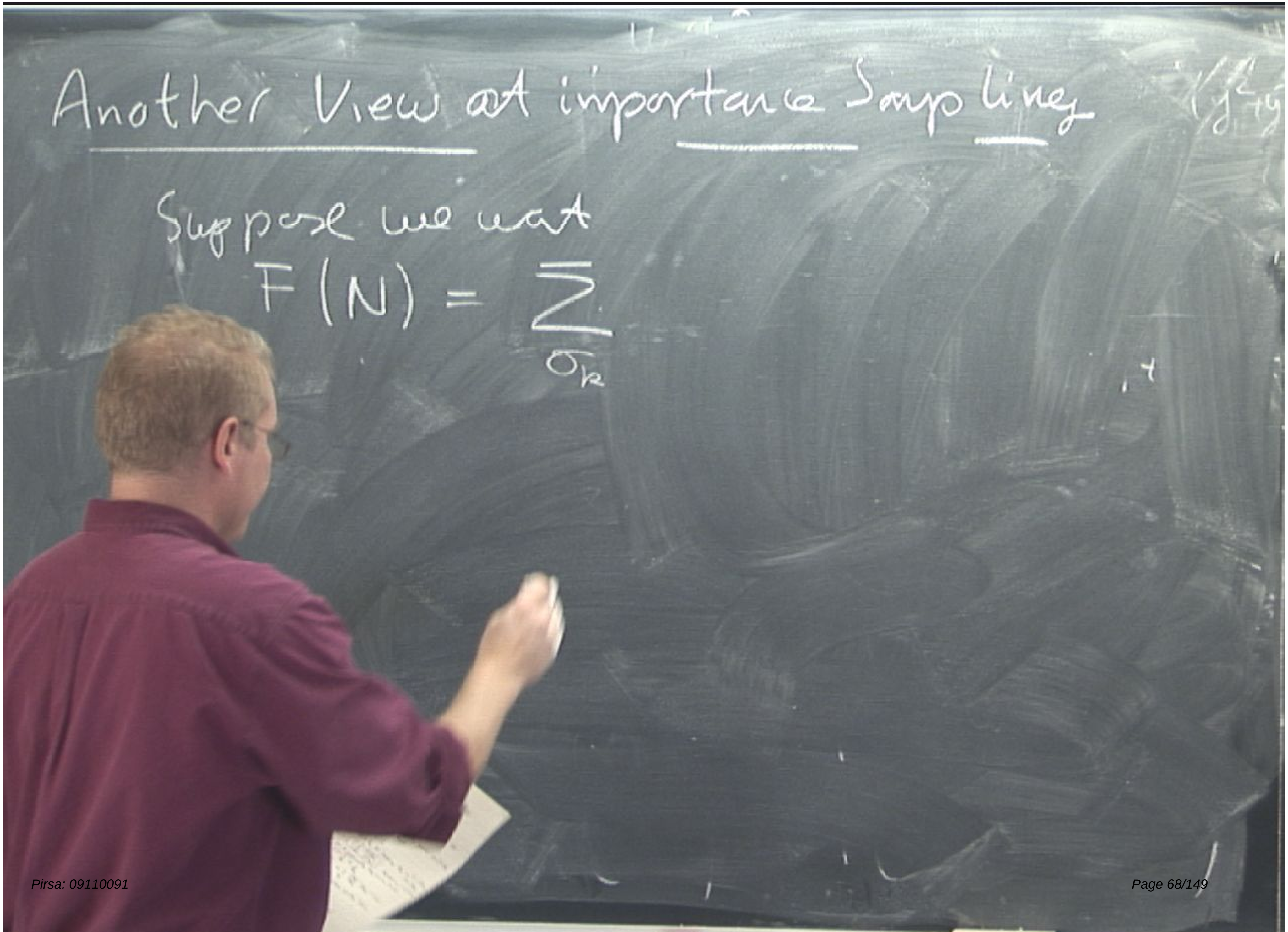
Unfortunately this is not feasible for importance sampling

Another View at importance Sampling

Another View at importance Sampling

Suppose we want

$$F(N) = \sum$$


Another View at importance Sampling
Suppose we want
$F(N) = \sum_{\sigma_k}$

Another View at importance Sampling

Suppose we want

$$F(N) = \sum_{j=1}^{N} f(\sigma_j)$$

Another View at importance Sampling

Suppose we want

$$F(N) = \sum_{j=1}^{N} f(\sigma_j) \quad \text{but } N \sim 2^{2000}$$

Another View at importance Sampling

Suppose we want

$$F(N) = \sum_{j=1}^{N} f(\sigma_j) \quad \text{but } N \sim 2^{2000}$$

So

$$F(N) = N\left(\frac{1}{M}\sum_{k=1}^{M} f(\sigma_k)\right)$$

Another View at importance Sampling

Suppose we want

$$F(N) = \sum_{j=1}^{N} f(\sigma_j) \quad \text{but } N \sim 2^{2000}$$

So

$$F(N) = N\left(\frac{1}{M}\sum_{k=1}^{M} f(\sigma_k)\right)$$

σ_k are chosen uniformly

Another View at importance Sampling

Suppose we want

$$F(N) = \sum_{j=1}^{N} f(\sigma_j) \quad \text{but } N \sim 2^{2000}$$

So

$$F(N) = N\left(\frac{1}{M}\sum_{k=1}^{M} f(\sigma_k)\right)$$

σ_k are chosen uniformly

We claim

$$\langle f \rangle = \sum_{k=1}^{M} f(\sigma_k)\,\frac{1}{M}$$

2) The probability for choosing σ_k is $p(\sigma_k) = \frac{1}{N}$

The probability for choosing σ_k is $p(\sigma_k) = \frac{1}{N}$

So

$$F(N) \simeq \frac{1}{M} \left(\sum_{k=1}^{M} \frac{f(\sigma_k)}{\quad} \right.$$

The probability for choosing σ_k is $p(\sigma_k) = \frac{1}{N}$

So

$$F(N) \simeq \frac{1}{M}\left(\sum_{k=1}^{M} \frac{f(\sigma_k)}{p(\sigma_k)}\right)$$

This works even for non-uniform $p(\sigma_k)$

The probability for choosing σ_k is $p(\sigma_k) = \frac{1}{N}$

So $\boxed{F(N) \simeq \frac{1}{M}\left(\sum_{k=1}^{M} \frac{f(\sigma_k)}{p(\sigma_k)}\right)}$

This works even for non-uniform $p(\sigma_k)$

why

2) The probability for choosing σ_k is $p(\sigma_k) = \frac{1}{N}$

So $\boxed{F(N) \simeq \frac{1}{M}\left(\sum_{k=1}^{M} \frac{f(\sigma_k)}{p(\sigma_k)}\right)}$

is works even for non-uniform $p(\sigma_k)$

ny? $M \to \infty$ then σ_k occurs $M P(\sigma_k)$ times

The probability for choosing σ_k is $p(\sigma_k) = \frac{1}{N}$

So $$\boxed{F(N) \simeq \frac{1}{M}\left(\sum_{k=1}^{M} \frac{f(\sigma_k)}{p(\sigma_k)}\right)}$$

This works even for non-uniform $p(\sigma_k)$

why? $M \to \infty$ then σ_k occurs $M\,P(\sigma_k)$ times

So $$\frac{1}{M}\sum_{k=1}^{M} \frac{f(\sigma_k)}{p(\sigma_k)} \longrightarrow \frac{1}{M}\sum_{\sigma_k} \frac{f(\sigma_k)}{p(\sigma_k)}$$

The probability for choosing σ_k is $p(\sigma_k) = \frac{1}{N}$

So $$\boxed{F(N) \simeq \frac{1}{M}\left(\sum_{k=1}^{M} \frac{f(\sigma_k)}{p(\sigma_k)}\right)}$$

This works even for non-uniform $p(\sigma_k)$ ✓

why? $M \to \infty$ then σ_k occurs $M\,P(\sigma_k)$ times

So $$\frac{1}{M}\sum_{k=1}^{M} \frac{f(\sigma_k)}{p(\sigma_k)} \longrightarrow \frac{1}{M}\sum_{\sigma_k} \frac{f(\sigma_k)}{p(\sigma_k)}\, M\,P(\sigma_k)$$

$$= F(N)$$

Suppose $\quad p(\sigma_R) = \dfrac{f(\sigma_R)}{F(N)}$

Suppose $p(\sigma_k) = \dfrac{f(\sigma_k)}{F(N)}$

$$\langle f \rangle = \frac{1}{M} \sum_{k=1}^{M} \frac{f(\sigma_k)}{f(\sigma_k)/F(N)} = F(N)$$

σ^2

Suppose $\quad p(\sigma_k) = \dfrac{f(\sigma_k)}{F(N)}$

$$\langle f \rangle = \frac{1}{M} \sum_{k=1}^{M} \frac{f(\sigma_k)}{f(\sigma_k)/F(N)} = F(N)$$

$$\sigma^2 = \frac{1}{M} \sum_{k=1}^{M} \left(\frac{f(\sigma_k)}{f(\sigma_k)/F(N)} \right)^2 - F^2$$

$$= F^2 - F^2 = 0$$

The probability for choosing σ_k is $p(\sigma_k) = \frac{1}{N}$

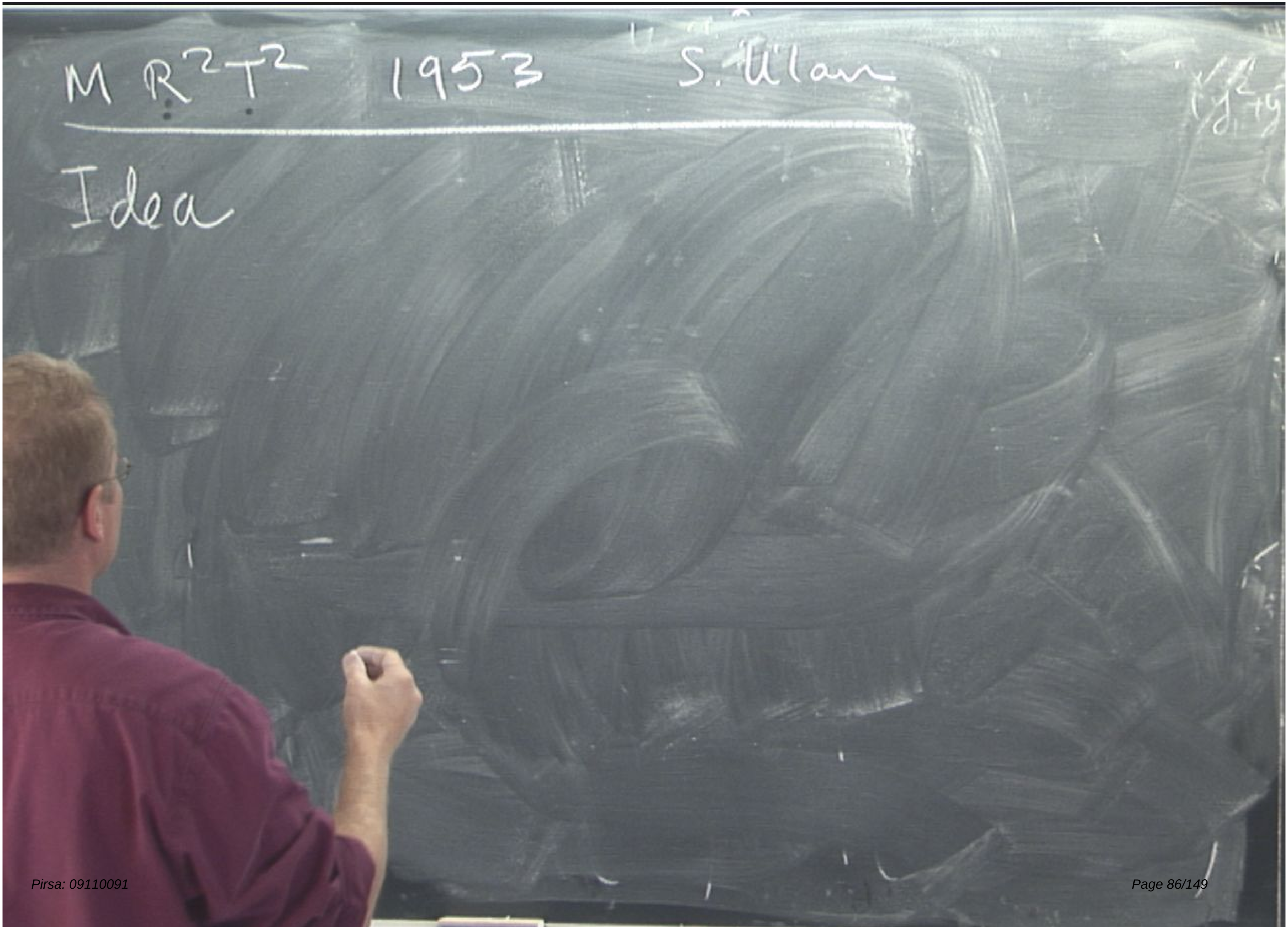
So $$\boxed{F(N) \simeq \frac{1}{M}\left(\sum_{k=1}^{M} \frac{f(\sigma_k)}{p(\sigma_k)}\right)}$$ ← time average

This works even for non-uniform $p(\sigma_k)$ ✓

why? $M \to \infty$ then σ_k occurs $M\,P(\sigma_k)$ times

So $$\frac{1}{M}\sum_{k=1}^{M} \frac{f(\sigma_k)}{p(\sigma_k)} \longrightarrow \frac{1}{M}\sum_{\sigma_k} \frac{f(\sigma_k)}{p(\sigma_k)} M\,P(\sigma_k) = F(N)$$

$M \quad R^2 T^2$


M R²T² 1953 S. Ulam
Idea

MR^2T^2 1953 S. Ulam

Idea: Replace the problem with another problem from Markov Chains that gives the same averages.

M R^2T^2 1953 S. Ulam

Idea: Replace the problem with another problem from Markov Chains that gives the same averages.

Setup a random walk on the x-axis or in the space of configurations where x_i / configs are visited with the desired distrib.

$M R^2 T^2$ 1953 S. Ulam

Idea: Replace the problem with another problem from Markov Chains that gives the same averages

Setup a random walk on the x-axis or in the space of configurations where x_i / configs are visited with the desired distrib.

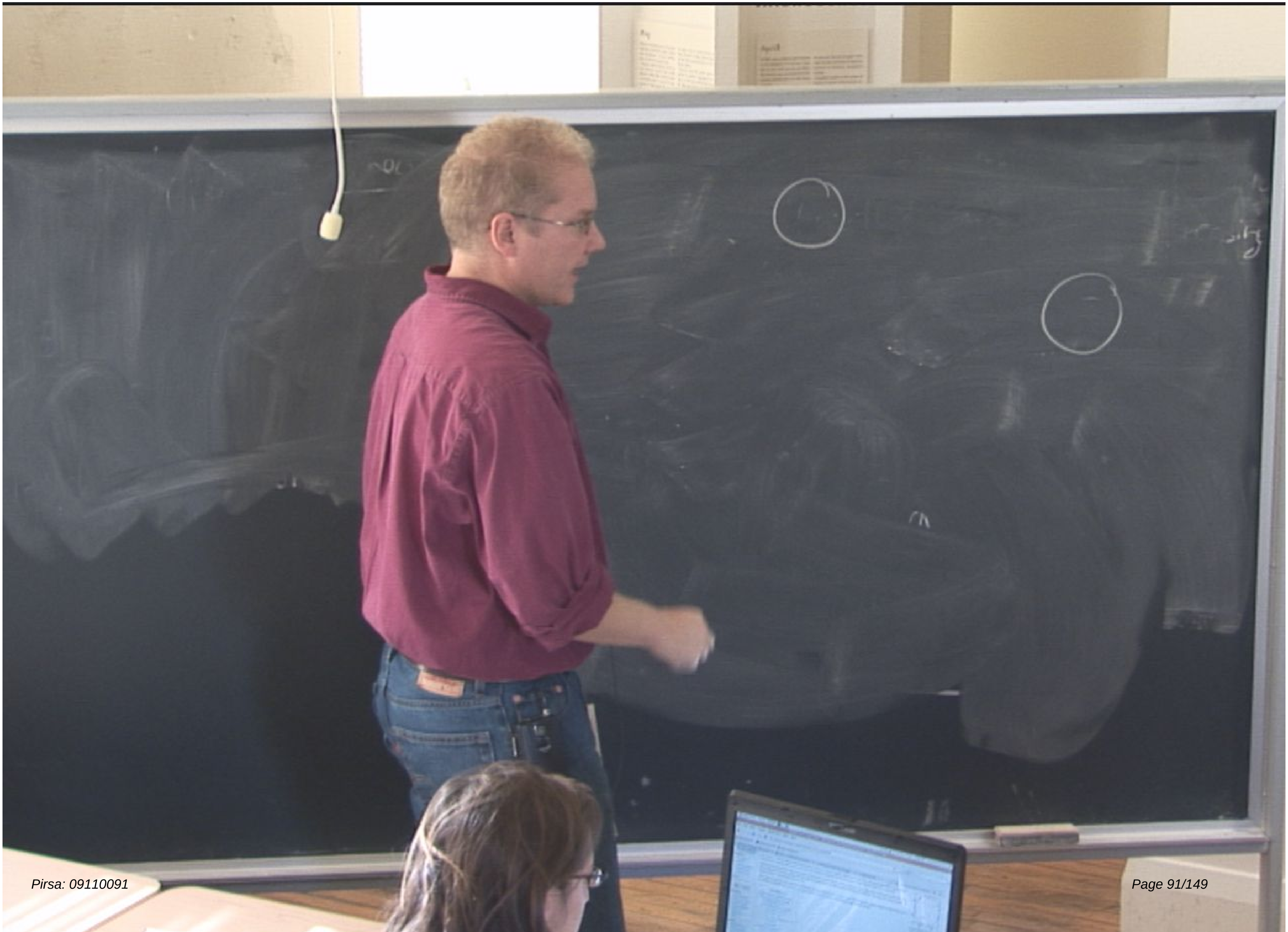
- Artificial time

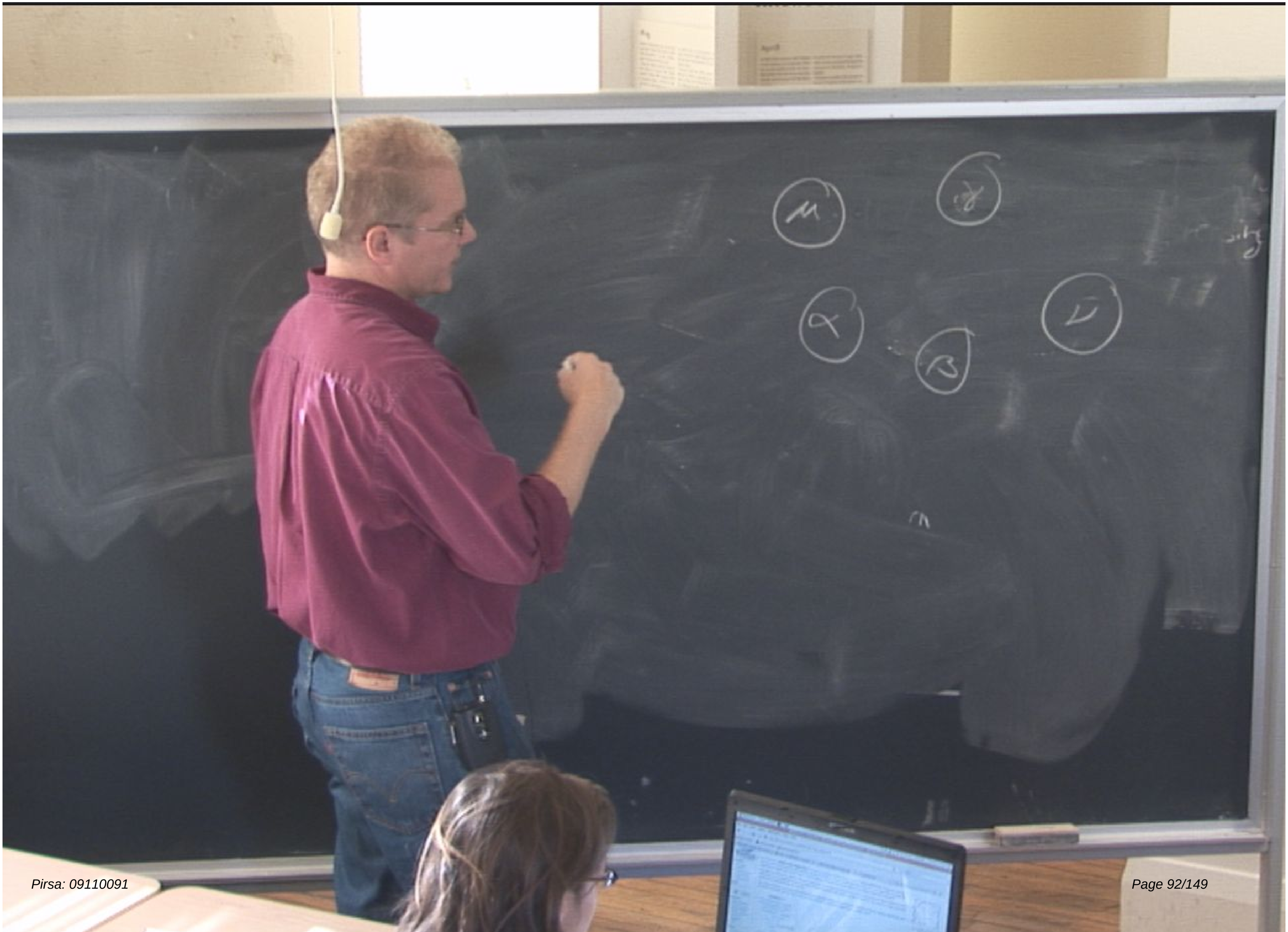
MR^2T^2 1953 S. Ulam

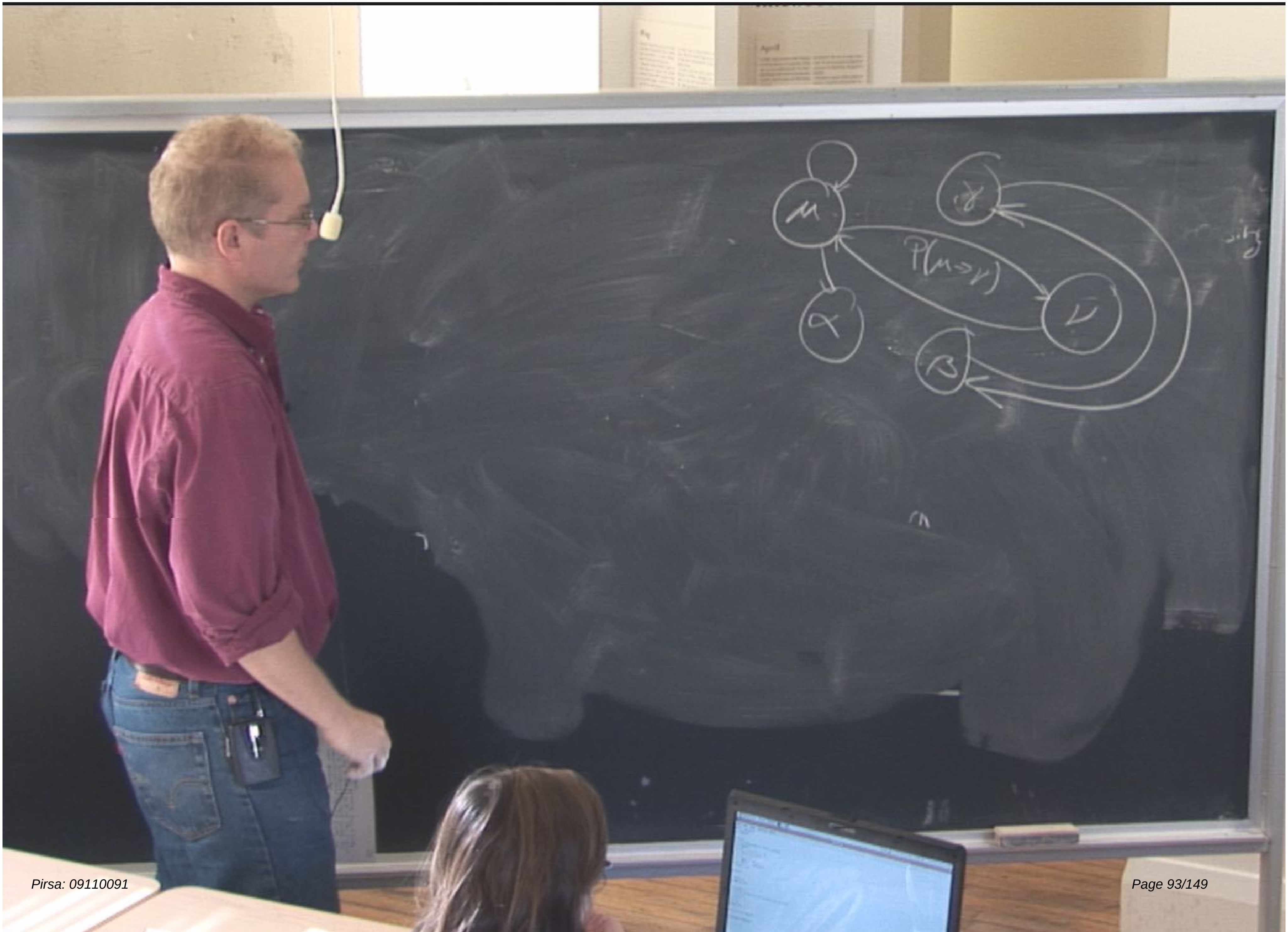
Idea: Replace the problem with another problem from Markov Chains that gives the same averages.

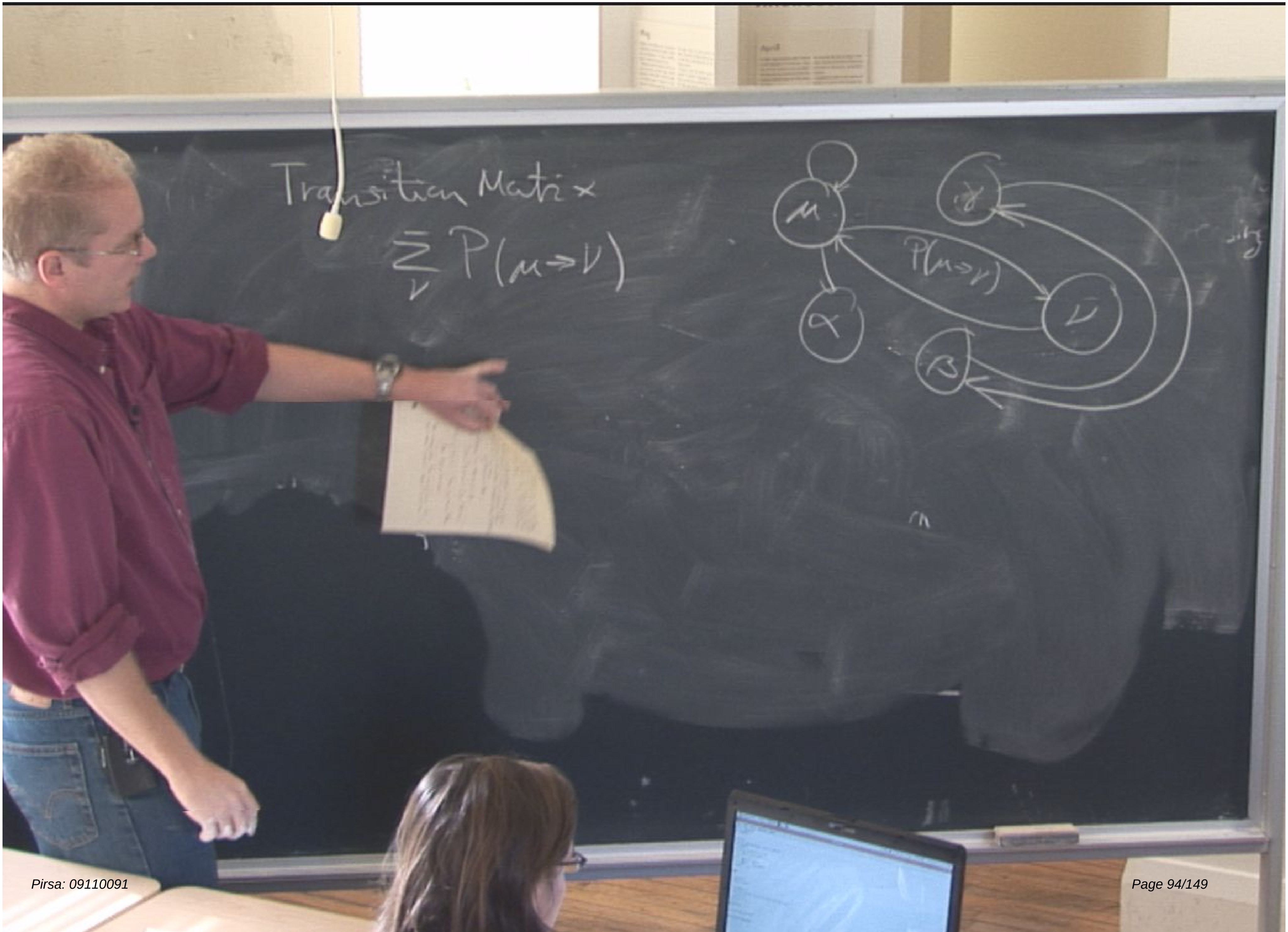
Setup a random walk on the x-axis or in the space of configurations where x_i / configs are visited with the desired distrib.

- Artificial time

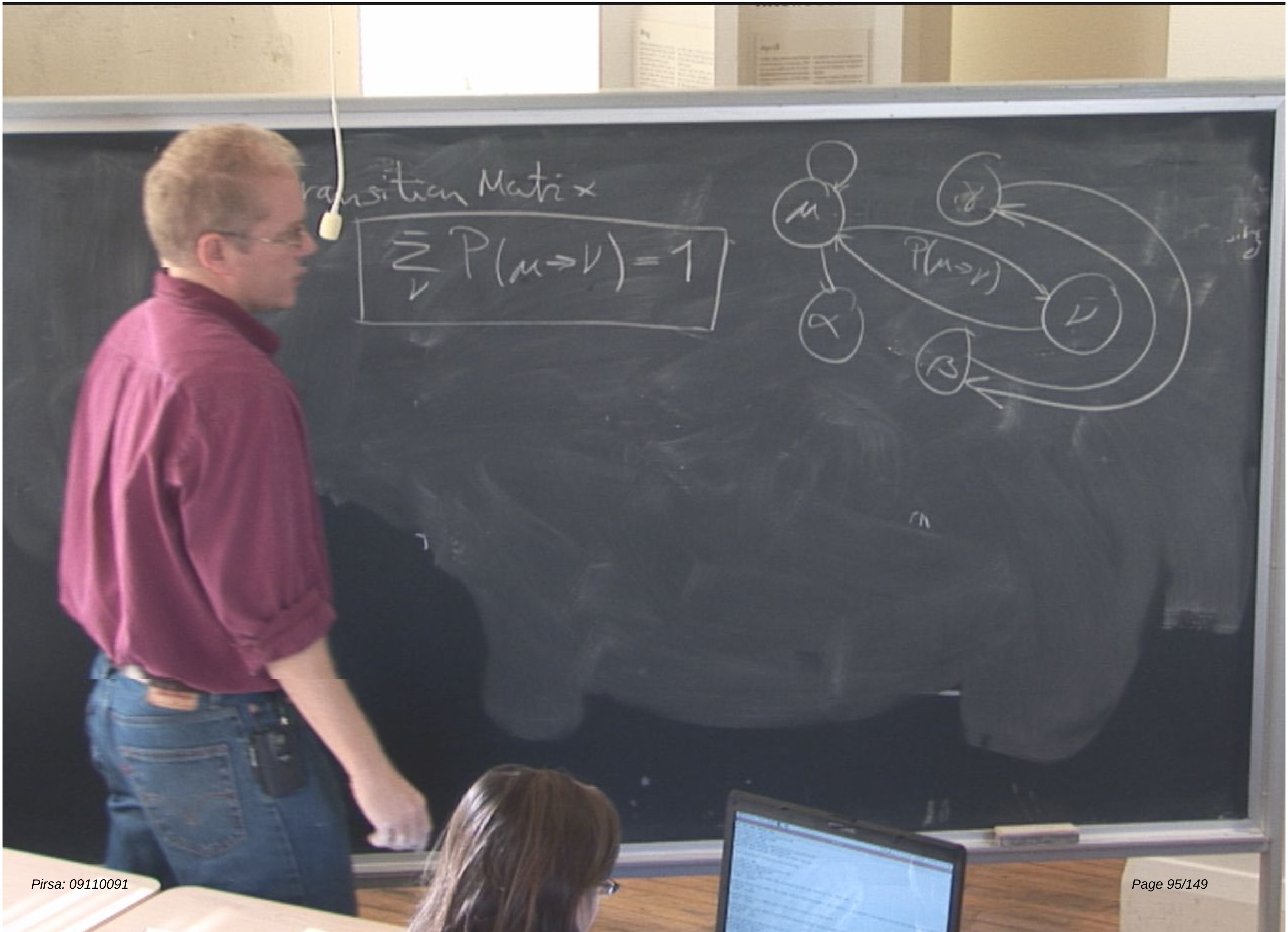


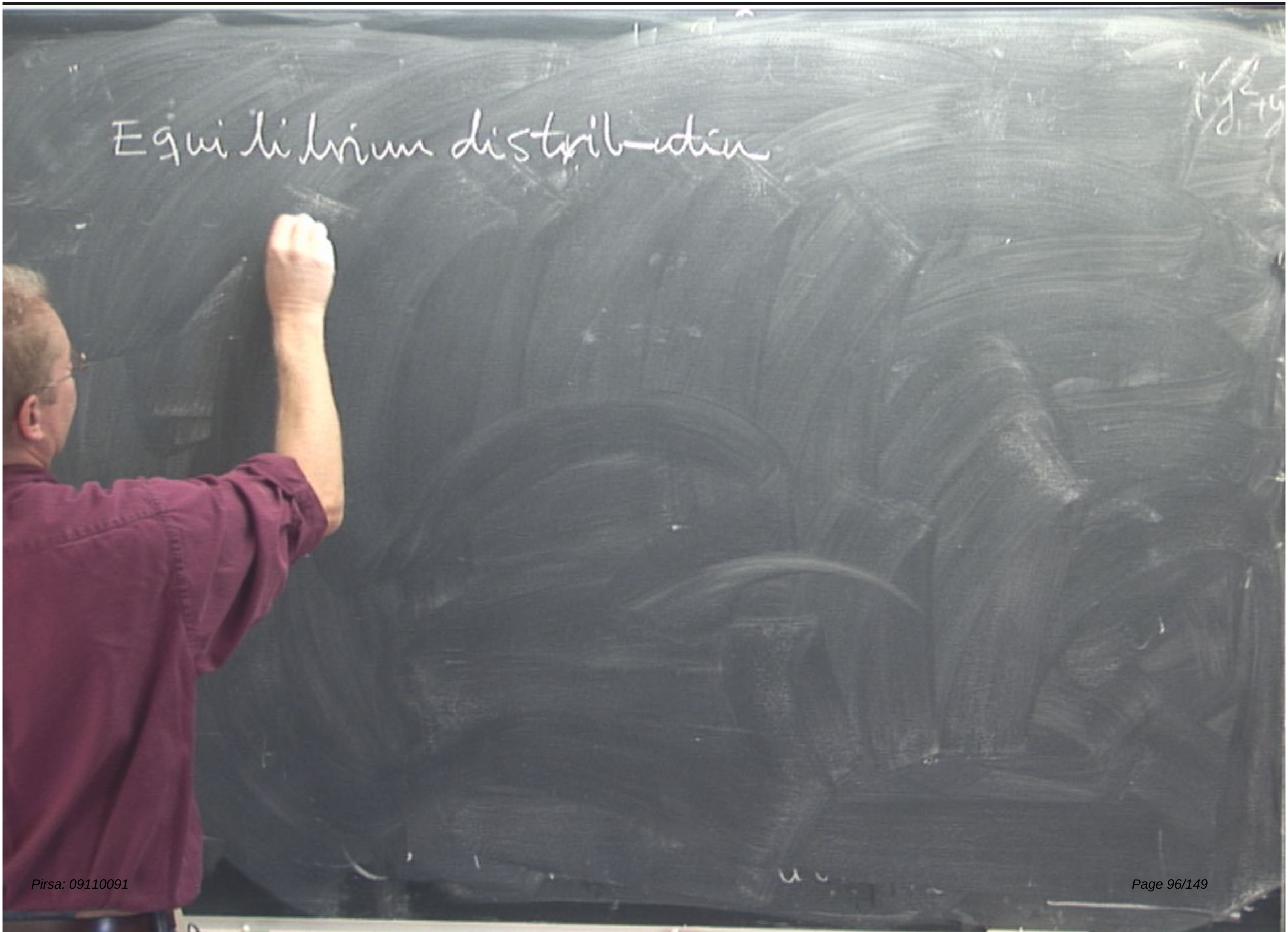



μ
γ
P(μ→ν)
ν
α
β



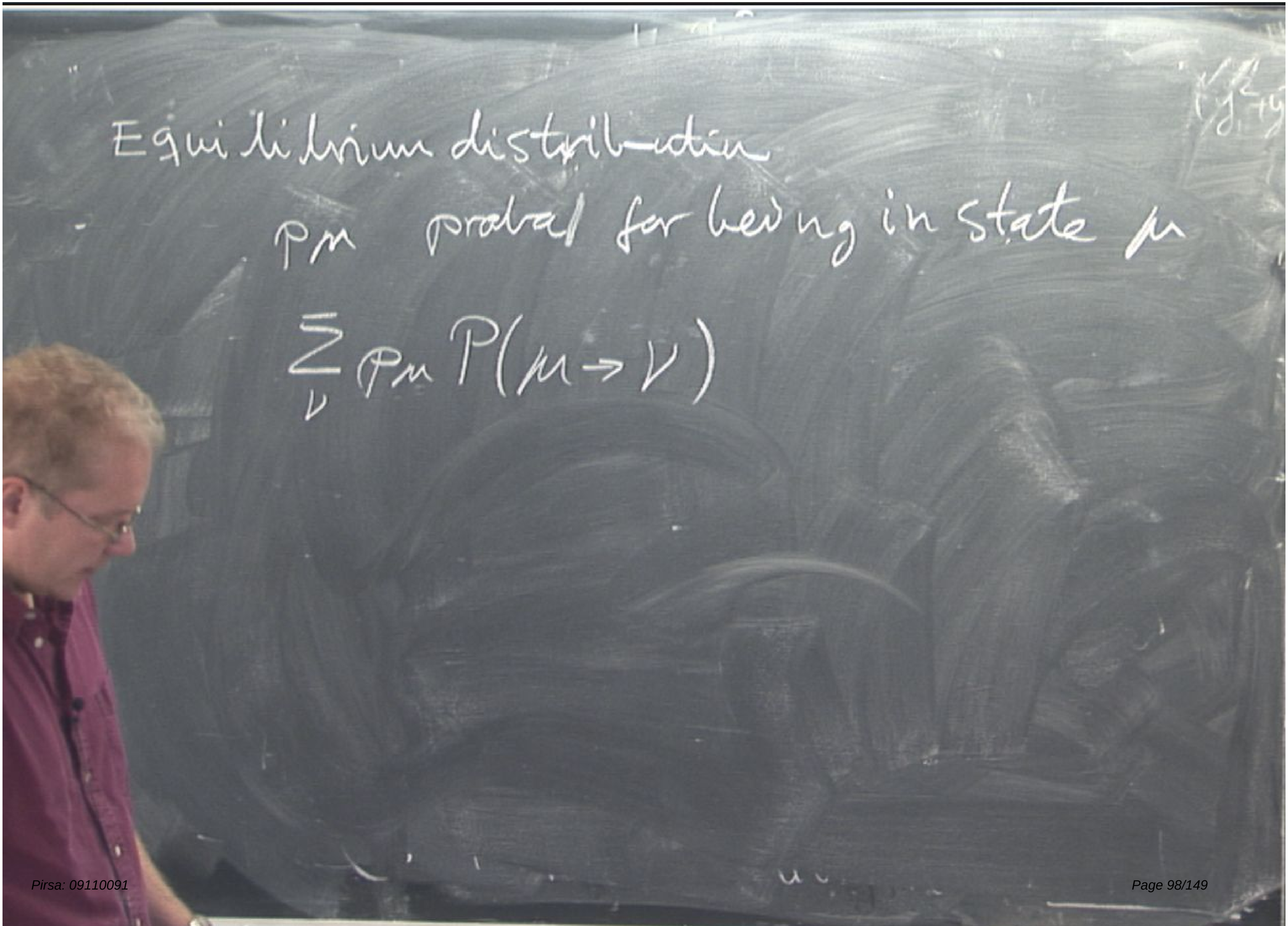
Transition Matrix
$\sum_\nu P(\mu \Rightarrow \nu)$
$P(\mu \Rightarrow \nu)$
μ
ν
α
β
γ

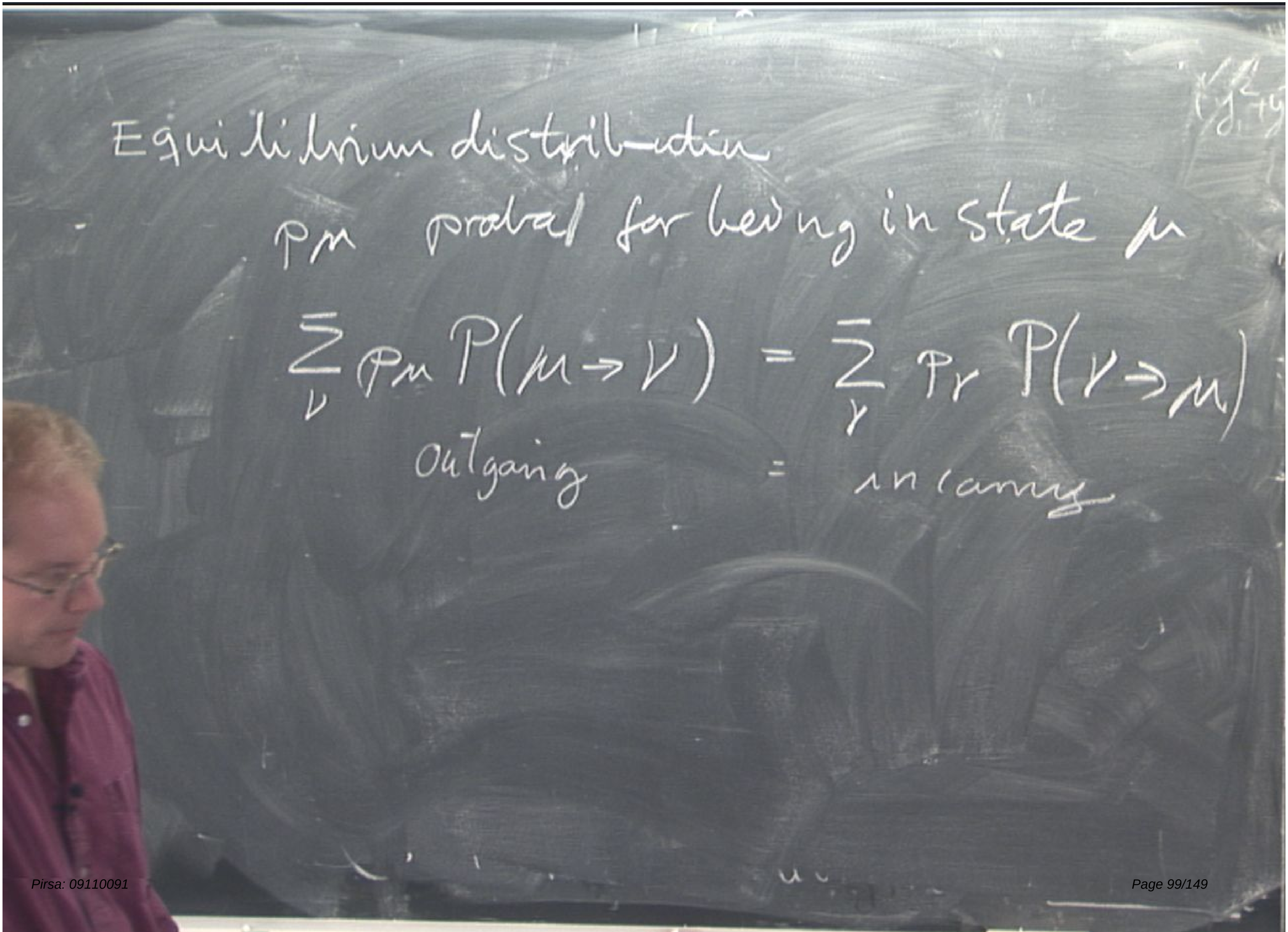

Matrix
$\sum_{\nu} P(\mu \to \nu) = 1$
μ
α
γ
ν
β
$P(\mu \to \nu)$


Equilibrium distribution

Equilibrium distribution

- p_μ probab for being in state μ


Equilibrium distribution
p_μ probab for being in state μ
$\sum_\nu p_\mu P(\mu \to \nu)$


Equilibrium distribution
p_μ probab for being in state μ
$$\sum_\nu p_\mu P(\mu \to \nu) = \sum_\nu p_\nu P(\nu \to \mu)$$
outgoing = incoming

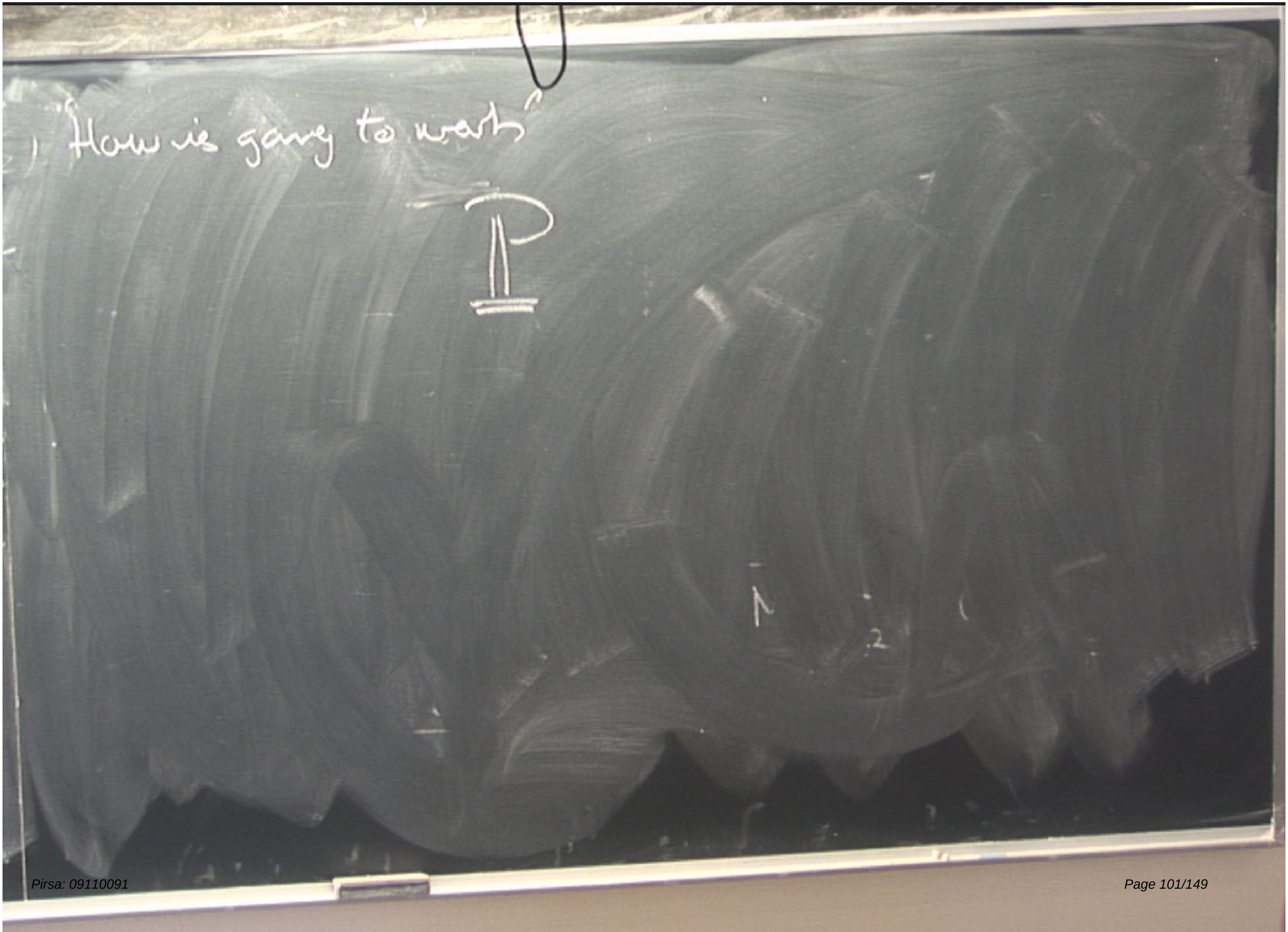
Equilibrium distribution

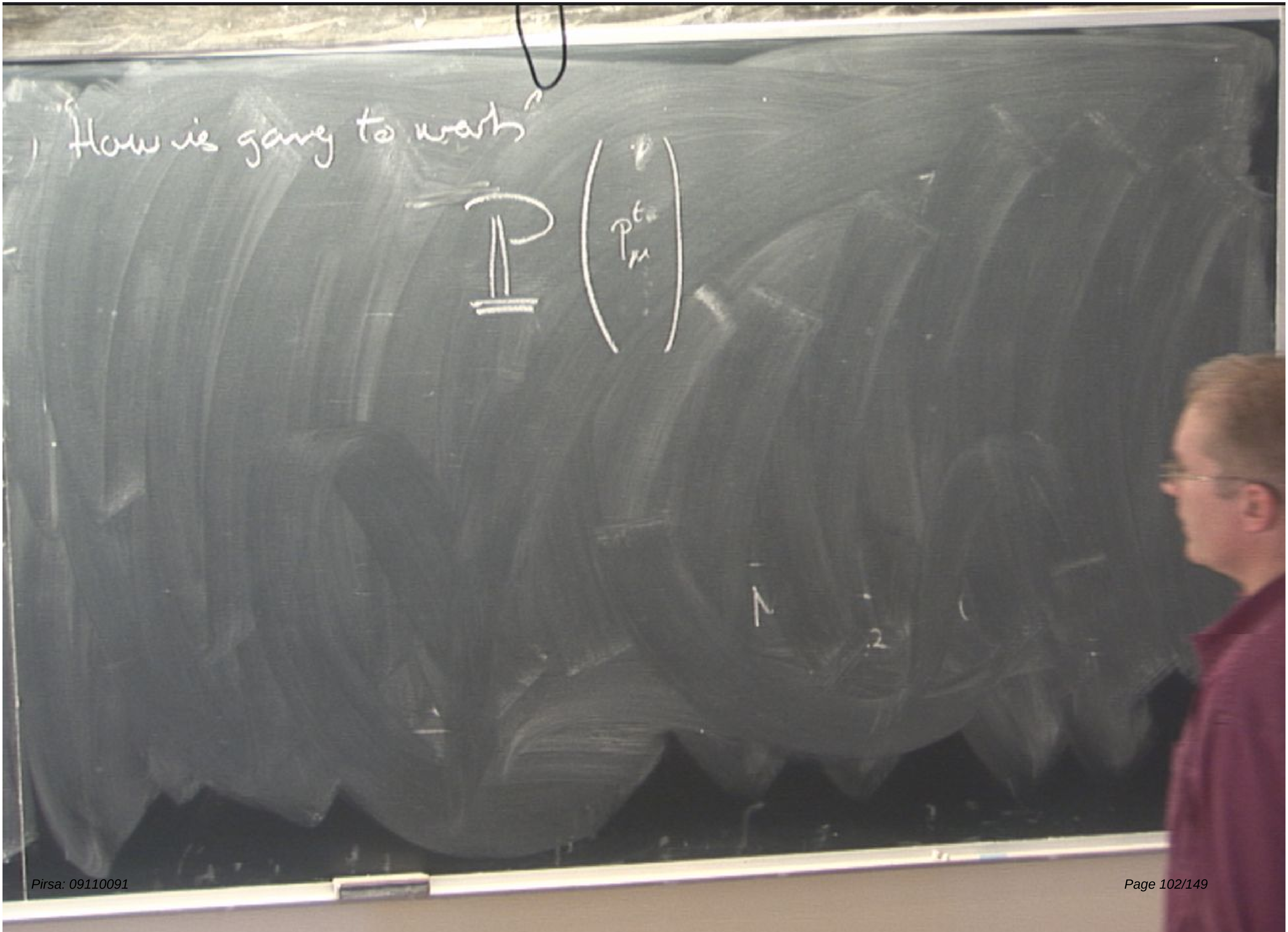
- p_μ probab for being in state μ

$$\sum_\nu p_\mu P(\mu \to \nu) = \sum_\nu p_\nu P(\nu \to \mu)$$

outgoing = incoming

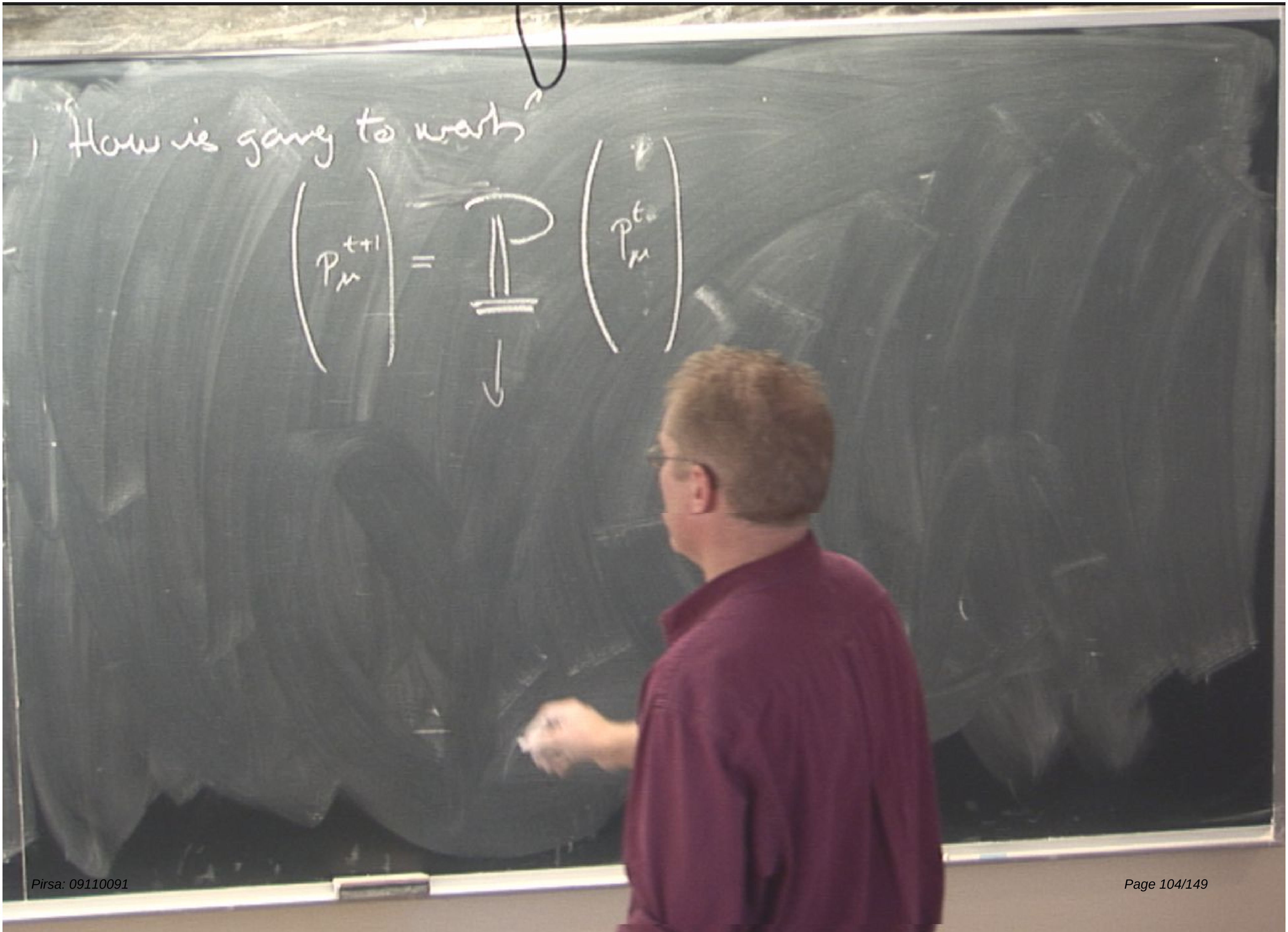
$$\boxed{p_\mu = \sum_\nu p_\nu P(\nu \to \mu)}$$


How is going to work?


How is going to work?
$\mathbb{P}\left(p_\mu^{t}\right)$

How is going to work?

$$\begin{pmatrix} \\ p_\mu^{t+1} \\ \\ \end{pmatrix} = \mathbb{P} \begin{pmatrix} \\ p_\mu^{t} \\ \\ \end{pmatrix}$$


How is going to work
$\left(p_\mu^{t+1} \right) = \mathbb{P} \left(p_\mu^{t} \right)$

How is going to work?

$$\begin{pmatrix} p_\mu^{t+1} \end{pmatrix} = \mathbb{P} \begin{pmatrix} p_\mu^{t} \end{pmatrix}$$

$$\downarrow$$

$$\begin{pmatrix} p^{*} \end{pmatrix} = \mathbb{P} \begin{pmatrix} p^{*} \end{pmatrix}$$

How is going to work?

$$\begin{pmatrix} p_\mu^{t+1} \end{pmatrix} = \mathbb{P} \begin{pmatrix} p_\mu^{t} \end{pmatrix}$$

$$\downarrow$$

$$\begin{pmatrix} p^{*} \end{pmatrix} = \mathbb{P} \begin{pmatrix} p^{*} \end{pmatrix}$$

Equilibrium

How is going to work?

$$\left(p_\mu^{t+1}\right) = \mathbb{P}\left(p_\mu^{t}\right)$$

$$\downarrow$$

Equilibrium

$$\left(p^{*}\right) = \mathbb{P}\left(p^{*}\right)$$

limit $\left(p^{*}\right) = \mathbb{P}^{n}\left(p^{*}\right)$ limit cycles

How is going to work?

$$\begin{pmatrix} p_\mu^{t+1} \end{pmatrix} = \mathbb{P} \begin{pmatrix} p_\mu^{t} \end{pmatrix}$$

$$\downarrow$$

Equilibrium

$$\begin{pmatrix} p^{*} \end{pmatrix} = \mathbb{P} \begin{pmatrix} p^{*} \end{pmatrix}$$

limit $\begin{pmatrix} p^{*} \end{pmatrix} = \mathbb{P}^{n} \begin{pmatrix} p^{*} \end{pmatrix}$ limit (cycle)

avoid

How is going to work?

$$\begin{pmatrix} p_\mu^{t+1} \end{pmatrix} = \underline{\underline{\mathbb{P}}} \begin{pmatrix} p_\mu^{t} \end{pmatrix}$$

$\downarrow$

$$\begin{pmatrix} p^* \end{pmatrix} = \underline{\mathbb{P}} \begin{pmatrix} p^* \end{pmatrix}$$

Equilibrium

want

$$\begin{pmatrix} p^* \end{pmatrix} \neq \underline{\mathbb{P}}^n \begin{pmatrix} p^* \end{pmatrix}$$

limit cycles

avoid

How is going to work?

$$\begin{pmatrix} p_\mu^{t+1} \end{pmatrix} = \mathbb{P} \begin{pmatrix} p_\mu^{t} \end{pmatrix}$$

$$\downarrow$$

$$\begin{pmatrix} p^* \end{pmatrix} = \mathbb{P} \begin{pmatrix} p^* \end{pmatrix}$$

Equilibrium

want $\cancel{\begin{pmatrix} p^* \end{pmatrix} = \mathbb{P}^n \begin{pmatrix} p^* \end{pmatrix}}$ limit cycles

avoid

Impose detailed balance

$$p_\mu P(\mu \to \nu) = p_\nu P(\nu \to \mu)$$

Impose detailed balance

$$p_\mu P(\mu \to \nu) = p_\nu P(\nu \to \mu)$$

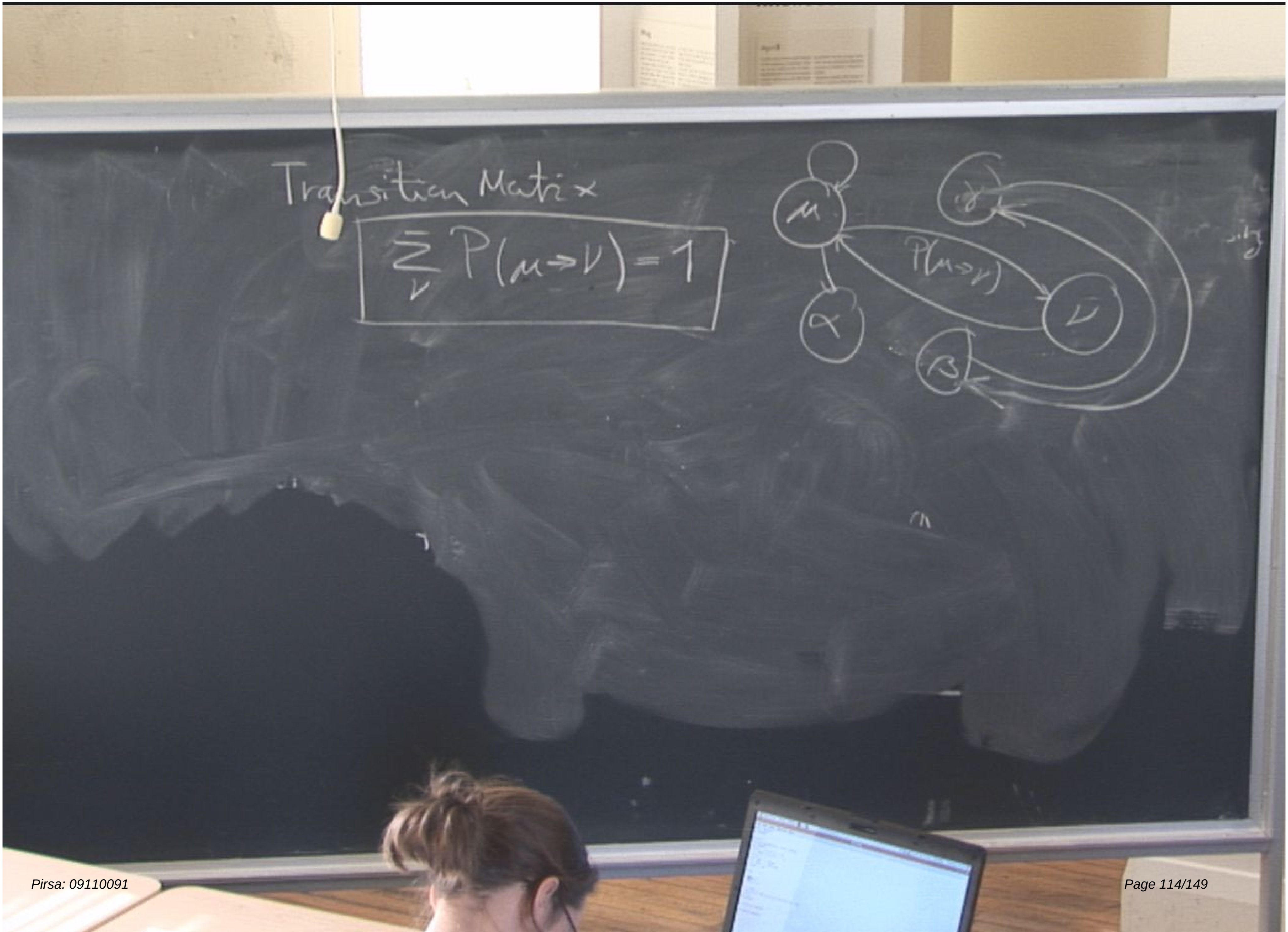
sufficient condition

Impose detailed balance

$$p_\mu P(\mu \to \nu) = p_\nu P(\nu \to \mu)$$

sufficient condition

Ergodicity


Transition Matrix
$\sum_{\nu} P(\mu \to \nu) = 1$
$P(\mu \to \nu)$
μ
α
ν
β
γ

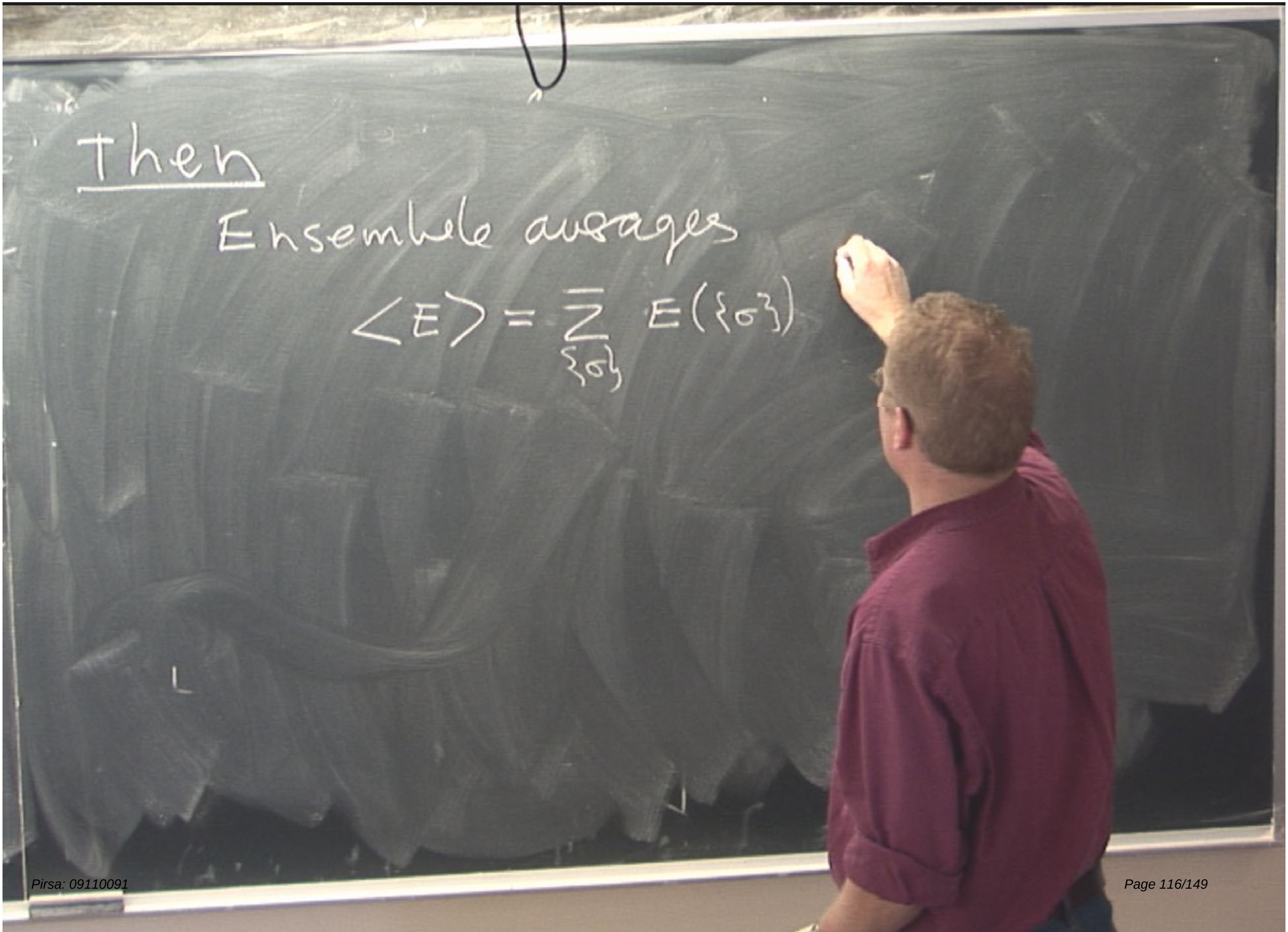
Impose detailed balance

$$p_\mu P(\mu \to \nu) = p_\nu P(\nu \to \mu)$$

sufficient condition

Ergodicity

All points/states are reachable from any starting points


Then
Ensemble averages
$\langle E \rangle = \sum_{\{\sigma\}} E(\{\sigma\})$

Then

Ensemble averages

$$\langle E \rangle = \sum_{\{\sigma\}} E(\{\sigma\}) \frac{e^{-E(\{\sigma\})/kT}}{Z}$$

Then

Ensemble averages

$$\langle E \rangle = \sum_{\{\sigma\}} E(\{\sigma\}) \, \frac{e^{-E(\{\sigma\})/kT}}{Z}$$

becomes equivalent to a time average

Then

Ensemble averages

$$\langle E \rangle = \sum_{\{\sigma\}} E(\{\sigma\}) \frac{e^{-E(\{\sigma\})/kT}}{Z}$$

becomes equivalent to a time average

$$= \sum_{t}$$

Then

Ensemble averages

$$\langle E \rangle = \sum_{\{\sigma\}} E(\{\sigma\}) \frac{e^{-E(\{\sigma\})/kT}}{Z}$$

becomes equivalent to a time average

$$\langle E \rangle = \frac{1}{M} \sum_{t} E(\{\sigma^t\})$$

Then

Ensemble averages

$$\langle E \rangle = \sum_{\{\sigma\}} E(\{\sigma\}) \frac{e^{-E(\{\sigma\})/kT}}{Z}$$

becomes equivalent to a time average

$$\langle E \rangle = \frac{1}{M} \sum_{t=1}^{M} E(\{\sigma^t\})$$

Then

Ensemble averages

$$\langle E \rangle = \sum_{\{\sigma\}} E(\{\sigma\}) \frac{e^{-E(\{\sigma\})/kT}}{Z}$$

becomes equivalent to a time average

$$\langle E \rangle = \frac{1}{M} \sum_{t=1}^{M} E(\{\sigma^t\})$$

Then

Ensemble averages

$$\langle E \rangle = \sum_{\{\sigma\}} E(\{\sigma\}) \, \frac{e^{-E(\{\sigma\})/kT}}{Z}$$

becomes equivalent to a time average

$$\langle E \rangle = \frac{1}{M} \sum_{t=1}^{M} E(\{\sigma^t\})$$

Define $\mathbb{P}$

Define $\mathbb{P}$

Not completely specified

$$\frac{P(\mu \Rightarrow \nu)}{P(\nu \Rightarrow \mu)} = \frac{p_\nu}{p_\mu}$$

Define $\mathbb{P}$

Not completely specified

$$\frac{P(\mu \Rightarrow \nu)}{P(\nu \Rightarrow \mu)} = \frac{p_\nu}{p_\mu}$$

only the ratio that is defined


Boltzman distrib.

- Boltzman distrib.

$$\frac{P(\mu \to \nu)}{P(\nu \to \mu)} = \frac{p_\mu}{p_\nu} = e^{-\beta(E_\nu - E_\mu)}$$

- Boltzman distrib.

$$\frac{P(\mu\to\nu)}{P(\nu\to\mu)} = \frac{p_\mu}{p_\nu} = e^{-\beta(E_\nu - E_\mu)}$$

(1) $$P(\mu\to\nu) = \frac{e^{-\frac{1}{2}\beta(E_\nu - E_\mu)}}{e^{-\frac{1}{2}\beta(E_\nu - E_\mu)} + e^{\frac{1}{2}\beta(E_\nu - E_\mu)}}$$

Boltzman distrib.

$$\frac{P(\mu\to\nu)}{P(\nu\to\mu)} = \frac{p_\mu}{p_\nu} = e^{-\beta(E_\nu - E_\mu)}$$

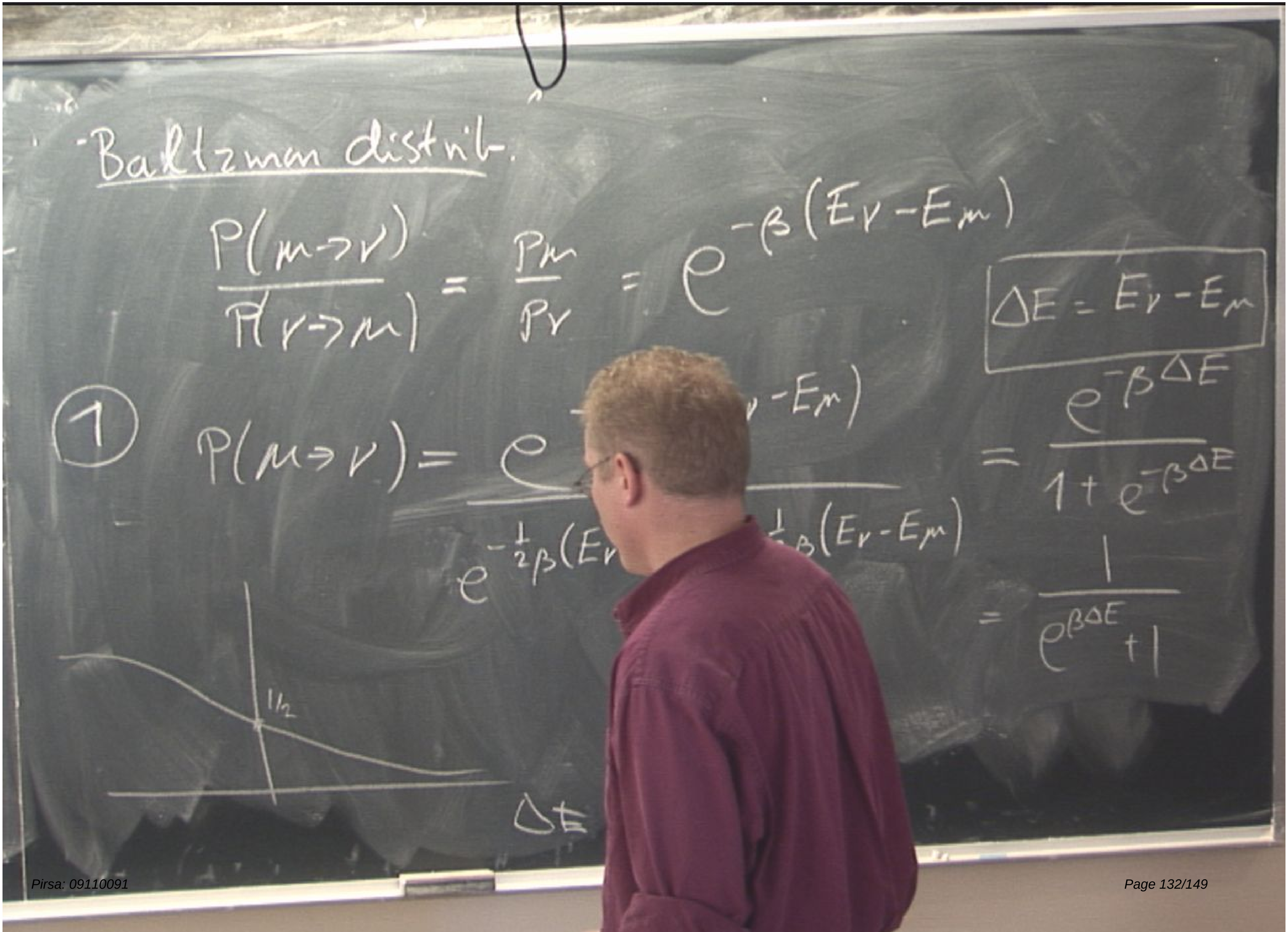
① $$P(\mu\to\nu) = \frac{e^{-\frac{1}{2}\beta(E_\nu - E_\mu)}}{e^{-\frac{1}{2}\beta(E_\nu - E_\mu)} + e^{\frac{1}{2}\beta(E_\nu - E_\mu)}} = \frac{e^{-\beta\Delta E}}{1+e^{-\beta\Delta E}} = \frac{1}{e^{\beta\Delta E}+1}$$

- Baltzman distrib.

$$\frac{P(\mu\to\nu)}{P(\nu\to\mu)} = \frac{p_\mu}{p_\nu} = e^{-\beta(E_\nu - E_\mu)}$$

$$\boxed{\Delta E = E_\nu - E_\mu}$$

(1) $$P(\mu\to\nu) = \frac{e^{-\frac{1}{2}\beta(E_\nu - E_\mu)}}{e^{-\frac{1}{2}\beta(E_\nu - E_\mu)} + e^{\frac{1}{2}\beta(E_\nu - E_\mu)}} = \frac{e^{-\beta\Delta E}}{1 + e^{-\beta\Delta E}} = \frac{1}{e^{\beta\Delta E} + 1}$$


Boltzman distrib.
$\frac{P(\mu\to\nu)}{P(\nu\to\mu)} = \frac{p_\mu}{p_\nu} = e^{-\beta(E_\nu - E_\mu)}$
$\Delta E = E_\nu - E_\mu$
1
$P(\mu\to\nu) =$
$-E_\mu)$
$e^{-\frac{1}{2}\beta(E_\nu}$
$\beta(E_\nu - E_\mu)$
$= \frac{e^{-\beta\Delta E}}{1+e^{-\beta\Delta E}}$
$= \frac{1}{e^{\beta\Delta E}+1}$
1/2
ΔE

$$\Delta E = E_\nu - E_\mu$$

②

$$P(\mu \to \nu) = \min(1, e^{-\beta \Delta E})$$

$\Delta E = E_\nu - E_\mu$

(2)

$$P(\mu \to \nu) = \min(1, e^{-\beta \Delta E})$$

$\Delta E < 0 \quad P(\mu \to \nu) = 1$

$$\Delta E = E_\nu - E_\mu$$

②

$$P(\mu \to \nu) = \min(1, e^{-\beta \Delta E})$$

$\Delta E < 0$ $\quad P(\mu \to \nu) = 1$

but then $P(\nu \to \mu)$

$$\Delta E = E_\nu - E_\mu$$

(2)

$$P(\mu \to \nu) = \min\left(1, e^{-\beta \Delta E}\right)$$

$\Delta E < 0$ $\quad P(\mu \to \nu) = 1$

but then $P(\nu \to \mu) = e^{\beta \Delta E}$

$$\frac{P(\mu \to \nu)}{P(\nu \to \mu)} = e^{-\beta \Delta E}$$

$\Delta E = E_\nu - E_\mu$

②

$$P(\mu \to \nu) = \min(1, e^{-\beta \Delta E})$$

$\Delta E < 0$ $\quad P(\mu \to \nu) = 1$

but then $P(\nu \to \mu) = e^{\beta \Delta E}$ $\quad e^{-\beta \Delta E}$

$\Delta E > 0$ $\quad P(\mu \to \nu) = e^{-\beta \Delta E}$

$P(\nu \to \mu) = 1$ $\quad e^{-\beta \Delta E}$

$\Delta E = E_\nu - E_\mu$

(2)

$$P(\mu \to \nu) = \min(1, e^{-\beta \Delta E})$$

$\underline{\Delta E < 0}$ $\quad P(\mu \to \nu) = 1$

but then $P(\nu \to \mu) = e^{\beta \Delta E}$

$$\frac{P(\mu \to \nu) = 1}{P(\nu \to \mu) = e^{\beta \Delta E}} = e^{-\beta \Delta E}$$

$\Delta E > 0$

$$\frac{P(\mu \to \nu) = e^{-\beta \Delta E}}{P(\nu \to \mu) = 1} = e^{-\beta \Delta E}$$

$$\Delta E = E_\nu - E_\mu$$

(2)

$$P(\mu \to \nu) = \min\left(1, e^{-\beta \Delta E}\right)$$

$\Delta E < 0$

$$\frac{P(\mu \to \nu) = 1}{\text{but then } P(\nu \to \mu) = e^{\beta \Delta E}} = e^{-\beta \Delta E}$$

$\Delta E > 0$

$$\frac{P(\mu \to \nu) = e^{-\beta \Delta E}}{P(\nu \to \mu) = 1} = e^{-\beta \Delta E}$$

$\Delta E = E_\nu - E_\mu$

(2)

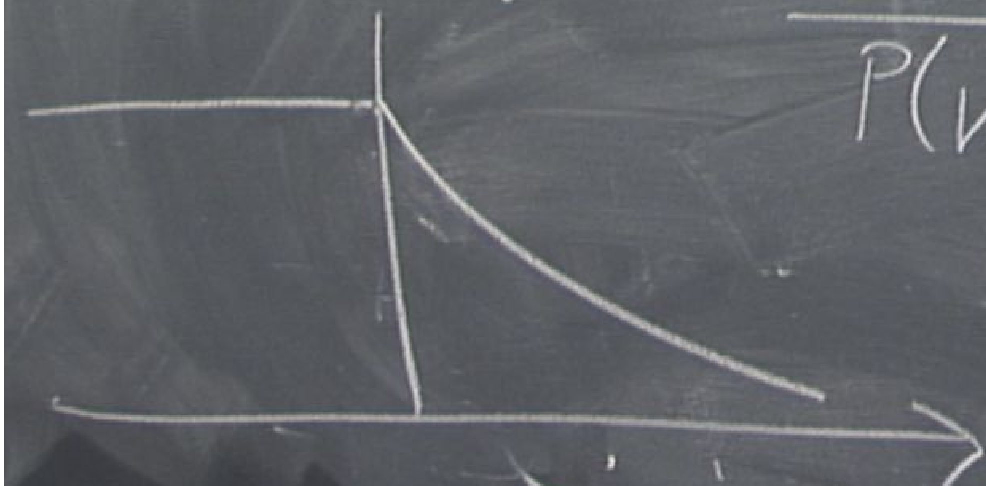
$$P(\mu \to \nu) = \min\left(1, e^{-\beta \Delta E}\right)$$

$\Delta E < 0$

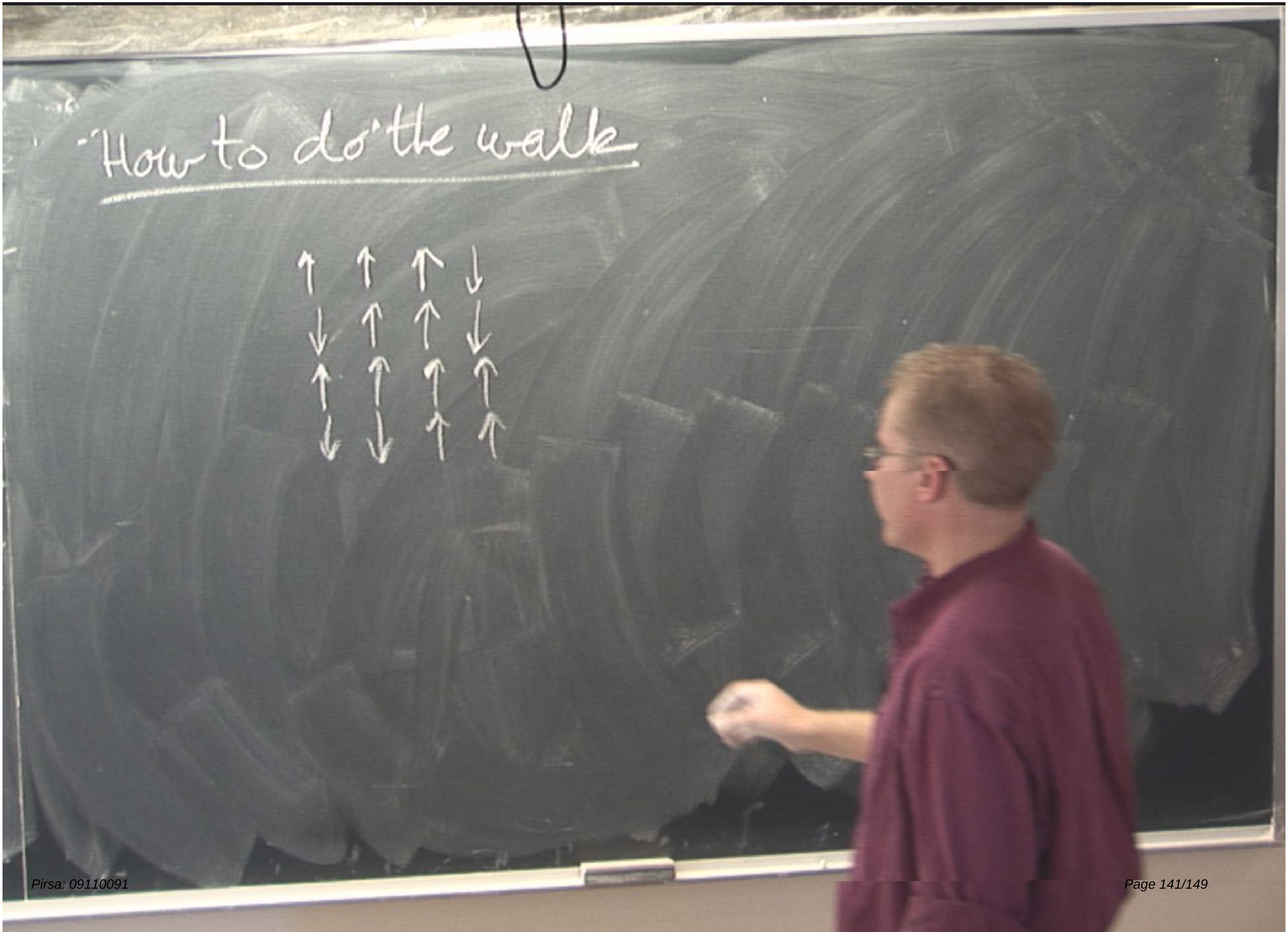
$$\frac{P(\mu \to \nu) = 1}{\text{but then } P(\nu \to \mu) = e^{\beta \Delta E}} = e^{-\beta \Delta E}$$

$\Delta E > 0$

$$\frac{P(\mu \to \nu) = e^{-\beta \Delta E}}{P(\nu \to \mu) = 1} = e^{-\beta \Delta E}$$



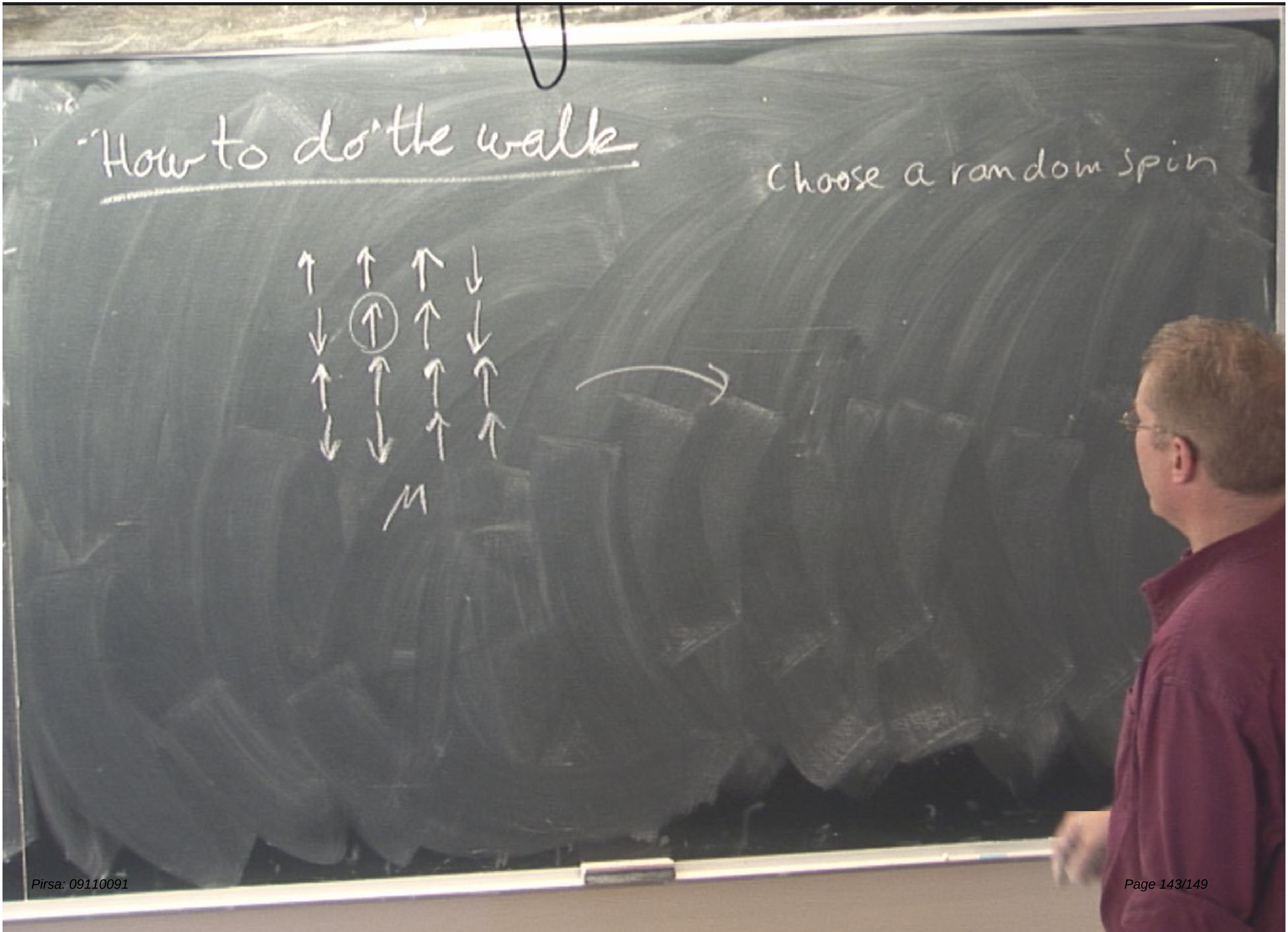
compare

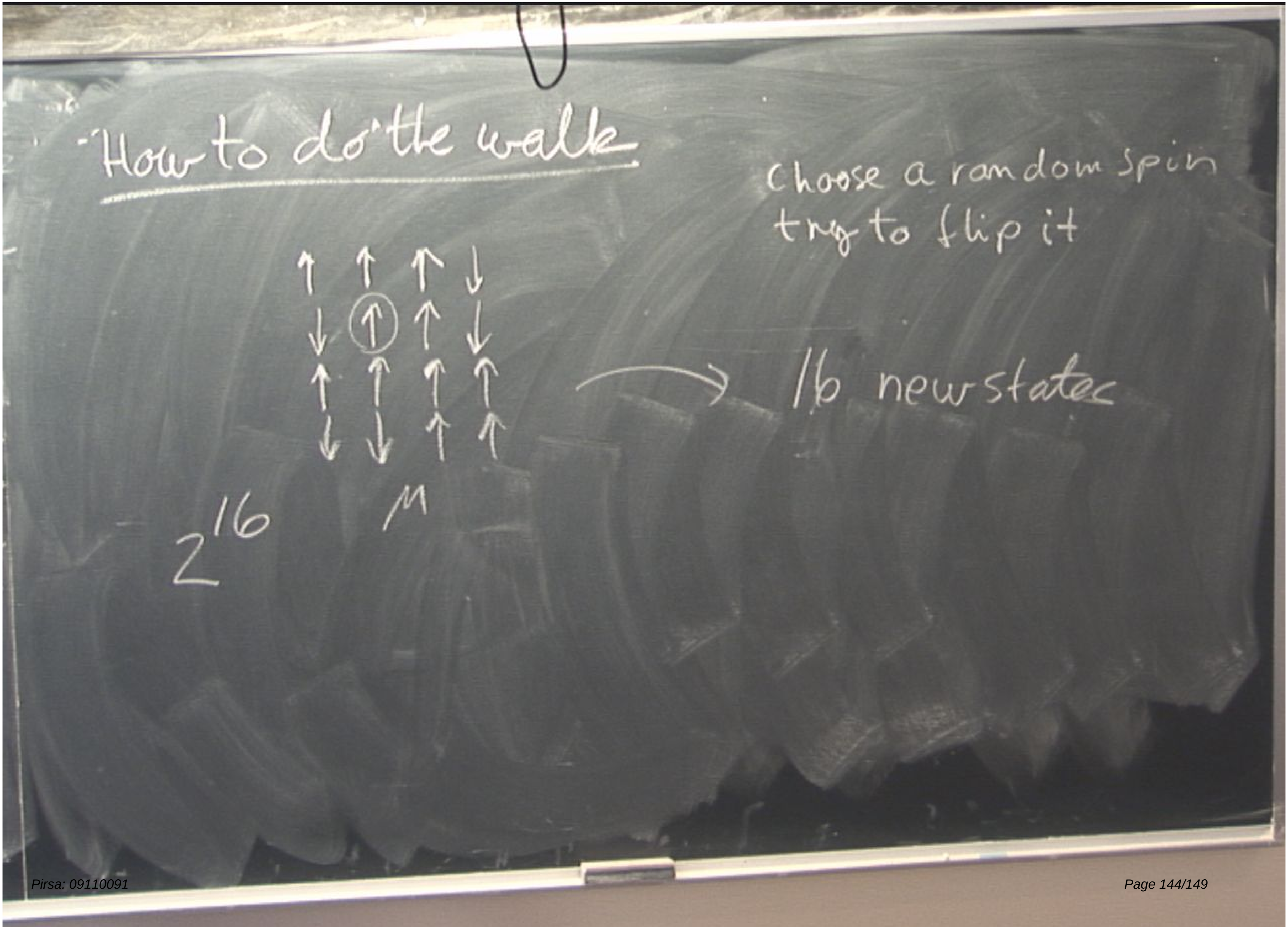

How to do the walk

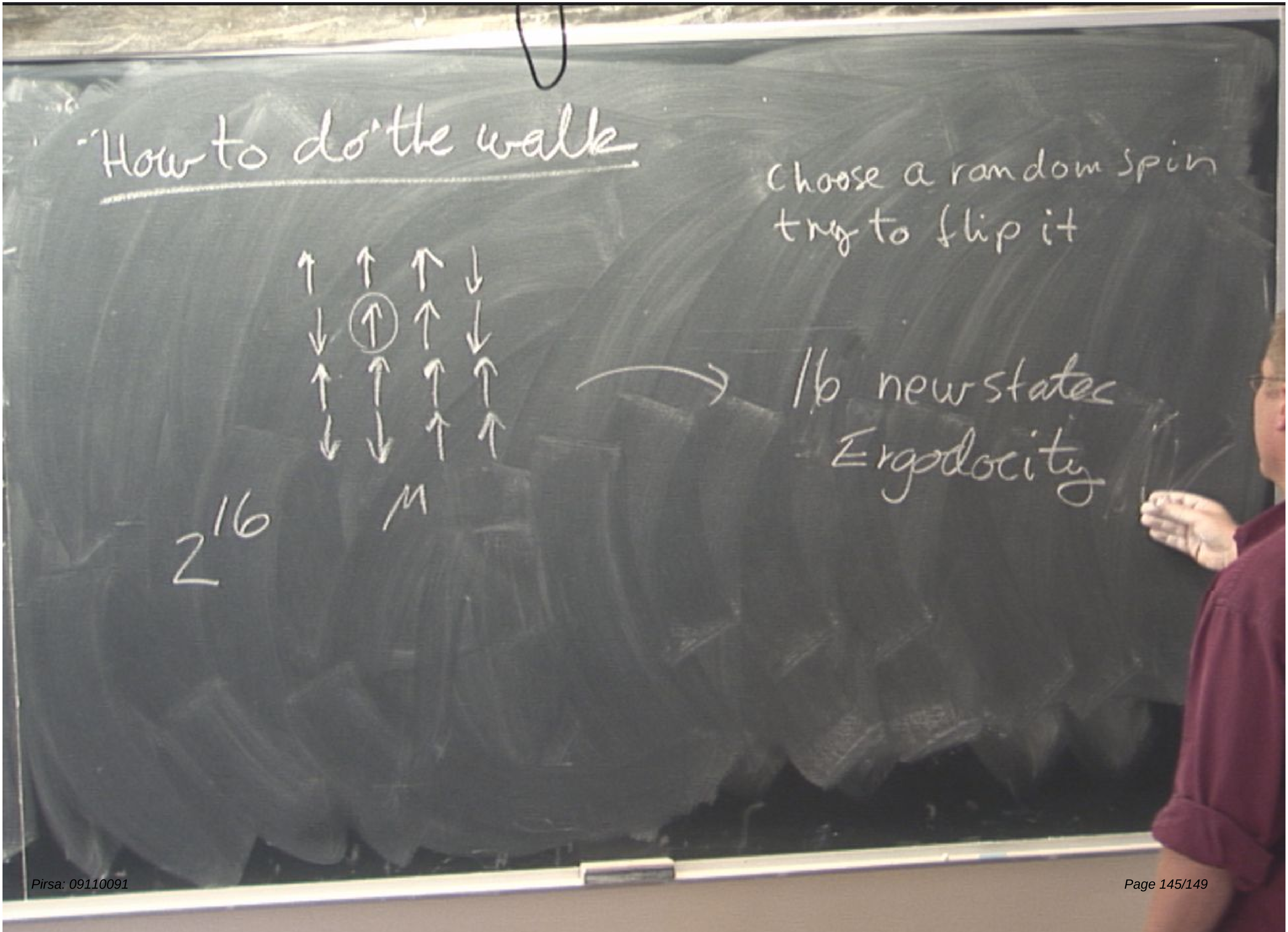
How to do the walk

$$\begin{array}{cccc} \uparrow & \uparrow & \uparrow & \downarrow \\ \downarrow & \uparrow & \uparrow & \downarrow \\ \uparrow & \uparrow & \uparrow & \uparrow \\ \downarrow & \downarrow & \uparrow & \uparrow \end{array} \longrightarrow \nu\,?$$

μ


How to do the walk
Choose a random spin


How to do the walk
Choose a random spin
try to flip it
16 new states
2^{16}
M


How to do the walk
Choose a random spin
try to flip it
16 new states
Ergodocity
2^16
M

How to do the walk

Choose a random spin
try to flip it

$$\begin{matrix} \uparrow & \uparrow & \uparrow & \downarrow \\ \downarrow & \uparrow & \uparrow & \downarrow \\ \uparrow & \uparrow & \uparrow & \uparrow \\ \downarrow & \downarrow & \uparrow & \uparrow \end{matrix} \longrightarrow \text{16 new states}$$

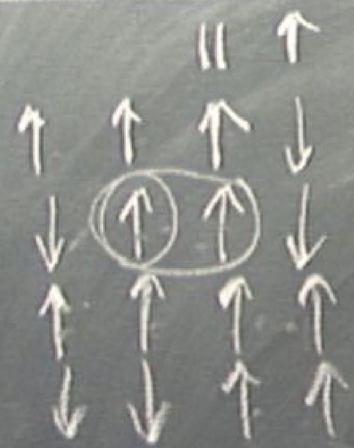
Ergodocity OK

2^{16}

M

$$\begin{matrix} \uparrow & \uparrow & \uparrow & \downarrow \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \uparrow & \uparrow & \uparrow & \uparrow \\ \downarrow & \downarrow & \uparrow & \uparrow \end{matrix}$$

How to do the walk



M

2^{16}

Choose a random spin
try to flip it

16 new states
Ergodocity OK

How to do the walk

11 ↑ 5 ↓

↑ ↑ ↑ ↓
↓ ↑ ↑ ↓
↑ ↑ ↑ ↑
↓ ↓ ↑ ↑

M

2^{16}

Choose a random spin
try to flip it

16 new states
Ergodocity OK

↑ ↑ ↑ ↓
↓ ↓ ↓ ↓
↑ ↑ ↑ ↑
↓ ↓ ↑ ↑

9↑ 7↓

parity of ↑
is conserved

How to do the walk

Choose a random spin
try to flip it

$\boxed{\text{Slow!}}$

16 new states
Ergodocity OK

11 ↑ 5 ↓

↑ ↑ ↑ ↓
↓ ↑ ↑ ↓
↑ ↑ ↑ ↑
↓ ↓ ↑ ↑

M

2^{16}

lots of freedom

↑ ↑ ↑ ↓
↓ ↓ ↓ ↓
↑ ↑ ↑ ↑
↓ ↓ ↑ ↑

9↑ 7↓
parity of ↑
is conserved