

Title: Statistical Mechanics (PHYS 602) - Lecture 9

Date: Oct 08, 2009 10:30 AM

URL: <http://pirsa.org/09100130>

Abstract:

# Probability distribution

$$\Gamma\tau = 2MkT$$

Whenever this relation is satisfied,  $p$  has the right variance,  $MkT$ , and the right Maxwell-Boltzmann probability distribution.

$$\rho(\mathbf{p}) = \left(\frac{\beta}{2\pi M}\right)^{3/2} \exp[-\beta p^2/(2M)]$$

More generally, if we have a Hamiltonian,  $H(\mathbf{p}, \mathbf{r})$ , for the one-particle system, the Maxwell-Boltzmann distribution takes the form

$$\rho(\mathbf{p}, \mathbf{r}) = \exp[-\beta H(\mathbf{p}, \mathbf{r})]/Z, \quad \text{v.7}$$

where, in the simplest case the Hamiltonian is

$$H(\mathbf{p}, \mathbf{r}) = p^2/(2M) + U(\mathbf{r})$$

Maxwell and Boltzmann expected that, in appropriate circumstances, if they waited long enough, a Hamiltonian system would get to equilibrium and they would end up with a Maxwell-Boltzmann probability distribution

Question: Should we not be able to derive this distribution from classical mechanics alone? Maybe we should have to assume that we must long enough to reach equilibrium? Anything more?

## Statistical and Hamiltonian Dynamics

We have that the equilibrium  $\rho = \exp(-\beta H)/Z$ . How can this arise from time dependence of a system? One very important possible time-dependence is given by Hamiltonian mechanics

$$\frac{dq_\alpha}{dt} = \frac{\partial \mathcal{H}}{\partial p_\alpha}$$
$$\frac{dp_\alpha}{dt} = -\frac{\partial \mathcal{H}}{\partial q_\alpha}$$

The simplest case is a particle moving in a potential field with a Hamiltonian

$$\mathcal{H} = \mathbf{p}^2/(2M) + U(\mathbf{r}) \quad \text{and consequently equations of motion}$$

$$\frac{d\mathbf{p}}{dt} = -\nabla U$$

$$\frac{d\mathbf{r}}{dt} = \mathbf{p}/M$$

The statistical mechanics of such situations is given by a probability density function  $\rho(\mathbf{p}, \mathbf{r}, t)$  such that the probability of finding the particle in a volume element  $d\mathbf{p} d\mathbf{r}$  about  $\mathbf{p}, \mathbf{r}$  at time  $t$  is  $\rho(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} d\mathbf{r}$ . The next question is, what is the time-dependence of this probability density? Or maybe, how do we get equilibrium statistical mechanics as a consequence of this

## Calculation concluded .....

$$\partial_t \rho(x, t) + \rho(x, t) \sum_j (\partial_{x_j} V_j) + \sum_j V_j \partial_{x_j} \rho(x, t) = 0$$

The Hamiltonian case is special. There are two kinds of coordinates  $x_j = q_\alpha$  with  $V_j = \partial_{p_\alpha} H$  and  $x_j = p_\alpha$  with  $V_j = -\partial_{q_\alpha} H$ . In that case, the divergence term  $(\partial_{q_\alpha} \partial_{p_\alpha} H - \partial_{p_\alpha} \partial_{q_\alpha} H) \rho = 0$ . This result, called **Liouville's** theorem, says that the size of the volume element is independent of time. As a result the probability density obeys a special equation, with no divergence term

$$\partial_t \rho(p, q, t) + \sum_\alpha [(\partial_{p_\alpha} H) \partial_{q_\alpha} - (\partial_{q_\alpha} H) \partial_{p_\alpha}] \rho(p, q, t) = 0 \quad \text{v.10}$$

The time derivative of any function of  $p$  and  $q$  is given in Hamiltonian mechanics by

$$dX(p, q)/dt = \sum_\alpha [(\partial_{p_\alpha} H) \partial_{q_\alpha} - (\partial_{q_\alpha} H) \partial_{p_\alpha}] X(p, q) \quad \text{v.11}$$

In particular if  $H$  is the Hamiltonian, assumed to be a function of the  $p$ 's and  $q$ 's but not containing any other  $t$ -dependence, then  $dH/dt=0$ , i.e.  $H$  is independent of time.



## Calculation Continued .....

$$\int dx g(x) \rho(x, t+dt) = \int dx g(x+dt V(x, t)) \rho(x, t)$$

expand to first order in dt

$$\int dx g(x) \rho(x, t) + dt \int dx g(x) \partial_t \rho(x, t) = \int dx g(x) \rho(x, t) + \int dx dt V(x, t) [d_x g(x)] \rho(x, t)$$

throw away the things that cancel against each other to get

$$\int dx g(x) \partial_t \rho(x, t) - \int dx V(x, t) [\partial_x g(x)] \rho(x, t) = 0$$

integrate by parts on the right hand side, using the fact that  $\rho(x, t)$  vanishes at  $x = \pm$  infinity

$$\int dx g(x) \{ \partial_t \rho(x, t) + \partial_x [V(x, t) \rho(x, t)] \} = 0$$

Notice that  $g(x)$  is arbitrary. If this left hand side is going to always vanish, the  $\{ \}$  must vanish. We then conclude that  $\partial_t \rho(x, t) + \partial_x [V(x, t) \rho(x, t)] = 0$ . That's for one coordinate,  $x$ . If there are lots of coordinates this equation reads

$$\partial_t \rho(x, t) + \rho(x, t) \sum_j (\partial_{x_j} V_j) + \sum_j V_j \partial_{x_j} \rho(x, t) = 0 \quad \text{v.9}$$

We call the second term on the left the **divergence** term. It describes the dilation of the volume element by the changes in the  $x$ 's caused by the time development. The last term is the direct result of the time-change in each coordinate  $X(t)$ . Now we have **the general result for the time development of the probability density**. We go look at the Hamiltonian

## Calculation concluded .....

$$\partial_t \rho(x, t) + \rho(x, t) \sum_j (\partial_{x_j} V_j) + \sum_j V_j \partial_{x_j} \rho(x, t) = 0$$

The Hamiltonian case is special. There are two kinds of coordinates  $x_j = q_\alpha$  with  $V_j = \partial_{p_\alpha} H$  and  $x_j = p_\alpha$  with  $V_j = -\partial_{q_\alpha} H$ . In that case, the divergence term  $(\partial_{q_\alpha} \partial_{p_\alpha} H - \partial_{p_\alpha} \partial_{q_\alpha} H) \rho = 0$ . This result, called **Liouville's** theorem, says that the size of the volume element is independent of time. As a result the probability density obeys a special equation, with no divergence term

$$\partial_t \rho(p, q, t) + \sum_\alpha [(\partial_{p_\alpha} H) \partial_{q_\alpha} - (\partial_{q_\alpha} H) \partial_{p_\alpha}] \rho(p, q, t) = 0 \quad \text{v.10}$$

The time derivative of any function of  $p$  and  $q$  is given in Hamiltonian mechanics by

$$dX(p, q)/dt = \sum_\alpha [(\partial_{p_\alpha} H) \partial_{q_\alpha} - (\partial_{q_\alpha} H) \partial_{p_\alpha}] X(p, q) \quad \text{v.11}$$

In particular if  $H$  is the Hamiltonian, assumed to be a function of the  $p$ 's and  $q$ 's but not containing any other  $t$ -dependence, then  $dH/dt=0$ , i.e.  $H$  is independent of time.

$$\partial_t \dot{x} = -V(x, t)$$

$$\frac{\partial}{\partial x} = + \frac{\partial \mathcal{H}}{\partial p_x}$$

$$\frac{\partial}{\partial p_x} = - \frac{\partial \mathcal{H}}{\partial x}$$



## Calculation concluded .....

$$\partial_t \rho(x, t) + \rho(x, t) \sum_j (\partial_{x_j} V_j) + \sum_j V_j \partial_{x_j} \rho(x, t) = 0$$

The Hamiltonian case is special. There are two kinds of coordinates  $x_j = q_\alpha$  with  $V_j = \partial_{p_\alpha} H$  and  $x_j = p_\alpha$  with  $V_j = -\partial_{q_\alpha} H$ . In that case, the divergence term  $(\partial_{q_\alpha} \partial_{p_\alpha} H - \partial_{p_\alpha} \partial_{q_\alpha} H) \rho = 0$ . This result, called **Liouville's** theorem, says that the size of the volume element is independent of time. As a result the probability density obeys a special equation, with no divergence term

$$\partial_t \rho(p, q, t) + \sum_\alpha [(\partial_{p_\alpha} H) \partial_{q_\alpha} - (\partial_{q_\alpha} H) \partial_{p_\alpha}] \rho(p, q, t) = 0 \quad \text{v.10}$$

The time derivative of any function of  $p$  and  $q$  is given in Hamiltonian mechanics by

$$dX(p, q)/dt = \sum_\alpha [(\partial_{p_\alpha} H) \partial_{q_\alpha} - (\partial_{q_\alpha} H) \partial_{p_\alpha}] X(p, q) \quad \text{v.11}$$

In particular if  $H$  is the Hamiltonian, assumed to be a function of the  $p$ 's and  $q$ 's but not containing any other  $t$ -dependence, then  $dH/dt=0$ , i.e.  $H$  is independent of time.



$$\partial_t \tilde{x} = V(x, t)$$

$$\frac{\partial g}{\partial x} = + \frac{\partial \mathcal{H}}{\partial p_x}$$

$$\frac{\partial p}{\partial x} = - \frac{\partial \mathcal{H}}{\partial g_x}$$

$$V_t = - \frac{\partial \mathcal{H}}{\partial g_t}$$

$$V_g = + \frac{\partial \mathcal{H}}{\partial p}$$

$$\partial_p V_p + \partial_q V_q =$$

$$\partial_p \left( -\frac{\partial \mathcal{H}}{\partial q} \right) + \partial_q \left( \frac{\partial \mathcal{H}}{\partial p} \right) =$$

$$= 0$$

$$\partial_x V = V(x,$$

$$\partial_x = + \frac{\partial \mathcal{H}}{\partial p_x}$$

$$\partial_x = - \frac{\partial \mathcal{H}}{\partial q_x}$$

$$V_p = - \frac{\partial \mathcal{H}}{\partial q}$$

$$V_q = + \frac{\partial \mathcal{H}}{\partial p}$$

## Calculation concluded .....

$$\partial_t \rho(x, t) + \rho(x, t) \sum_j (\partial_{x_j} V_j) + \sum_j V_j \partial_{x_j} \rho(x, t) = 0$$

The Hamiltonian case is special. There are two kinds of coordinates  $x_j = q_\alpha$  with  $V_j = \partial_{p_\alpha} H$  and  $x_j = p_\alpha$  with  $V_j = -\partial_{q_\alpha} H$ . In that case, the divergence term  $(\partial_{q_\alpha} \partial_{p_\alpha} H - \partial_{p_\alpha} \partial_{q_\alpha} H) \rho = 0$ . This result, called **Liouville's** theorem, says that the size of the volume element is independent of time. As a result the probability density obeys a special equation, with no divergence term

$$\partial_t \rho(p, q, t) + \sum_\alpha [(\partial_{p_\alpha} H) \partial_{q_\alpha} - (\partial_{q_\alpha} H) \partial_{p_\alpha}] \rho(p, q, t) = 0 \quad \text{v.10}$$

The time derivative of any function of  $p$  and  $q$  is given in Hamiltonian mechanics by

$$dX(p, q)/dt = \sum_\alpha [(\partial_{p_\alpha} H) \partial_{q_\alpha} - (\partial_{q_\alpha} H) \partial_{p_\alpha}] X(p, q) \quad \text{v.11}$$

In particular if  $H$  is the Hamiltonian, assumed to be a function of the  $p$ 's and  $q$ 's but not containing any other  $t$ -dependence, then  $dH/dt=0$ , i.e.  $H$  is independent of time.



$$\partial_r V_p + \partial_q V_q =$$

$$\partial_p \left( -\frac{\partial \mathcal{H}}{\partial q} \right) + \partial_q \left( \frac{\partial \mathcal{H}}{\partial p} \right) =$$

$$= 0$$

$$\frac{\partial \mathcal{H}}{\partial x} = -\frac{p}{\hbar}$$

$$V_r = -\frac{\partial \mathcal{H}}{\partial q}$$

$$V_q = +\frac{\partial \mathcal{H}}{\partial p}$$

$$\partial_x \mathcal{H} = V(x)$$

$$\frac{\partial \mathcal{H}}{\partial p_x} = +\frac{\partial \mathcal{H}}{\partial p_x}$$

$$\frac{\partial \mathcal{H}}{\partial q_x} = -\frac{\partial \mathcal{H}}{\partial q_x}$$

$$\mathcal{H} = \frac{p^2}{2m}$$



$$\partial_r V_p + \partial_q V_q =$$

$$\left(-\frac{\partial \mathcal{H}}{\partial q}\right) + \partial_q \left(\frac{\partial \mathcal{H}}{\partial p}\right) =$$

$$= 0$$

$$\frac{\partial \mathcal{H}}{\partial t} = -\frac{p}{\hbar}$$

$$\rho$$

$$\hbar \partial_t \rho = [\rho, \mathcal{H}]$$

$$\partial_r V_p + \partial_q V_q =$$

$$\left(-\frac{\partial \mathcal{H}}{\partial q}\right) + \partial_q \left(\frac{\partial \mathcal{H}}{\partial p}\right) =$$

$$= 0$$

$$\frac{\partial \mathcal{H}}{\partial t} = -\frac{p}{\hbar}$$

$$\hbar \partial_\hbar \rho = [\rho, \mathcal{H}]$$

$$\rho(0) = |\alpha\rangle\langle\alpha|$$



$$\partial_r V_r + \partial_q V_q =$$

$$\left(-\frac{\partial \mathcal{H}}{\partial q}\right) + \partial_q \left(\frac{\partial \mathcal{H}}{\partial p}\right) =$$

$$= 0$$

$$\frac{\partial \mathcal{H}}{\partial t} = -\frac{p}{\hbar}$$

$$\hbar \partial_t \rho = [\rho, \mathcal{H}]$$

$$\rho|0\rangle = |\alpha\rangle \langle \alpha| \frac{1}{2}$$

$$+ |\beta\rangle \langle \beta| \frac{1}{2}$$

$$0 = \langle \alpha | \beta \rangle$$

$$\rho \quad \text{and} \quad \partial_\mu \rho = [\rho, \psi]$$

$$\rho(0) = |\alpha\rangle\langle\alpha|^{\frac{1}{2}} + |\beta\rangle\langle\beta|^{\frac{1}{2}}$$

$$\text{trace } \rho = 1 \quad 0 = \langle\alpha|\beta\rangle$$

$$\langle x \rangle = \text{trace}(\rho x)$$



$$\frac{\delta \rho}{\delta p} =$$

$$\rho$$

$$i\hbar \partial_t \rho = [\rho, H]$$

$$\rho(0) = |\alpha\rangle\langle\alpha|^{\frac{1}{2}} + |\beta\rangle\langle\beta|^{\frac{1}{2}}$$

$$\text{trace } \rho = 1 \quad 0 = \langle\alpha|\beta\rangle$$

$$\langle x \rangle = \text{trace}(\rho x)$$

$$S = -\text{trace}(\rho \ln \rho)$$

## Calculation concluded .....

$$\partial_t \rho(x, t) + \rho(x, t) \sum_j (\partial_{x_j} V_j) + \sum_j V_j \partial_{x_j} \rho(x, t) = 0$$

The Hamiltonian case is special. There are two kinds of coordinates  $x_j = q_\alpha$  with  $V_j = \partial_{p_\alpha} H$  and  $x_j = p_\alpha$  with  $V_j = -\partial_{q_\alpha} H$ . In that case, the divergence term  $(\partial_{q_\alpha} \partial_{p_\alpha} H - \partial_{p_\alpha} \partial_{q_\alpha} H) \rho = 0$ . This result, called **Liouville's** theorem, says that the size of the volume element is independent of time. As a result the probability density obeys a special equation, with no divergence term

$$\partial_t \rho(p, q, t) + \sum_\alpha [(\partial_{p_\alpha} H) \partial_{q_\alpha} - (\partial_{q_\alpha} H) \partial_{p_\alpha}] \rho(p, q, t) = 0 \quad \text{v.10}$$

The time derivative of any function of  $p$  and  $q$  is given in Hamiltonian mechanics by

$$dX(p, q)/dt = \sum_\alpha [(\partial_{p_\alpha} H) \partial_{q_\alpha} - (\partial_{q_\alpha} H) \partial_{p_\alpha}] X(p, q) \quad \text{v.11}$$

In particular if  $H$  is the Hamiltonian, assumed to be a function of the  $p$ 's and  $q$ 's but not containing any other  $t$ -dependence, then  $dH/dt=0$ , i.e.  $H$  is independent of time.

# Poisson Bracket

The **Poisson Bracket** is Defined by

$$\{f, g\} = \sum_{\alpha} \left[ \frac{\partial f}{\partial q_{\alpha}} \frac{\partial g}{\partial p_{\alpha}} - \frac{\partial g}{\partial q_{\alpha}} \frac{\partial f}{\partial p_{\alpha}} \right]$$

$$\partial_t \rho = \{H, \rho\} \quad \text{for any function of } p\text{'s and } q\text{'s, with no explicit time-dependence}$$
$$dX/dt = \{X, H\}$$

These Poisson brackets are rather like the commutators of quantum mechanics. For example  $\{\{f, g\}, h\} + \{\{h, f\}, g\} + \{\{g, h\}, f\} = 0$ . The same kind of relation is true for operators in quantum theory with  $\{$  and  $\}$  replaced by  $[$  and  $]$ . The bracket relations for classical time-dependence are very much like the time-dependence of operators and density matrices in quantum theory. This relation between quantum mechanics and the canonical version of classical mechanics is quite surprising and turns out to be quite deep.

Note the difference in sign between the relation for  $dX/dt$  and the one for  $\partial_t \rho$ . I have gotten one of those signs wrong many times in my life. Think for a bit about why we write  $\partial_t \rho$  but  $dX/dt$ . Why?



---

## Liouville's Theorem and conservation of phase space volume

This theorem is the statement that volume elements in phase space are conserved in the course of a motion. That's the consequence of the vanishing divergence term. So if you start off with a  $\rho(p,q,0)$  at time zero which is non-zero in some region of phase space, of volume  $\Omega(0)$ , after a time,  $\rho(p,q,t)$  will progress to occupy a different region of phase space of volume  $\Omega(t)$ . According to Liouville's theorem the volumes of the regions before and after will be unchanged,  $\Omega(t) = \Omega(0)$ . Imagine that the probability density was constant within the region you picked, i.e.  $\rho(p,q,0) = 1/\Omega(0)$ . Then, by the same argument to any tiny subvolume of the original region would also have its volume preserved under the transformation, and therefore the value of  $\rho(p,q,t)$  within the new region would be the same as the one within the old. Within the new region,  $\rho(p,q,t) = 1/\Omega(0)$ .



## Any function of $H$ will do

To achieve equilibrium we can have the probability density be **any** function of the Hamiltonian and the other conserved quantities in the system. The following functions are in broad use. We assume one type of particle, with number  $N$ .

**Canonical ensemble:**  $\rho = \exp(-\beta H)/Z(\beta)$      $N$  is fixed

This is the right ensemble to use if a small system with a known number of particles is weakly coupled to a larger system so that it might exchange energy but not particles with the larger system. This is OK to use if there are no other important conserved variables, beyond the ones mentioned. If you have a system in motion, watch out for the momentum. If you have a galaxy, watch out for angular momentum.

**Microcanonical ensemble:**  $\rho = \delta(E-H)/\chi(E)$      $N$  is fixed

This is the right ensemble to use if the energy and number of particles in a small system are known. This is OK to use if there are no other important conserved variables, beyond the ones mentioned. However, small systems with fixed energy can often have other hidden conserved things.

$$\Omega(\vec{p}, \vec{q})$$

$$A = 0$$

$$\frac{\partial \Gamma}{\partial t} = -\frac{p}{t}$$

$$p(p, q, t)$$

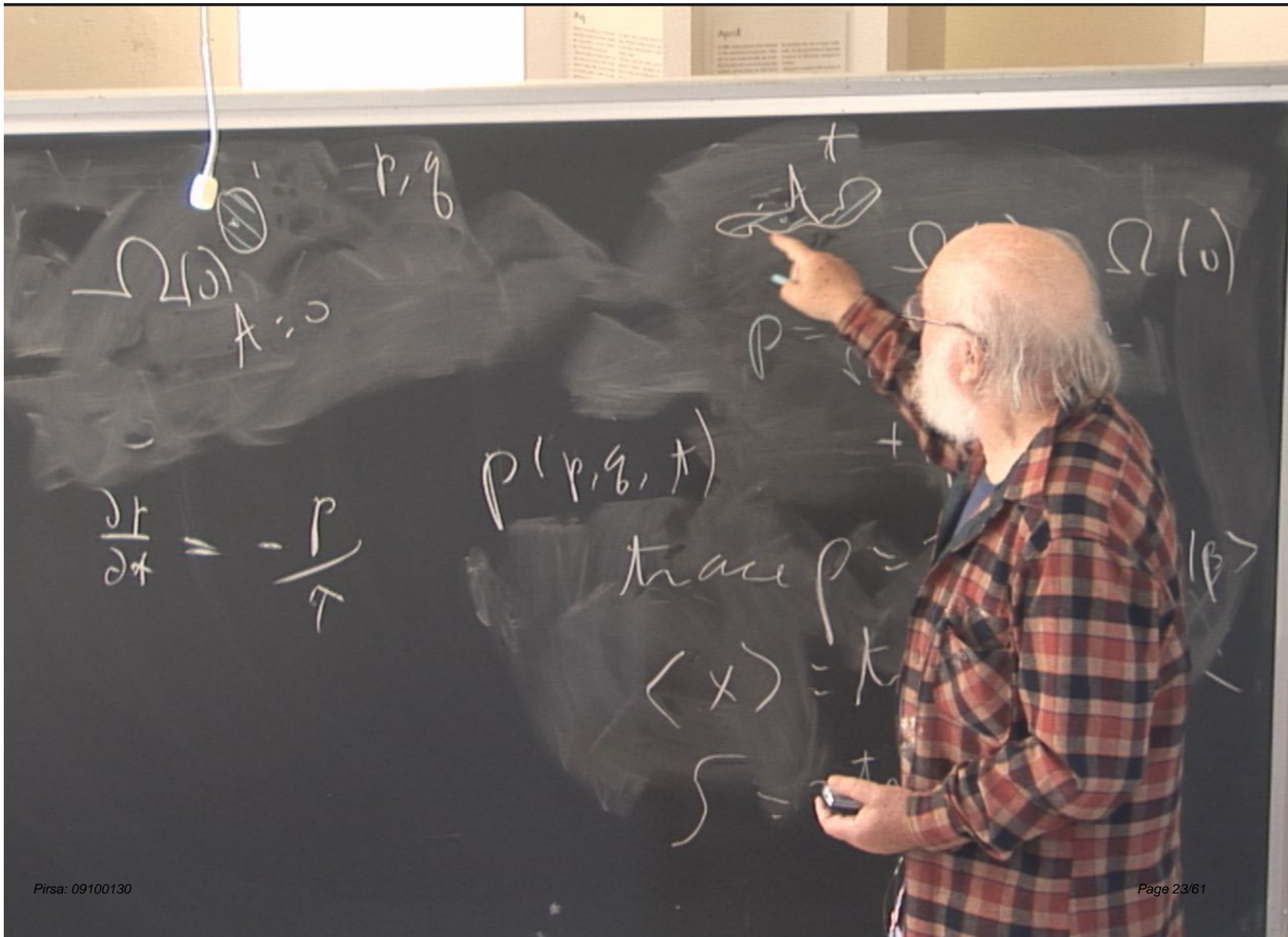
$$+ |\beta\rangle \langle \beta| \frac{1}{2}$$

$$\text{trace } \rho = 1 \quad 0 = \langle x \rangle$$

$$\langle x \rangle = \text{trace}(\rho x)$$

$$S = -\text{trace}(\rho \ln \rho)$$





$$\Omega(t) \quad p, q$$

$$\frac{\partial t}{\partial t} = -\frac{p}{T}$$

$$\rho(p, q, t)$$

$$\text{trace } \rho = 1 \quad 0 = \langle x | \rho \rangle$$

$$\langle x \rangle = \text{trace}(\rho x)$$

$$S = -\text{trace}(\rho \ln \rho)$$

$$\Omega(t) = \Omega(0)$$

$$P = \frac{1}{\Omega(t)} = \frac{1}{\Omega(0)}$$

$$+ | \beta \rangle \langle \beta | \frac{1}{2}$$



## Liouville's Theorem and conservation of functions of the energy

$$\partial_t \rho(p, q, t) + \sum_{\alpha} [(\partial_{p_{\alpha}} H) \partial_{q_{\alpha}} - (\partial_{q_{\alpha}} H) \partial_{p_{\alpha}}] \rho(p, q, t) = 0$$

If  $\rho(p, q, 0)$  is any function of the Hamiltonian, e.g.  $\rho(p, q, 0) = Z^{-1} \exp[-\beta H(p, q)]$  then this same functional form will hold for all times, assuming that the Hamiltonian has no explicit time dependence.  $\rho(p, q, 0) = f(H(p, q))$  implies that  $\rho(p, q, t) = f(H(p, q))$  for any  $t$ .

Further, if  $\rho$  is any function of a time-independent  $H$  and of any other conserved function of  $p$  and  $q$  (but not  $t$ ) then  $\rho$  will be a solution of our equation. Thus, not only is the Boltzmann function a solution describing the equilibrium time-dependence of a Hamiltonian system, there are many other solutions as well.

Classical mechanics is not enough to specify a unique equilibrium probability density in a classical system. Something else is needed in addition.

## Liouville's Theorem and conservation of phase space volume

This theorem is the statement that volume elements in phase space are conserved in the course of a motion. That's the consequence of the vanishing divergence term. So if you start off with a  $\rho(p,q,0)$  at time zero which is non-zero in some region of phase space, of volume  $\Omega(0)$ , after a time,  $\rho(p,q,t)$  will progress to occupy a different region of phase space of volume  $\Omega(t)$ . According to Liouville's theorem the volumes of the regions before and after will be unchanged,  $\Omega(t) = \Omega(0)$ . Imagine that the probability density was constant within the region you picked, i.e.  $\rho(p,q,0) = 1/\Omega(0)$ . Then, by the same argument to any tiny subvolume of the original region would also have its volume preserved under the transformation, and therefore the value of  $\rho(p,q,t)$  within the new region would be the same as the one within the old. Within the new region,  $\rho(p,q,t) = 1/\Omega(0)$ .

## Liouville's Theorem and conservation of functions of the energy

$$\partial_t \rho(p, q, t) + \sum_{\alpha} [(\partial_{p_{\alpha}} H) \partial_{q_{\alpha}} - (\partial_{q_{\alpha}} H) \partial_{p_{\alpha}}] \rho(p, q, t) = 0$$

If  $\rho(p, q, 0)$  is any function of the Hamiltonian, e.g.  $\rho(p, q, 0) = Z^{-1} \exp[-\beta H(p, q)]$  then this same functional form will hold for all times, assuming that the Hamiltonian has no explicit time dependence.  $\rho(p, q, 0) = f(H(p, q))$  implies that  $\rho(p, q, t) = f(H(p, q))$  for any  $t$ .

Further, if  $\rho$  is any function of a time-independent  $H$  and of any other conserved function of  $p$  and  $q$  (but not  $t$ ) then  $\rho$  will be a solution of our equation. Thus, not only is the Boltzmann function a solution describing the equilibrium time-dependence of a Hamiltonian system, there are many other solutions as well.

Classical mechanics is not enough to specify a unique equilibrium probability density in a classical system. Something else is needed in addition.



$\Omega(x)$   
 $A \propto$   
 $p, q$   
 $\rho(p, q) = [H^2(p, q)]$   
 $\rho = \frac{1}{\Omega(x)} = \frac{1}{\Omega(0)}$

$\frac{dt}{dx} = -\frac{p}{q}$   
 $\rho(p, q, x)$   
 $+ |\beta\rangle \langle \beta| \frac{1}{2}$   
 $\text{trace } \rho = 1$   
 $0 = \langle \alpha | \beta \rangle$   
 $\langle x \rangle = \text{trace}(\rho x)$   
 $S = -\text{trace}(\rho \ln \rho)$

## Liouville's Theorem and conservation of functions of the energy

$$\partial_t \rho(p, q, t) + \sum_{\alpha} [(\partial_{p_{\alpha}} H) \partial_{q_{\alpha}} - (\partial_{q_{\alpha}} H) \partial_{p_{\alpha}}] \rho(p, q, t) = 0$$

If  $\rho(p, q, 0)$  is any function of the Hamiltonian, e.g.  $\rho(p, q, 0) = Z^{-1} \exp[-\beta H(p, q)]$  then this same functional form will hold for all times, assuming that the Hamiltonian has no explicit time dependence.  $\rho(p, q, 0) = f(H(p, q))$  implies that  $\rho(p, q, t) = f(H(p, q))$  for any  $t$ .

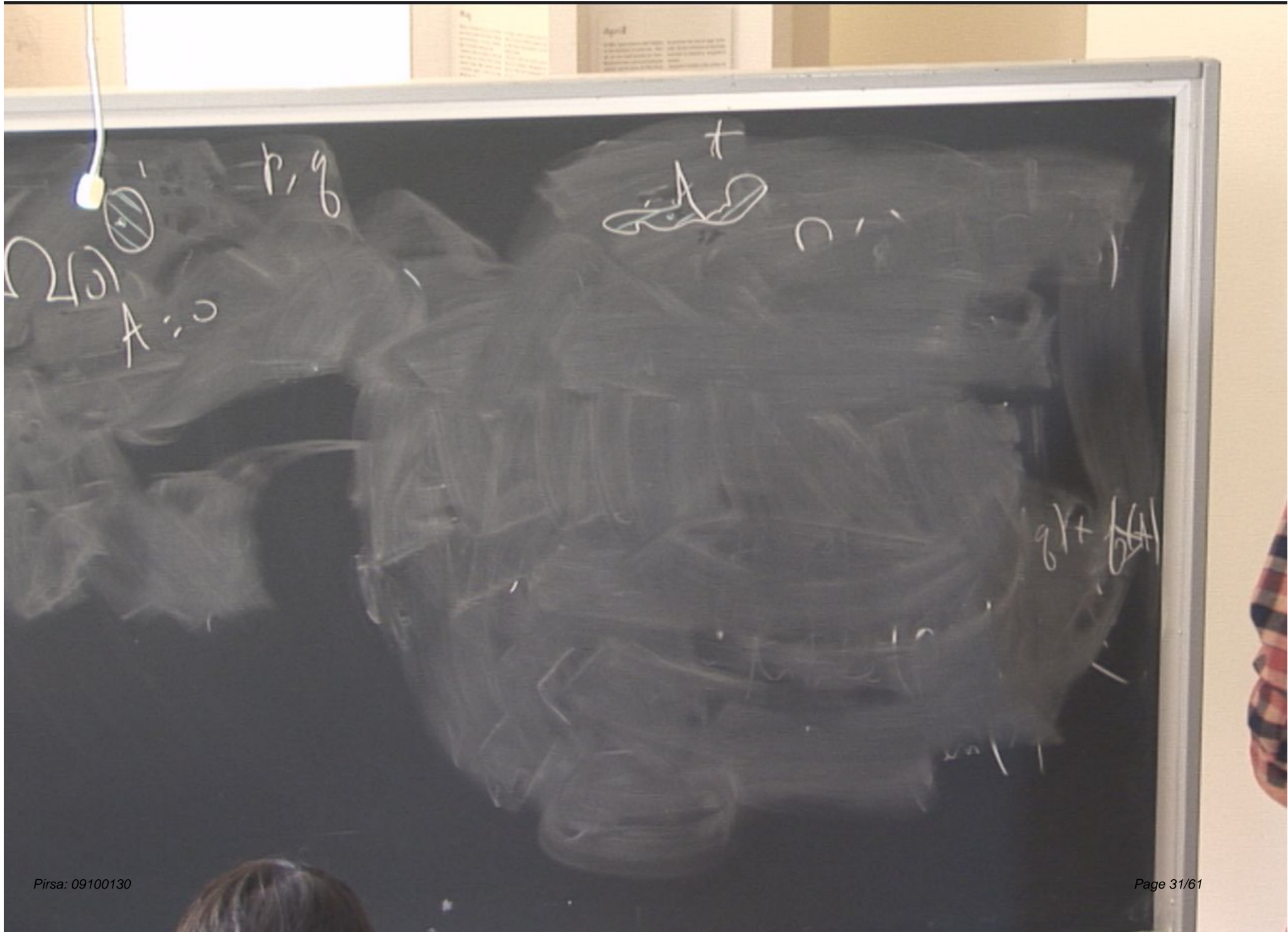
Further, if  $\rho$  is any function of a time-independent  $H$  and of any other conserved function of  $p$  and  $q$  (but not  $t$ ) then  $\rho$  will be a solution of our equation. Thus, not only is the Boltzmann function a solution describing the equilibrium time-dependence of a Hamiltonian system, there are many other solutions as well.

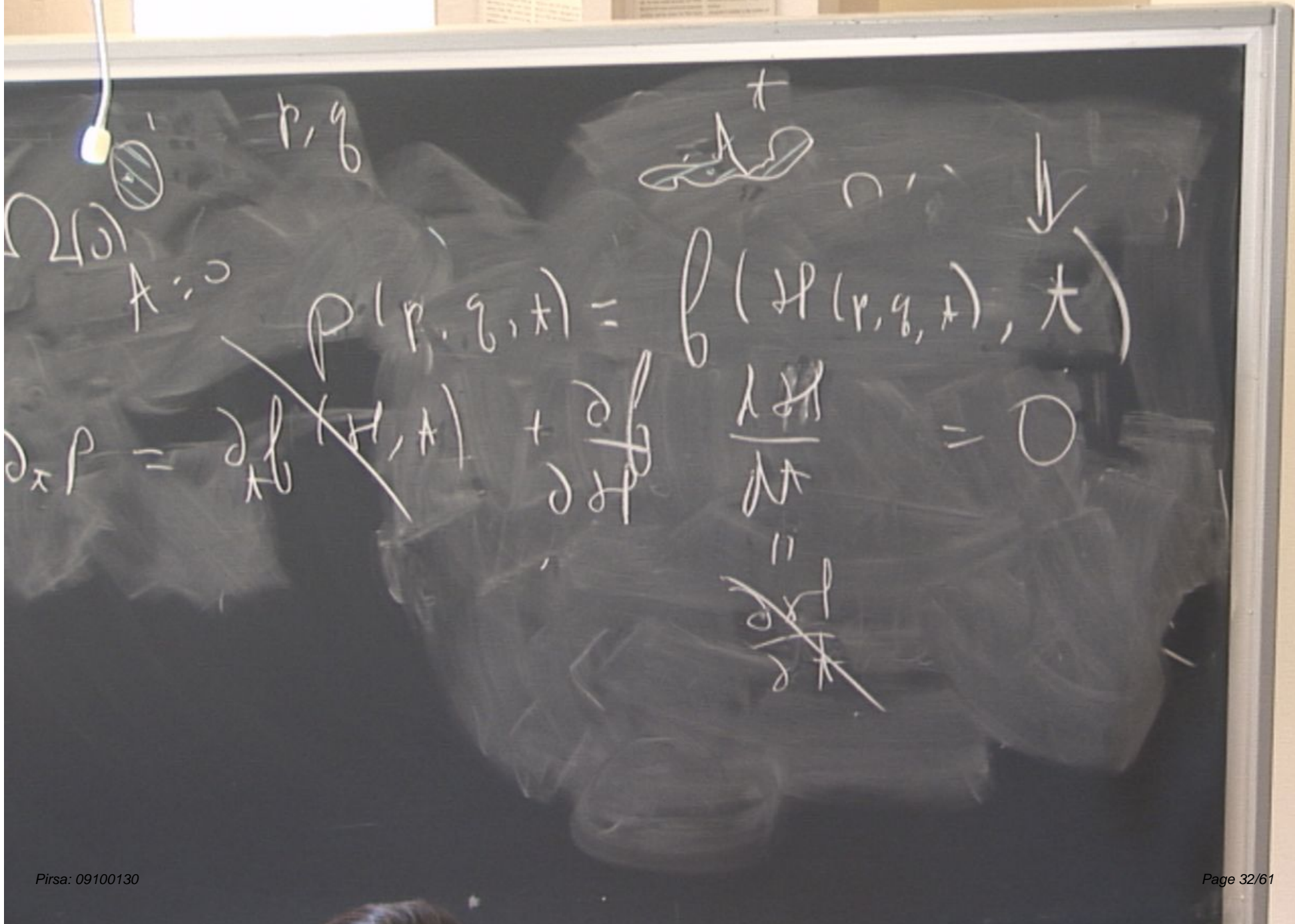
Classical mechanics is not enough to specify a unique equilibrium probability density in a classical system. Something else is needed in addition.



$\Omega(t)$   
 $A \propto$   
 $p, q$   
 $\rho(p, q) = [H(p, q)]^2$   
 $P = \frac{1}{\Omega(t)} = \frac{1}{\Omega(0)}$   
 $\frac{dH}{dt} = \frac{\partial H}{\partial t}$   
 $H(p, q, t)$   
 $H(p, q) = \frac{p^2}{2m} + V(q)$   
 $\langle x \rangle = \text{trace}(\rho x)$   
 $\int = -\text{trace}(\rho \ln \rho)$







## Liouville's Theorem and conservation of functions of the energy

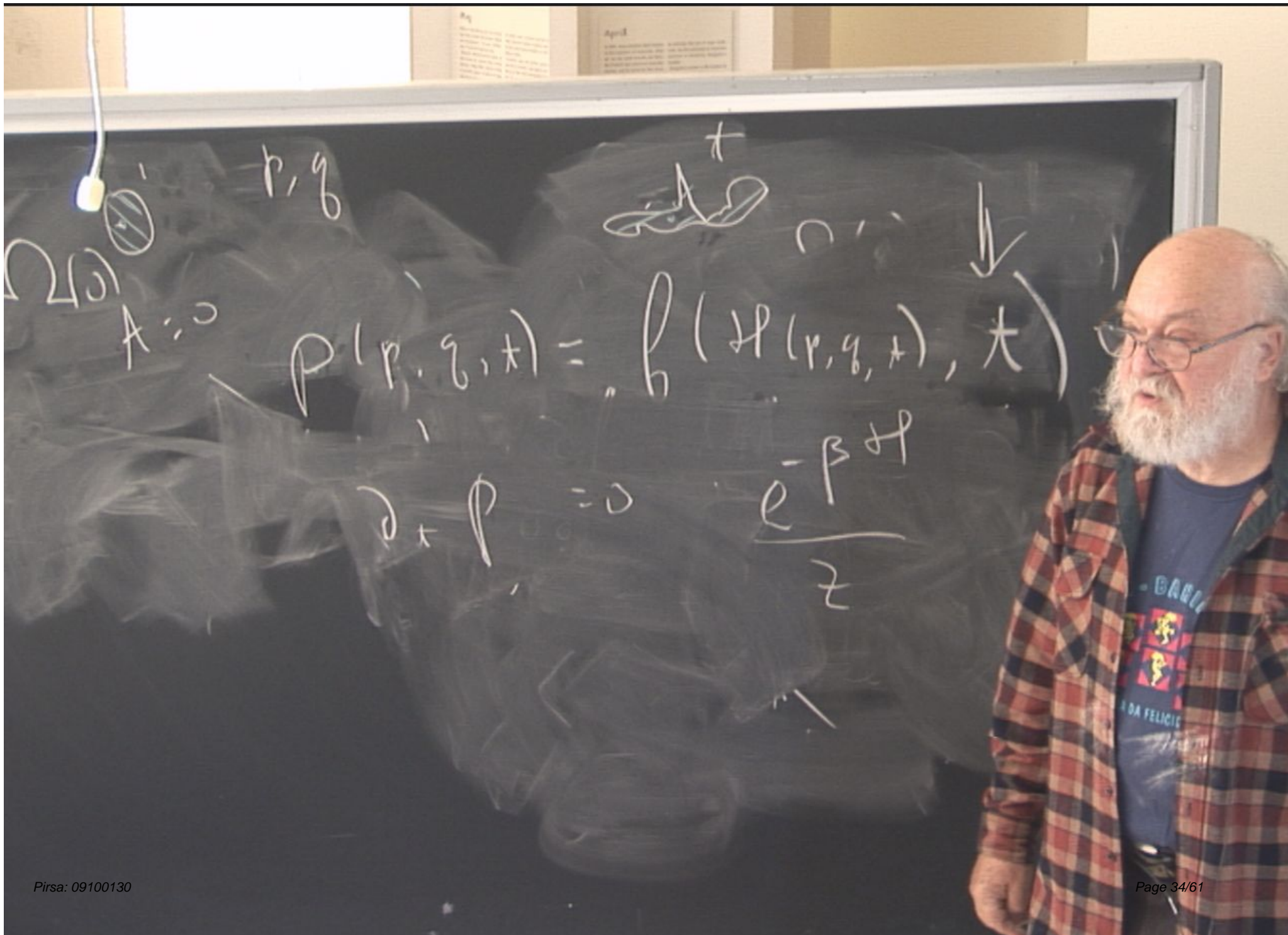
$$\partial_t \rho(p, q, t) + \sum_{\alpha} [(\partial_{p_{\alpha}} H) \partial_{q_{\alpha}} - (\partial_{q_{\alpha}} H) \partial_{p_{\alpha}}] \rho(p, q, t) = 0$$

If  $\rho(p, q, 0)$  is any function of the Hamiltonian, e.g.  $\rho(p, q, 0) = Z^{-1} \exp[-\beta H(p, q)]$  then this same functional form will hold for all times, assuming that the Hamiltonian has no explicit time dependence.  $\rho(p, q, 0) = f(H(p, q))$  implies that  $\rho(p, q, t) = f(H(p, q))$  for any  $t$ .

Further, if  $\rho$  is any function of a time-independent  $H$  and of any other conserved function of  $p$  and  $q$  (but not  $t$ ) then  $\rho$  will be a solution of our equation. Thus, not only is the Boltzmann function a solution describing the equilibrium time-dependence of a Hamiltonian system, there are many other solutions as well.

Classical mechanics is not enough to specify a unique equilibrium probability density in a classical system. Something else is needed in addition.





## Any function of $H$ will do

To achieve equilibrium we can have the probability density be **any** function of the Hamiltonian and the other conserved quantities in the system. The following functions are in broad use. We assume one type of particle, with number  $N$ .

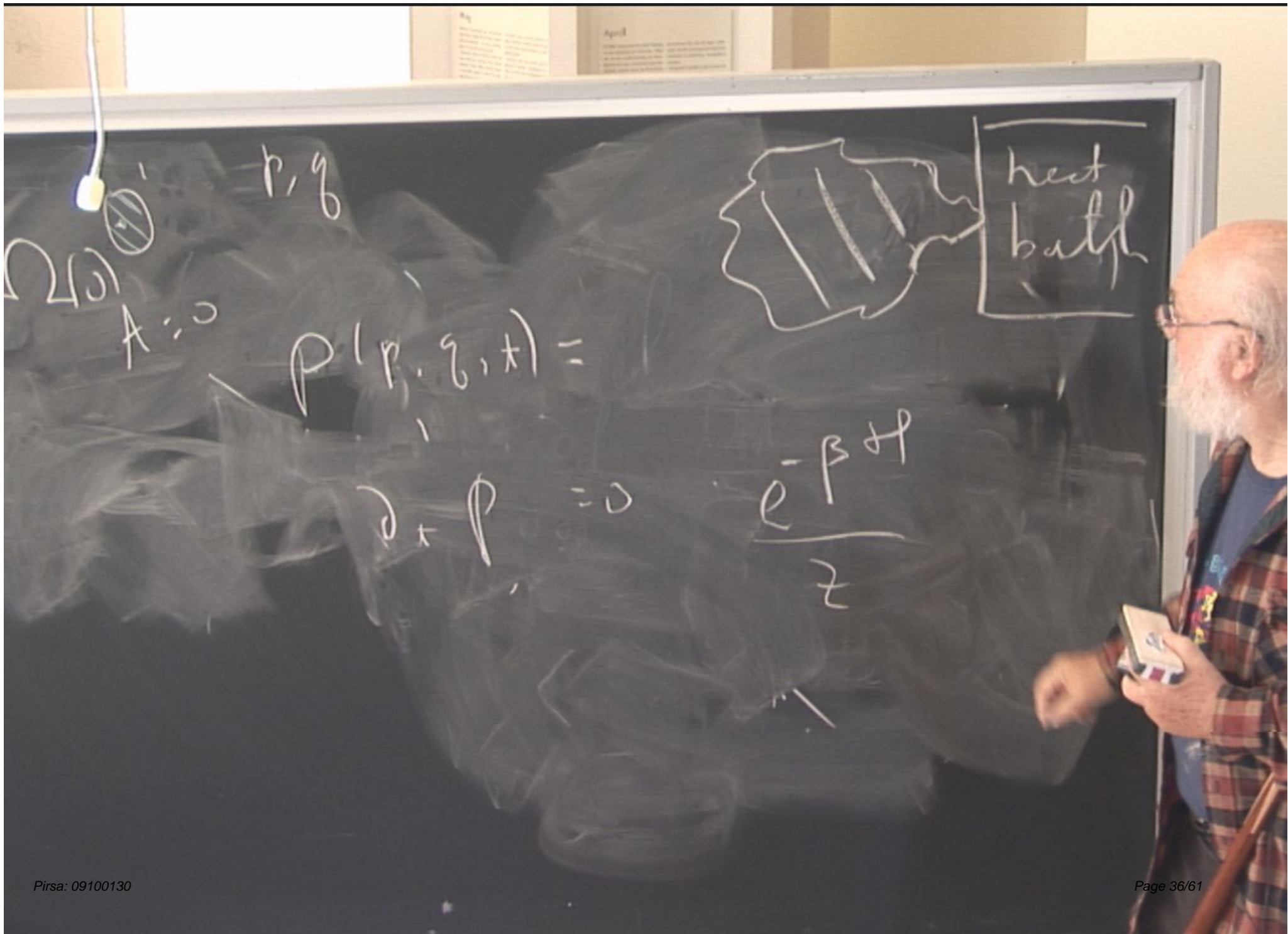
**Canonical ensemble:**  $\rho = \exp(-\beta H)/Z(\beta)$      $N$  is fixed

This is the right ensemble to use if a small system with a known number of particles is weakly coupled to a larger system so that it might exchange energy but not particles with the larger system. This is OK to use if there are no other important conserved variables, beyond the ones mentioned. If you have a system in motion, watch out for the momentum. If you have a galaxy, watch out for angular momentum.

**Microcanonical ensemble:**  $\rho = \delta(E-H)/\chi(E)$      $N$  is fixed

This is the right ensemble to use if the energy and number of particles in a small system are known. This is OK to use if there are no other important conserved variables, beyond the ones mentioned. However, small systems with fixed energy can often have other hidden conserved things.





$p, q$   
 $A = 0$

$$P(p, q, t) =$$

$$\partial_t P = 0$$

$$\frac{e^{-\beta H}}{Z}$$





## Any function of $H$ will do, continued

Grand Canonical ensemble:  $\rho = \exp[-\beta(H-\mu N)]/\Xi(\beta,\mu)$

This is the right ensemble to use if a small system is weakly coupled to a larger system so that it might exchange energy and particles with the larger system. This is OK to use if there are no other important conserved variables, beyond the ones mentioned. However watch out for many different kinds of particles, each type has a conserved  $N$ . Classical mechanics is provided with extra factors of  $1/N!$  in  $\rho$  for each different kind of particle. Particle statistics automatically does this in quantum theory.

For large systems, and for most purposes, all ensembles are equivalent to one another.  
Why?

## Any function of $H$ will do

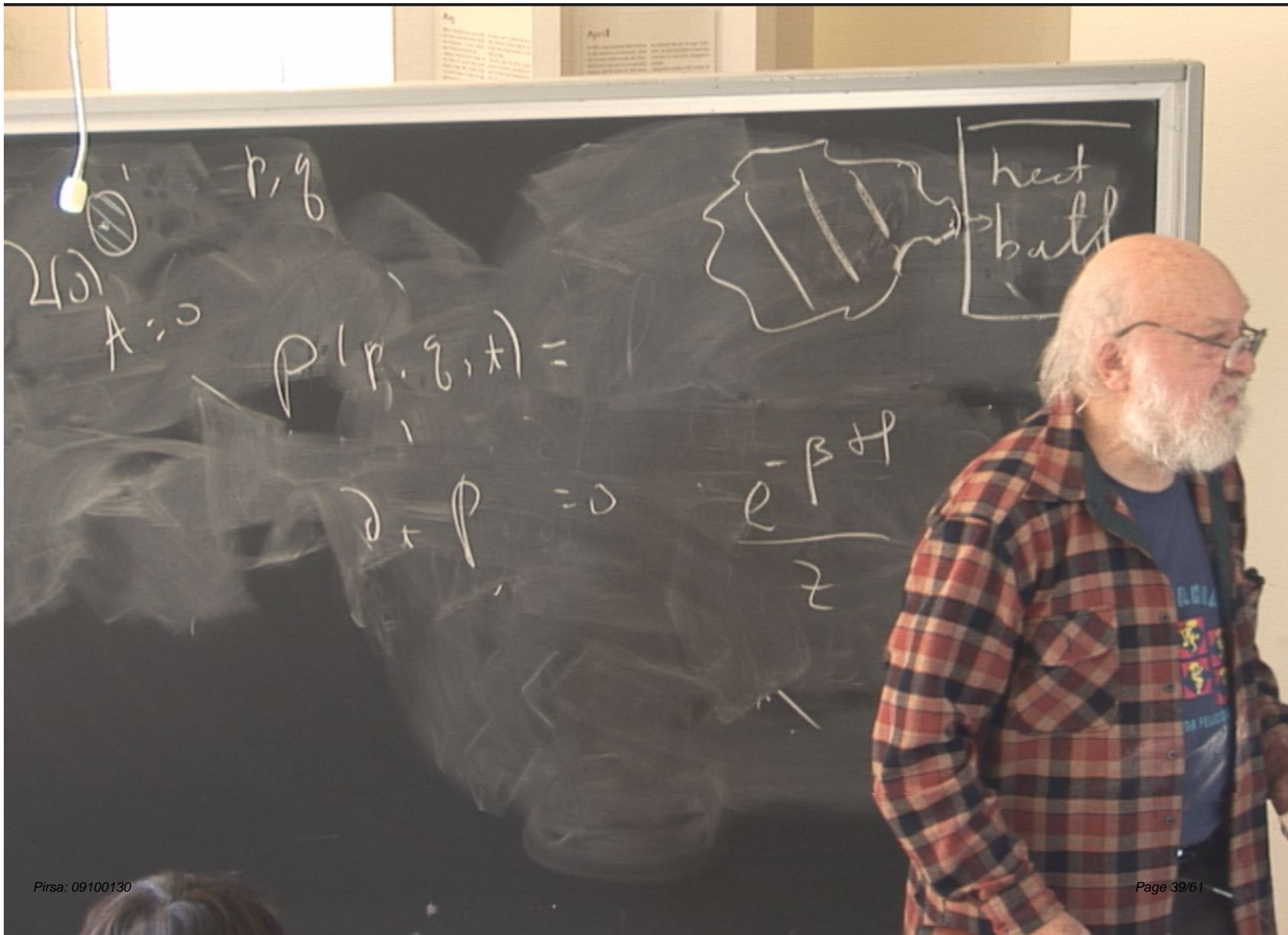
To achieve equilibrium we can have the probability density be **any** function of the Hamiltonian and the other conserved quantities in the system. The following functions are in broad use. We assume one type of particle, with number  $N$ .

**Canonical ensemble:**  $\rho = \exp(-\beta H)/Z(\beta)$      $N$  is fixed

This is the right ensemble to use if a small system with a known number of particles is weakly coupled to a larger system so that it might exchange energy but not particles with the larger system. This is OK to use if there are no other important conserved variables, beyond the ones mentioned. If you have a system in motion, watch out for the momentum. If you have a galaxy, watch out for angular momentum.

**Microcanonical ensemble:**  $\rho = \delta(E-H)/\chi(E)$      $N$  is fixed

This is the right ensemble to use if the energy and number of particles in a small system are known. This is OK to use if there are no other important conserved variables, beyond the ones mentioned. However, small systems with fixed energy can often have other hidden conserved things.





## Any function of $H$ will do

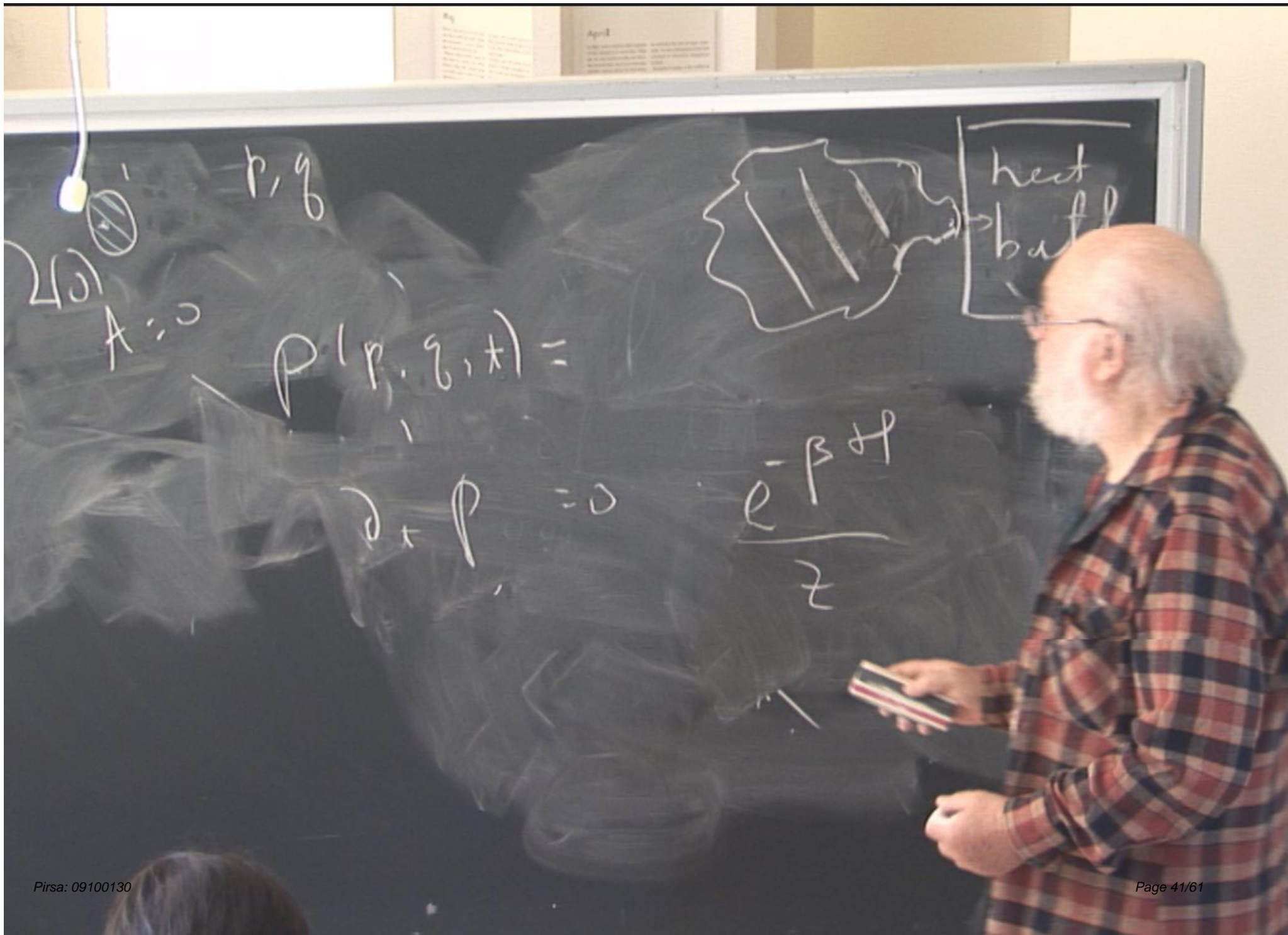
To achieve equilibrium we can have the probability density be **any** function of the Hamiltonian and the other conserved quantities in the system. The following functions are in broad use. We assume one type of particle, with number  $N$ .

**Canonical ensemble:**  $\rho = \exp(-\beta H)/Z(\beta)$      $N$  is fixed

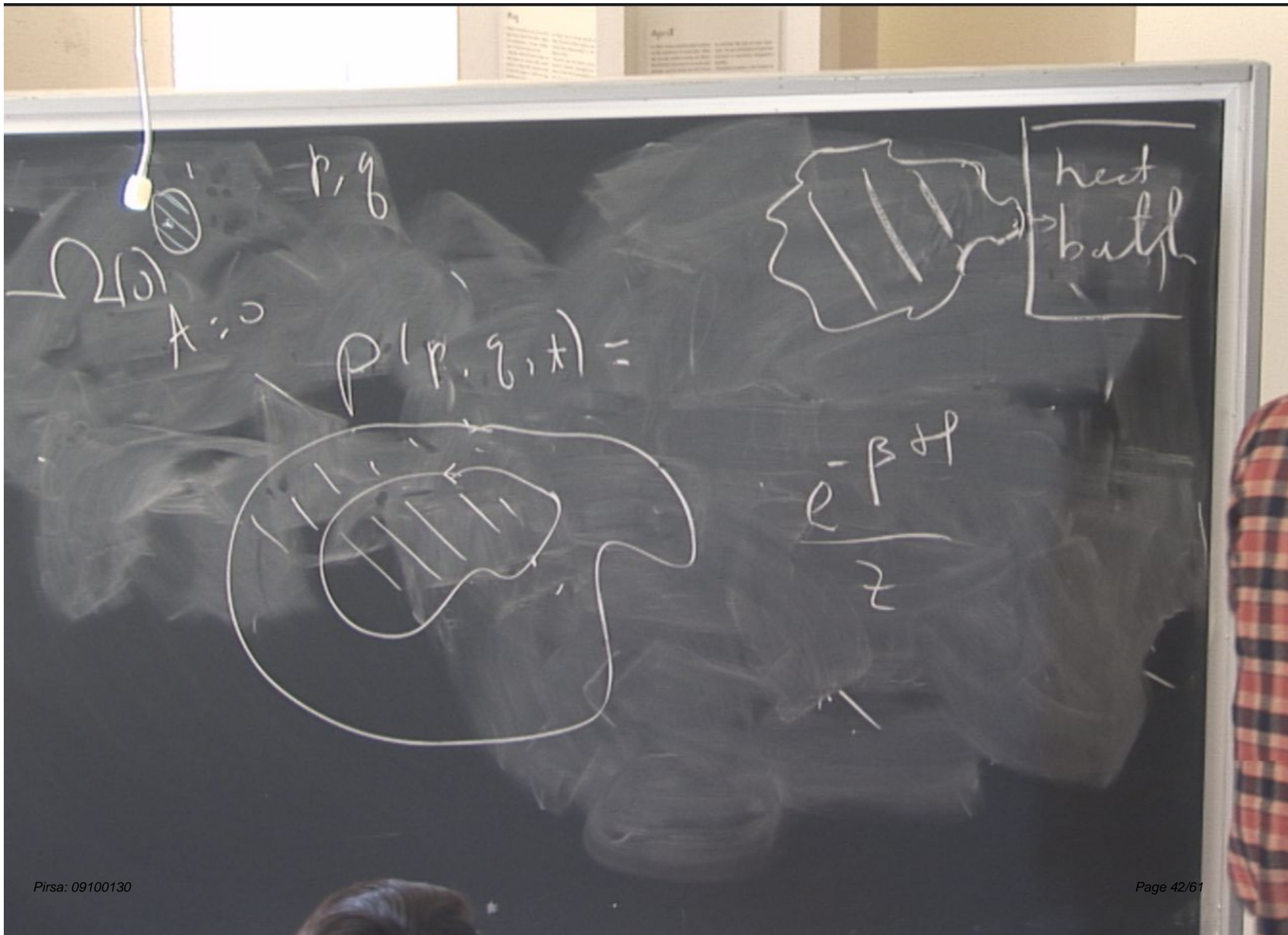
This is the right ensemble to use if a small system with a known number of particles is weakly coupled to a larger system so that it might exchange energy but not particles with the larger system. This is OK to use if there are no other important conserved variables, beyond the ones mentioned. If you have a system in motion, watch out for the momentum. If you have a galaxy, watch out for angular momentum.

**Microcanonical ensemble:**  $\rho = \delta(E-H)/\chi(E)$      $N$  is fixed

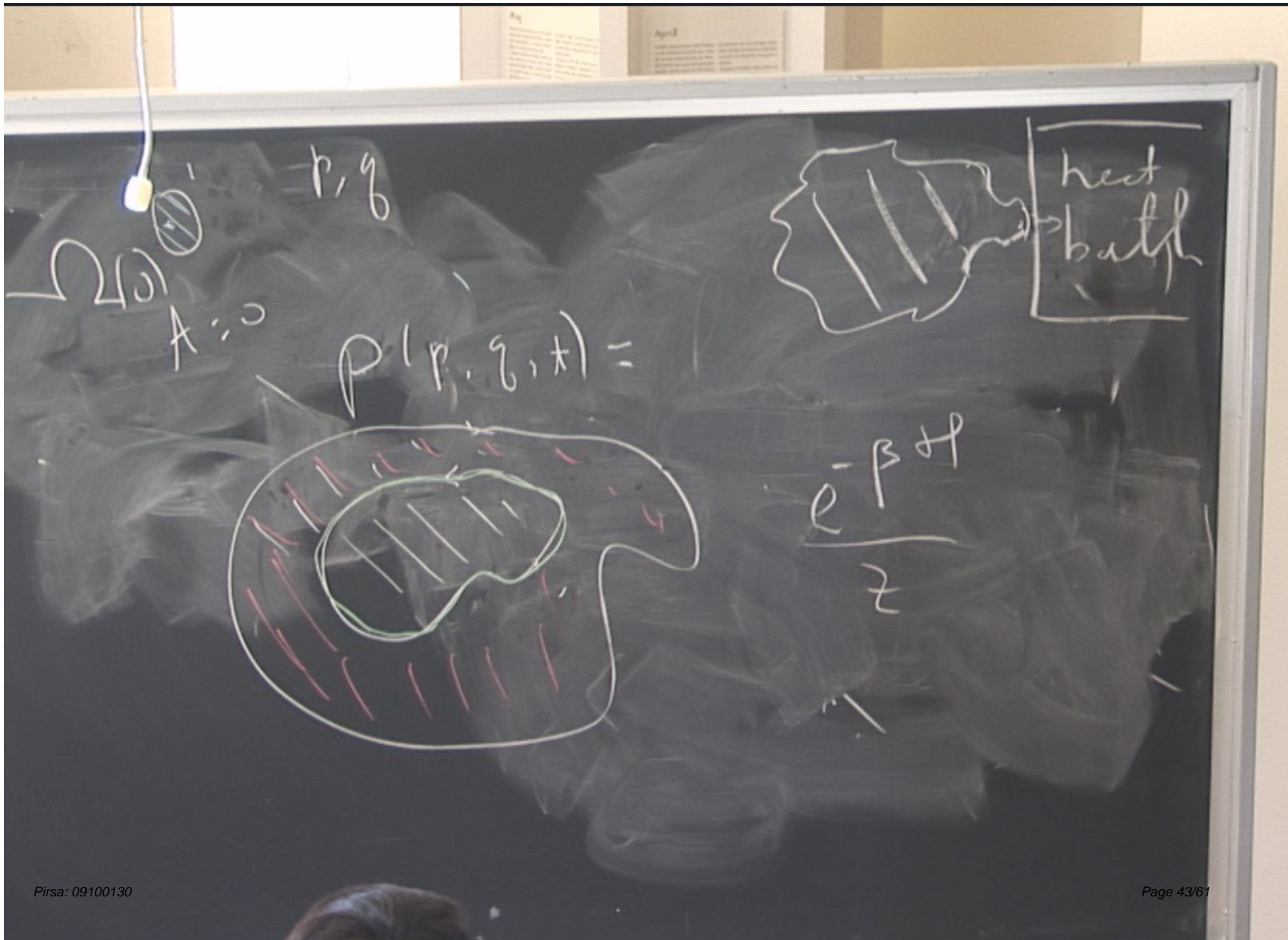
This is the right ensemble to use if the energy and number of particles in a small system are known. This is OK to use if there are no other important conserved variables, beyond the ones mentioned. However, small systems with fixed energy can often have other hidden conserved things.











## Any function of $H$ will do, continued

Grand Canonical ensemble:  $\rho = \exp[-\beta(H-\mu N)]/\Xi(\beta,\mu)$

This is the right ensemble to use if a small system is weakly coupled to a larger system so that it might exchange energy and particles with the larger system. This is OK to use if there are no other important conserved variables, beyond the ones mentioned. However watch out for many different kinds of particles, each type has a conserved  $N$ . Classical mechanics is provided with extra factors of  $1/N!$  in  $\rho$  for each different kind of particle. Particle statistics automatically does this in quantum theory.

For large systems, and for most purposes, all ensembles are equivalent to one another.  
Why?

## Say it again

For the simplest case in which  $H = p^2/(2M) + U(\mathbf{r})$ , the result of Hamiltonian mechanics is that the probability distribution  $\rho(\mathbf{p}, \mathbf{r}, t)$  has the time dependence

$$\partial_t \rho(\mathbf{p}, \mathbf{r}, t) + (\mathbf{p}/M) \cdot \nabla_{\mathbf{r}} \rho(\mathbf{p}, \mathbf{r}, t) - (\nabla_{\mathbf{r}} U) \cdot \nabla_{\mathbf{p}} \rho(\mathbf{p}, \mathbf{r}, t) = 0$$

A time independent solution of this equation would be that  $\rho$  could be any function of  $H$ . This result stands in apparent contradiction to our knowledge of statistical mechanics which tells us that the probability distribution should be the Maxwell-Boltzmann distribution, i.e. one which is exponential in  $H$ . **What additional information should we bring to bear on this situation?**

We already have a hint from the Brownian motion calculation that this calculation might give the Maxwell-Boltzmann result. Let's go back to that and see what equation we get for  $\rho$ . The Einstein model for Brownian motion is

$$dp/dt = \dots + \eta(t) - p/\tau$$

where ... might stand for additional terms coming from Hamiltonian mechanics. I plan to study successively the effect of the two terms in this model upon the equation for  $\rho$  and then put it all together.



## The friction term $-p/\tau$

We already know the effect of this term. It is a generalized velocity of the form included in equation v.7, with  $V(p,t) = -p/\tau$ . We therefore know immediately what effect this term has on the equation for the probability distribution. It gives, via equation v.9

$$\partial_t \rho(p,x,t) = -\partial_p [V \rho] = \dots + \partial_p [p/\tau \rho(p,x,t)]$$

where  $x$  is the variable conjugate to  $p$  in a Hamiltonian formulation. We hold on to this result to use in our later analysis. (Note that derivative  $\partial_p$  acts on everything.)

However, we cannot just look up our old result to get the effect of the other term, the stochastic kicks,  $\eta$ , in the Brownian model

$$dp/dt = \dots + \eta(t) - p/\tau$$

As we shall see in a bit, their average first order effect is zero but their effect to second order produces a result proportional to the time that they have been in action. Our old result does not include second order effects. So we shall go back almost to the beginning to assess the effect of these kicks upon the time-dependence.

## calculation of the effect of $dp/dt = \dots + \eta(t)$

Recall our old calculation of  $\partial_t \rho$ . In this situation, we are after an understanding of the behavior of the momentum  $p$ . We have two ways of calculating the average of a function of momentum at time  $t + \delta t$ . The first comes from computing  $\int dp g(p) \rho(p, t + \delta t)$ . That same average is obtained by taking the solution at time  $t + \delta t$ , which is of the form  $p(t + \delta t) = p(t) + \delta p$ . Here,  $\delta p$  is given by the effect of the stochastic term, so that

$$\delta p = \int_t^{t+\delta t} ds \eta(s)$$

Then the average at time  $t + \delta t$  can also be written as  $\langle \int dp g(p + \delta p) \rho(p, t) \rangle$ . Here, the average  $\langle \dots \rangle$  is an average over the possible values of the stochastic variables  $\eta(s)$  for  $s$  between  $t$  and  $t + \delta t$ . If we equate these two expressions, we find

$$\int dp g(p) \rho(p, t + \delta t) = \langle \int dp g(p + \delta p) \rho(p, t) \rangle$$

The right hand side can be rearranged by shifting the origin of integration and replacing the variable  $p$  by  $p - \delta p$ . Then this right side becomes  $\langle \int dp g(p) \rho(p - \delta p, t) \rangle$ , and

$$\int dp g(p) \rho(p, t + \delta t) = \langle \int dp g(p) \rho(p - \delta p, t) \rangle$$

Since  $g(p)$  is arbitrary, we it follows that

Grains of  
Pollen to  
Evidence  
for Atoms

How  
Big Is A  
Molecule?

$$\int \psi^* \psi \delta A + \delta A \int \psi^* \psi \delta A$$

$$= \int \psi^* \psi \delta A + \langle \delta A \rangle \int \psi^* \psi \delta A$$





Grains of  
Pollen to  
Evidence  
for Atoms

How  
Big Is A  
Molecule?

$$\int \psi^* \hat{H} \psi + \delta \int \psi^* \hat{H} \psi = 0$$

$$= \int \psi^* \hat{H} \psi + \langle \delta \psi \rangle \int \psi^* \hat{H} \psi$$

heat bath

calculation of the effect of  $dp/dt = \dots + \eta(t)$  continued

$$\rho(p, t + \delta t) = \langle \rho(p - \delta p, t) \rangle \quad \delta p = \int_t^{t+\delta t} ds \, \eta(s)$$

expand (the result is particularly simple because  $\delta p$  does not depend upon  $p$  (or  $q$ )).

$$\rho(p, t) + \delta t \, \partial_t \rho(p, t) = \langle \rho(p, t) \rangle + \langle \delta p \rangle \partial_p \rho(p, t) + \langle \delta p^2 \rangle (\partial_p)^2 \rho(p, t) / 2$$

The first terms on each side cancel, the average of  $\delta p$  is zero, and the rest gives

$$\partial_t \rho(p, t) = [ \langle \delta p^2 \rangle / (2\delta t) ] (\partial_p)^2 \rho(p, t)$$

The average has the value:

$$[\delta p]^2 = \int_t^{t+\delta t} du \int_t^{t+\delta t} ds \, \langle \eta(u) \eta(s) \rangle = \int_t^{t+\delta t} du \int_t^{t+\delta t} ds \, \Gamma \delta(u - s) = \Gamma \delta t$$

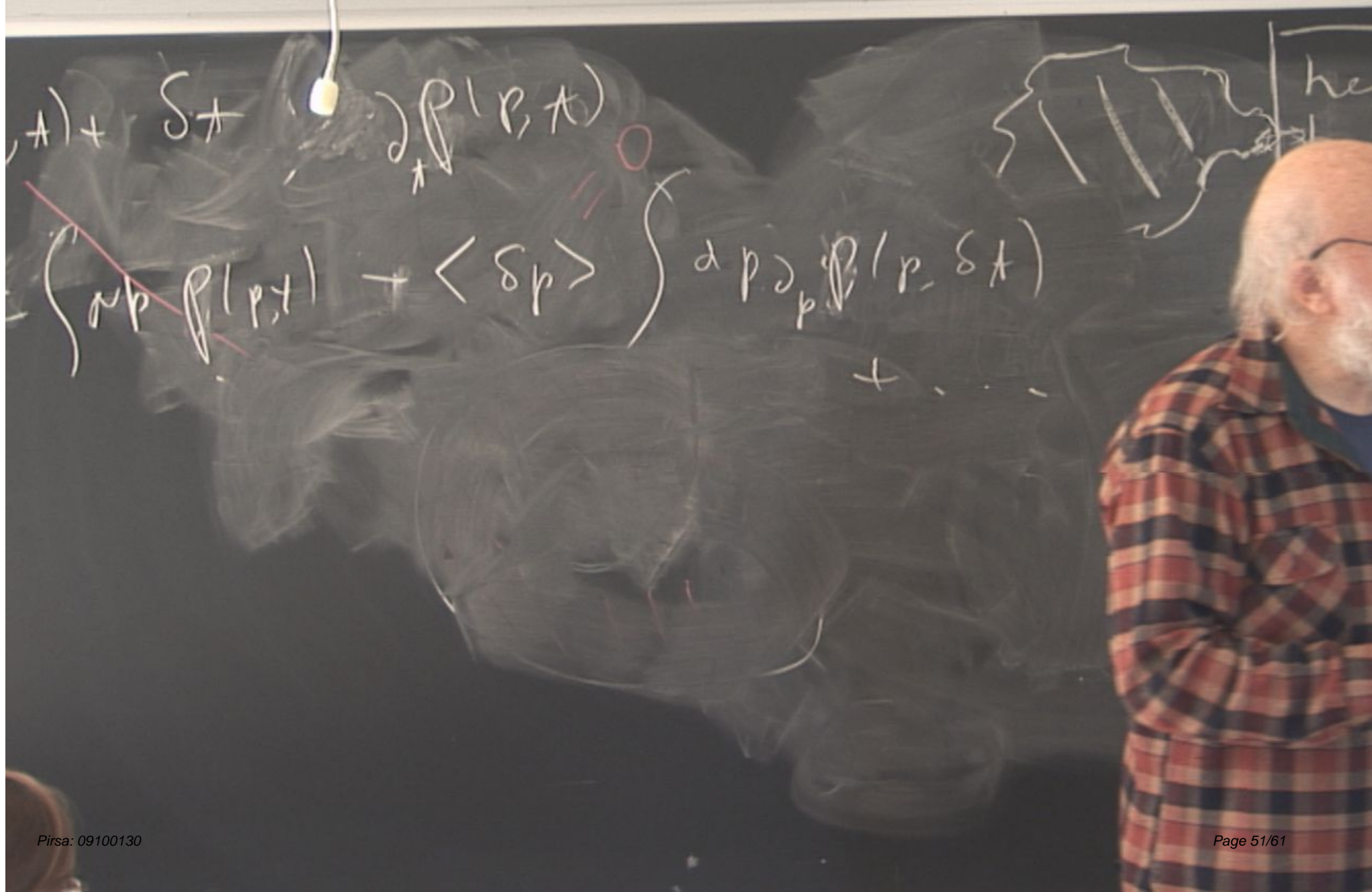
(recall that  $\langle \eta(t) \eta(s) \rangle = \Gamma \delta(t - s)$ ) so that we end up with the result

$$\partial_t \rho(p, t) = \dots + (\Gamma/2) (\partial_p)^2 \rho(p, t)$$

This describes a diffusion in momentum space.

(Notice that because it is diffusive, this result cannot be followed backward in time!

Notice also that we had to go to second order in our calculation of the effect of the random kicks because the second order effect here is linear in  $\delta t$  and we were holding on to all linear terms.)



$$\frac{\partial}{\partial t} \rho(p, t) + S \frac{\partial}{\partial p} \rho(p, t)$$

$$= \int dp \rho(p, t) - \langle \delta p \rangle \int dp \frac{\partial}{\partial p} \rho(p, t) + \dots$$



## calculation of the effect of $dp/dt = \dots + \eta(t)$

Recall our old calculation of  $\partial_t \rho$ . In this situation, we are after an understanding of the behavior of the momentum  $p$ . We have two ways of calculating the average of a function of momentum at time  $t + \delta t$ . The first comes from computing  $\int dp g(p) \rho(p, t + \delta t)$ . That same average is obtained by taking the solution at time  $t + \delta t$ , which is of the form  $p(t + \delta t) = p(t) + \delta p$ . Here,  $\delta p$  is given by the effect of the stochastic term, so that

$$\delta p = \int_t^{t+\delta t} ds \eta(s)$$

Then the average at time  $t + \delta t$  can also be written as  $\langle \int dp g(p + \delta p) \rho(p, t) \rangle$ . Here, the average  $\langle \dots \rangle$  is an average over the possible values of the stochastic variables  $\eta(s)$  for  $s$  between  $t$  and  $t + \delta t$ . If we equate these two expressions, we find

$$\int dp g(p) \rho(p, t + \delta t) = \langle \int dp g(p + \delta p) \rho(p, t) \rangle$$

The right hand side can be rearranged by shifting the origin of integration and replacing the variable  $p$  by  $p - \delta p$ . Then this right side becomes  $\langle \int dp g(p) \rho(p - \delta p, t) \rangle$ , and

$$\int dp g(p) \rho(p, t + \delta t) = \langle \int dp g(p) \rho(p - \delta p, t) \rangle$$

Since  $g(p)$  is arbitrary, we it follows that

## Effect of Brownian motion: toward a unique probability distribution

We put together our two different pieces of the Brownian time derivative equation and get:

$$\partial_t \rho(p,x,t) = \dots + (\Gamma/2) (\partial_p)^2 \rho(p,x,t) + \partial_p [p/\tau \rho(p,x,t)]$$

The ....'s refer to terms which might come from usual Hamiltonian mechanics. We shall put them aside for a moment. An equation, like to one here, obtained from averaging the stochastic mechanics of a Langevin equation, is called a **Fokker-Planck equation**.

We look for a time-independent solution of this equation to see what is the equilibrium behavior. Write  $\rho(p,x,t) = \exp[-Q(p,x)]$  and find

$$0 = \dots + \partial_p [ (\Gamma/2) (-\partial_p Q) + p/\tau ]$$

which has the solution  $Q = p^2 / (\Gamma \tau) + C(x)$ , where  $C(x)$  is a “constant” of integration. To get the usual Hamiltonian result, use equation v.6 to replace  $\Gamma \tau$ . Also, write  $C(x) = U(x)/(kT)$  since that is what comes from the usual Hamiltonian mechanics. Then,  $Q$  becomes  $p^2/(2MkT) + U(x)/(kT)$ , which is exactly the result of Hamiltonian mechanics.

Einstein's result shows that, in order to get the Maxwell-Boltzmann equilibrium result, we have to go beyond Hamiltonian mechanics and include some **stochastic** behavior. **This is surprise.**

$$\partial_x \rho(p, x, A) = \dots$$

$$+ \frac{p}{2} \partial_p \left[ -\partial_p + \frac{p}{\hbar} \right] \rho$$

+ ...

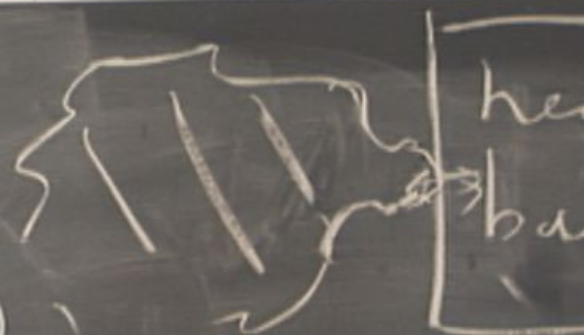




$$\partial_{\mu} \rho(p, x, A) =$$

$$+ \frac{p}{2} \partial_p \left[ -\partial_p + \frac{p}{\pi} \right] \rho$$

$$+ \dots - p = l^Q$$



$$\partial_{\mu} \rho(p, x, A) = \dots$$

$$+ \frac{p}{2} \partial_p \left[ -\partial_p + \frac{p}{\pi} \right] \rho$$



$$+ \dots - p = \ell$$

$$= \frac{p}{2} \partial_p e^{-Q(p, x)} \left[ -\partial_p Q(p, x) + \frac{p}{\pi} \right] = 0$$

## Effect of Brownian motion: toward a unique probability distribution

We put together our two different pieces of the Brownian time derivative equation and get:

$$\partial_t \rho(p,x,t) = \dots + (\Gamma/2) (\partial_p)^2 \rho(p,x,t) + \partial_p [p/\tau \rho(p,x,t)]$$

The ....'s refer to terms which might come from usual Hamiltonian mechanics. We shall put them aside for a moment. An equation, like to one here, obtained from averaging the stochastic mechanics of a Langevin equation, is called a **Fokker-Planck equation**.

We look for a time-independent solution of this equation to see what is the equilibrium behavior. Write  $\rho(p,x,t) = \exp[-Q(p,x)]$  and find

$$0 = \dots + \partial_p [ (\Gamma/2) (-\partial_p Q) + p/\tau ]$$

which has the solution  $Q = p^2 / (\Gamma \tau) + C(x)$ , where  $C(x)$  is a “constant” of integration. To get the usual Hamiltonian result, use equation v.6 to replace  $\Gamma \tau$ . Also, write  $C(x) = U(x)/(kT)$  since that is what comes from the usual Hamiltonian mechanics. Then,  $Q$  becomes  $p^2/(2MkT) + U(x)/(kT)$ , which is exactly the result of Hamiltonian mechanics.

Einstein's result shows that, in order to get the Maxwell-Boltzmann equilibrium result, we have to go beyond Hamiltonian mechanics and include some **stochastic** behavior. **This is surprise.**



$$\partial_p \rho(p, x, A) = \dots$$

$$+ \frac{p}{2} \partial_p \left[ -\partial_p + \frac{p}{\pi} \right] \rho$$



$$+ \dots - p = \ell$$

$$= \partial_p e^{-Q(p, x)} \left[ -\frac{p}{2} \partial_p Q(p, x) + \frac{p}{\pi} \right] = 0$$

$$\partial_p \rho(p, x, A) = \dots$$

$$+ \frac{p}{2} \partial_p \left[ -\partial_p + \frac{p}{\hbar} \right] \rho$$



$-Q(p, x)$

$+ \dots - p = \ell$

$$\partial_p e^{-Q(p, x)} \left[ -\frac{p}{i} \partial_p Q(p, x) + \frac{p}{\hbar} \right] = 0$$

$\equiv 0$

$$\partial_x \rho(p, x, A) = \left( \frac{p}{\hbar} \partial_x - \partial_x V \partial_p \right) \rho$$

$$+ \frac{p}{2} \partial_p \left[ -\partial_p + \frac{p}{\hbar} \right] \rho$$

$$- Q(p, x), \quad p = \ell$$

$$\partial_p e^{-Q(p, x)} \left[ -\frac{p}{\hbar} \partial_p Q(p, x) + \frac{p}{\hbar} \right] = 0$$

0



## Effect of Brownian motion: toward a unique probability distribution

We put together our two different pieces of the Brownian time derivative equation and get:

$$\partial_t \rho(p,x,t) = \dots + (\Gamma/2) (\partial_p)^2 \rho(p,x,t) + \partial_p [p/\tau \rho(p,x,t)]$$

The ....'s refer to terms which might come from usual Hamiltonian mechanics. We shall put them aside for a moment. An equation, like to one here, obtained from averaging the stochastic mechanics of a Langevin equation, is called a **Fokker-Planck equation**.

We look for a time-independent solution of this equation to see what is the equilibrium behavior. Write  $\rho(p,x,t) = \exp[-Q(p,x)]$  and find

$$0 = \dots + \partial_p [ (\Gamma/2) (-\partial_p Q) + p/\tau ]$$

which has the solution  $Q = p^2 / (\Gamma \tau) + C(x)$ , where  $C(x)$  is a “constant” of integration. To get the usual Hamiltonian result, use equation v.6 to replace  $\Gamma \tau$ . Also, write  $C(x) = U(x)/(kT)$  since that is what comes from the usual Hamiltonian mechanics. Then,  $Q$  becomes  $p^2/(2MkT) + U(x)/(kT)$ , which is exactly the result of Hamiltonian mechanics.

Einstein's result shows that, in order to get the Maxwell-Boltzmann equilibrium result, we have to go beyond Hamiltonian mechanics and include some **stochastic** behavior. **This is surprise.**