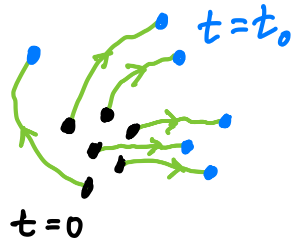
Unit 2-3: Liouville's Theorem

The concept of the density matrix will soon be expanded beyond the particular example of the *microcanonical* ensemble discussed in the previous section. It can also be generalized to *non-equilibrium* situations, where the density matrix varies with time, $\rho(q_i, p_i, t)$. We therefore want to see what general condition ρ must satisfy in order that $\dfrac{\partial \rho}{\partial t} = 0$, and so ρ is describing a steady, time-independent, state.



Consider an initial density ρ of points in phase space. As the systems represented by these initial points evolve in time, their trajectories give the density $\rho(t)$ at later times. Think of the points in ρ like particles in a fluid. The probability density ρ must obey a local *conservation* equation (think of the charge conservation equation of E&M),

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{u}) = 0 \tag{2.3.1}$$

where $\mathbf{u}$ is the "velocity" vector of the probability current $\rho\mathbf{u}$, that tells how the points in ρ flow in the $6N$ dimensional phase space.

The vector $\mathbf{u}$ is the $6N$ dimensional vector $\mathbf{u} \equiv (\dot{q}_1, \ldots, \dot{q}_{3N}, \dot{p}_1, \ldots, \dot{p}_{3N})$, and $\boldsymbol{\nabla} \equiv \left(\dfrac{\partial}{\partial q_1}, \ldots, \dfrac{\partial}{\partial q_{3N}}, \dfrac{\partial}{\partial p_1}, \ldots, \dfrac{\partial}{\partial p_{3N}} \right)$, so

$$\boldsymbol{\nabla} \cdot (\rho \mathbf{u}) \equiv \sum_{i=1}^{3N} \left[\frac{\partial}{\partial q_i}(\rho \dot{q}_i) + \frac{\partial}{\partial p_i}(\rho \dot{p}_i) \right] = \sum_{i=1}^{3N} \left[\frac{\partial \rho}{\partial q_i}\dot{q}_i + \rho \frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \rho}{\partial p_i}\dot{p}_i + \rho \frac{\partial \dot{p}_i}{\partial p_i} \right] \tag{2.3.2}$$

$$= \sum_{i=1}^{3N} \left(\left[\frac{\partial \rho}{\partial q_i}\dot{q}_i + \frac{\partial \rho}{\partial p_i}\dot{p}_i \right] + \rho \left[\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right] \right) \tag{2.3.3}$$

Now from Hamilton's equations of motion,

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \qquad \Rightarrow \qquad \frac{\partial \dot{q}_i}{\partial q_i} = \frac{\partial^2 \mathcal{H}}{\partial q_i \partial p_i}, \quad \frac{\partial \dot{p}_i}{\partial p_i} = -\frac{\partial^2 \mathcal{H}}{\partial p_i \partial q_i} \qquad \Rightarrow \qquad \frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} = 0 \tag{2.3.4}$$

and so,

$$\boldsymbol{\nabla} \cdot (\rho \mathbf{u}) = \sum_{i=1}^{3N} \left[\frac{\partial \rho}{\partial q_i}\dot{q}_i + \frac{\partial \rho}{\partial p_i}\dot{p}_i \right] = \sum_{i=1}^{3N} \left[\frac{\partial \rho}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \rho}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q_i} \right] \equiv [\rho, \mathcal{H}] \tag{2.3.5}$$

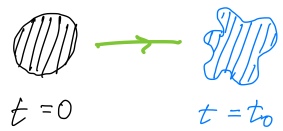
where $[\rho, \mathcal{H}]$ defines the *Poisson bracket* of the two observables ρ and $\mathcal{H}$ (in the correspondence of classical to quantum mechanics, the Poisson bracket becomes the commutator).

So the equation of conservation of probability in phase space, Eq. (2.3.1), becomes

$$\frac{\partial \rho}{\partial t} + [\rho, \mathcal{H}] = 0 \qquad \text{or} \qquad \boxed{\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left[\frac{\partial \rho}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial \rho}{\partial p_i}\frac{dp_i}{dt} \right] \equiv \frac{d\rho}{dt} = 0} \tag{2.3.6}$$

This is *Liouville's theorm*. Here the *total derivative* $d\rho/dt$, sometimes called the *convective derivative*, is just the total derivative of $\rho(q_i(t), p_i(t), t)$ with respect to t; $d\rho/dt$ tells how the value of ρ changes in time as seen by an observer who travels along with the system on its trajectory $\{q_i(t), p_i(t)\}$.

fluid



density of fluid stays constant
so volume of fluid stays constant
as drop deforms

Liouville's theorem, that $d\rho/dt = 0$, therefore says that the probability density in phase space ρ stays constant in time *as one flows along with the density*, just like the behavior of an incompressible fluid. This is a consequence of the probability conservation law of Eq. (2.3.1).

However, for ρ to describe *equilibrium*, the probability density must obey the stronger condition that $\partial\rho/\partial t = 0$, i.e. the probability for the system to be at any fixed point $\{q_i, p_i\}$ in phase space stays constant in time. Only when $\partial\rho/\partial t = 0$ will ensemble averages be independent of time.

To have $\dfrac{\partial\rho}{\partial t} = 0 \quad \Rightarrow \quad [\rho, \mathcal{H}] = 0$, and so for equilibrium ρ must satisfy,

$$[\rho, \mathcal{H}] = \sum_{i=1}^{3N} \left[\frac{\partial\rho}{\partial q_i}\frac{\partial\mathcal{H}}{\partial p_i} - \frac{\partial\rho}{\partial p_i}\frac{\partial\mathcal{H}}{\partial q_i} \right] = 0 \tag{2.3.7}$$

We will have $[\rho, \mathcal{H}] = 0$ provided $\rho(q_i, p_i)$ depends on the $\{q_i, p_i\}$ only via the function $\mathcal{H}[q_i, p_i]$, i.e. if $\rho = \rho(\mathcal{H}[q_i, p_i])$. Then we have,

$$\frac{\partial\rho}{\partial q_i} = \frac{\partial\rho}{\partial\mathcal{H}}\frac{\partial\mathcal{H}}{\partial q_i} \qquad \text{and} \qquad \frac{\partial\rho}{\partial p_i} = \frac{\partial\rho}{\partial\mathcal{H}}\frac{\partial\mathcal{H}}{\partial p_i} \tag{2.3.8}$$

so that,

$$[\rho, \mathcal{H}] = \sum_{i=1}^{3N} \left[\frac{\partial\rho}{\partial q_i}\frac{\partial\mathcal{H}}{\partial p_i} - \frac{\partial\rho}{\partial p_i}\frac{\partial\mathcal{H}}{\partial q_i} \right] = \sum_{i=1}^{3N} \left[\frac{\partial\rho}{\partial\mathcal{H}}\frac{\partial\mathcal{H}}{\partial q_i}\frac{\partial\mathcal{H}}{\partial p_i} - \frac{\partial\rho}{\partial\mathcal{H}}\frac{\partial\mathcal{H}}{\partial p_i}\frac{\partial\mathcal{H}}{\partial q_i} \right] = 0 \tag{2.3.9}$$

We already saw one example of such an equilibrium density matrix,

$$\rho(q_i, p_i) = C\,\delta(\mathcal{H}[q_i - p_i] - E) \qquad \textit{the microcanonical ensemble} \tag{2.3.10}$$

Anther choice that we will soon see is,

$$\rho(q_i, p_i) = C\,\mathrm{e}^{-\mathcal{H}[q_i, p_i]/k_B T} \qquad \textit{the canonical ensemble} \tag{2.3.11}$$

where in both cases C is an appropriate normalization constant.