

系综理论

June 14, 2018

[1] For a given macrostate (N, V, E) , a statistical system, at any time t , is equally likely to be in any one of an extremely large number of distinct microstates. As time passes, the system continually switches from one microstate to another. Over a reasonable span of time, all one observes is a behavior “averaged” over the variety of microstates through which the system passes. we can consider, at a single instant of time, a rather large number of systems - all being some sort of “mental copies” of the given system - which are characterized by the same macrostate as the original system but are, naturally enough, in all sorts of possible microstates. We expect that the average behavior of any system in this collection, which we call an ensemble, would be identical to the time-averaged behavior of the given system.

1 经典统计系综的概念

对于自由度为 s 的力学系统，其微观状态用 $2s$ 个变量，即 s 个广义坐标和 s 个广义动量描写：

$$q_1, \cdots, q_s; p_1, \cdots, p_s \tag{1}$$

系统的某一个确定的微观状态对应相空间中的一个点，称为**代表点**。系统微观状态随时间的变化遵从正则方程：

$$\begin{cases} \dot{q}_i = \frac{\partial H}{\partial p_i} \\ \dot{p}_i = -\frac{\partial H}{\partial q_i} \end{cases} \quad (2)$$

其中 $H = H(q_1, \dots, q_s; p_1, \dots, p_s)$ 为系统的哈密顿量。 H 除依赖于 (q_1, \dots, q_s) 以外，还可以依赖于若干外参量。 $(q_1, \dots, q_s; p_1, \dots, p_s)$ 是代表点在相空间中的位置， $(\dot{q}_1, \dots, \dot{q}_s; \dot{p}_1, \dots, \dot{p}_s)$ 是代表点在相空间中的运动速度。**相轨道**是由方程 (2) 决定的系统微观状态随时间的变化在相空间中所描出的轨迹。对于保守力学系统，(H 不显含时间)， H 及其微商是 $(q_1, \dots, q_s; p_1, \dots, p_s)$ 的单值函数。相轨道的运动方向完全由 $(\dot{q}_1, \dots, \dot{q}_s; \dot{p}_1, \dots, \dot{p}_s)$ 决定，经过相空间中任何一点只能有一条相轨道。相空间中，由不同初态出发的相轨道彼此不可能相交。

相体元：

$$d\Omega = dq_1 \cdots dq_s dp_1 \cdots dp_s \quad (3)$$

令

$$\rho(q_1, \dots, q_s; p_1, \dots, p_s, t) dq_1 \cdots dq_s dp_1 \cdots dp_s = \rho(q_1, \dots, q_s; p_1, \dots, p_s, t) d\Omega \quad (4)$$

代表 t 时刻系统的微观状态处于 $d\Omega$ 内的几率。其中 ρ 称为**系统微观状态的几率分布函数**或**几率密度**，它依赖于系统的微观状态，一般而言还与时间有关 (本节的讨论不限于平衡态)，并满足归一化条件：

$$\int \rho(q_1, \dots, q_s; p_1, \dots, p_s, t) d\Omega = \int \rho d\Omega = 1. \quad (5)$$

统计物理学的基本观点认为力学量的宏观观测值等于相应微观量对微观状态的统计平均值。令 $O = O(q_1, \dots, p_s)$ 代表任一力学量，

$$\bar{O} = \int O(q_1, \dots, p_s) \rho(q_1, \dots, p_s) d\Omega \quad (6)$$

不同微观状态在统计平均中的贡献由几率分布函数体现。

设想在一定的宏观条件下，系统所有可能的微观状态的总数为 \mathcal{N} (不要与系统所包含的粒子总数 N 相混淆。 \mathcal{N} 是一个极大的数，在经典统计中为 ∞)。令 $\tilde{\rho}d\Omega$ 代表微观状态处于 $d\Omega$ 内的态的数目，

$$\int \tilde{\rho}d\Omega = \mathcal{N} ,$$

而几率应正比于 $\tilde{\rho}d\Omega$ ，即

$$\rho d\Omega = \frac{\tilde{\rho}d\Omega}{\int \tilde{\rho}d\Omega} \propto \tilde{\rho}d\Omega . \quad (7)$$

把系统的每一个微观状态假想成一个处于该微观状态下的系统。由此， \mathcal{N} 个微观状态就对应于 \mathcal{N} 个假想的系统，它们各自处在相应的微观状态，微观状态处于 $d\Omega$ 内的系统数为 $\tilde{\rho}d\Omega$ 。这个假想的系统的集合称为**统计系综**，或简称为**系综**。

[2] The distribution of gas particles is most generally described as a six-dimensional (6-D) phase-space density, $f(x, y, z, p_x, p_y, p_z, t)$, known as the **distribution function**. It is directly related to **specific intensity**.

系综：假想的、和所研究的系统性质完全相同的、彼此独立、各自处于某一微观状态的大量系统的集合。系综是所研究的系统在一定宏观条件下的各种可能的微观状态的“化身”。 $\tilde{\rho}d\Omega$ 成为系综里微观状态处于 $d\Omega$ 内的系统数目，它也对应于相空间中相体元 $d\Omega$ 内的代表点数，而 $\tilde{\rho}$ 则是相空间中代表点密度。力学量对微观状态的统计平均可以完全等价地解释成对系综的平均。

[1] The microstate of a given classical system, at any time t , may be defined by specifying the instantaneous positions and momenta of all the particles constituting the system. If N is the number of particles in the system, the definition of a microstate requires the specification of $3N$ position coordinates q_1, q_2, \dots, q_{3N} and $3N$ momentum coordinates p_1, p_2, \dots, p_{3N} . The set of coordinates (q_i, p_i) , where $i = 1, 2, \dots, 3N$, may be regarded

as a point in a space of $6N$ dimensions. We refer to this space as the **phase space**, and the **phase point** (q_i, p_i) as a **representative point**, of the given system.

The coordinates q_i and p_i are functions of the time t and determined by the canonical equations of motion,

$$\begin{aligned}\dot{q}_i &= \frac{\partial H(q_i, p_i)}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H(q_i, p_i)}{\partial q_i}\end{aligned}\tag{8}$$

where $H(q_i, p_i)$ is the Hamiltonian of the system. As time passes, the set of coordinates (q_i, p_i) , which also defines the microstate of the system, undergoes a continual change. Correspondingly, our representative point in the phase space carves out a trajectory whose direction, at any time t , is determined by the velocity vector $\mathbf{v} \equiv (\dot{q}_i, \dot{p}_i)$, which in turn is given by the equations of motion.

An ensemble is said to be stationary if ρ does not depend explicitly on time, that is, at all times

$$\frac{\partial \rho}{\partial t} = 0 .\tag{9}$$

A stationary ensemble qualifies to represent a system in equilibrium.

[3] The simplest system to study is the dilute (nearly ideal) gas. A typical volume of gas contains of the order of 10^{23} particles, and kinetic theory tries to deduce the macroscopic properties of the gas from the time evolution of the set of atomic coordinates. At any time t , the **microstate of a system of N particles** is described by specifying the positions $\mathbf{q}_i(t)$, and momenta $\mathbf{p}_i(t)$, of all particles. The **microstate** corresponds to a **point $\mu(t)$** , in the $6N$ -dimensional **phase space** $\Gamma = \prod_{i=1}^N \{\mathbf{q}_i, \mathbf{p}_i\}$. The time evolution of this point is

governed by the **canonical equations**

$$\begin{aligned}\dot{q}_i &= \frac{\partial \mathcal{H}(q_i, p_i)}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial \mathcal{H}(q_i, p_i)}{\partial q_i}\end{aligned}\tag{10}$$

where the Hamiltonian $\mathcal{H}(\mathbf{q}, \mathbf{p})$ describes the total energy in terms of the set of coordinates $\mathbf{q} \equiv \{q_1, q_2, \dots, q_N\}$, and momenta $\mathbf{p} \equiv \{p_1, p_2, \dots, p_N\}$.

The microscopic equations of motion have **time reversal symmetry**, that is, **if all the momenta are suddenly reversed, $\mathbf{p} \rightarrow -\mathbf{p}$, at $t = 0$, the particles retrace their previous trajectory, $\mathbf{q}(t) = \mathbf{q}(-t)$** . This follows from the **invariance of \mathcal{H} under the transformation $T(\mathbf{p}, \mathbf{q}) \rightarrow (-\mathbf{p}, \mathbf{q})$** .

The macrostate M of an ideal gas in equilibrium is described by a small number of state functions such as E , T , P , and N . The space of macrostates is considerably smaller than the phase space spanned by microstates. Therefore, there must be a very large number of microstates μ corresponding to the same macrostate M .

This many-to-one correspondence suggests the introduction of a **statistical ensemble of microstates**. Consider \mathcal{N} **copies of a particular macrostate, each described by a different representative point $\mu(t)$ in the phase space Γ** . Let $d\mathcal{N}(\mathbf{q}, \mathbf{p}, t)$ equal the number of representative points in an infinitesimal volume $d\Gamma = \prod_{i=1}^N d^3p_i d^3q_i$ around the point (\mathbf{p}, \mathbf{q}) . A **phase space density $\rho(\mathbf{p}, \mathbf{q}, t)$** is defined by

$$\rho(\mathbf{p}, \mathbf{q}, t) d\Gamma = \lim_{\mathcal{N} \rightarrow \infty} \frac{d\mathcal{N}(\mathbf{p}, \mathbf{q}, t)}{\mathcal{N}} .\tag{11}$$

$\int d\Gamma \rho = 1$ and ρ is a properly normalized probability density function in phase space. To compute macroscopic values for various functions $\mathcal{O}(\mathbf{p}, \mathbf{q})$, the **ensemble averages** is

$$\langle \mathcal{O} \rangle = \int d\Gamma \rho(\mathbf{p}, \mathbf{q}, t) \mathcal{O}(\mathbf{p}, \mathbf{q}) .\tag{12}$$

When the **exact microstate μ is specified**, the system is said to be in a **pure state**. On the other hand, when our knowledge of the system is probabilistic, in the sense of its being taken from an ensemble with density $\rho(\Gamma)$, it is said to belong to a **mixed state**. It is **difficult to describe equilibrium in the context of a pure state**, since $\mu(t)$ is **constantly changing in time**. Equilibrium is more conveniently described for mixed states by examining the time evolution of the phase space density $\rho(t)$, which is governed by the Liouville equation.

1.1 Quantum states and the phase space

[1] The product of the uncertainties Δq and Δp in the simultaneous measurement of the canonically conjugate coordinates q and p would be of order \hbar :

$$(\Delta q \Delta p)_{\min} \sim \hbar \quad (13)$$

Around any point (q, p) in the (two-dimensional) phase space, there exists an area of order \hbar within which the position of the representative point cannot be pinpointed. In a phase space of $2N$ dimensions, the corresponding “volume of uncertainty” around any point would be of order \hbar^N . Therefore, it seems reasonable to regard the phase space as made up of elementary cells, of volume $\sim \hbar^N$, and to consider the various positions within such a cell as nondistinct.

2 Liouville 定理

Liouville 定理是关于系综几率密度 $\rho(q_1, \dots, p_s, t)$ 随时间演化所满足的方程。几率密度 $\rho(q_1, \dots, p_s, t)$ 与代表点密度 $\tilde{\rho}(q_1, \dots, p_s, t)$ 成正比。设 t 时刻，在相空间某处相体元

$d\Omega$ 内的一群代表点，其密度为 $\tilde{\rho}(q_1, \dots, q_s; p_1, \dots, p_s, t)$ 。当时间变到 $t + dt$ 时，这群代表点沿着由正则运动方程所规定的轨道各自独立地运动

$$t \longrightarrow t + dt, \quad (14)$$

$$q_i(t) \longrightarrow q_i(t + dt) = q_i + \dot{q}_i dt, \quad (15)$$

$$p_i(t) \longrightarrow p_i(t + dt) = p_i + \dot{p}_i dt, \quad (16)$$

这群代表点从相空间原来的区域 $d\Omega$ 移动到另区域 $d\Omega'$ ，后一处的密度为

$$\tilde{\rho}(q_1 + \dot{q}_1 dt, \dots, p_s + \dot{p}_s dt, t + dt),$$

$t + dt$ 与 t 时刻这两处的密度之差为

$$\begin{aligned} d\tilde{\rho} &= \tilde{\rho}(q_1 + \dot{q}_1 dt, \dots, p_s + \dot{p}_s dt, t + dt) - \tilde{\rho}(q_1, \dots, p_s, t) \\ &= \frac{\partial \tilde{\rho}}{\partial t} dt + \sum_i \left\{ \frac{\partial \tilde{\rho}}{\partial q_i} \dot{q}_i + \frac{\partial \tilde{\rho}}{\partial p_i} \dot{p}_i \right\} dt, \end{aligned} \quad (17)$$

即有

$$\frac{d\tilde{\rho}}{dt} = \frac{\partial \tilde{\rho}}{\partial t} + \sum_i \left\{ \frac{\partial \tilde{\rho}}{\partial q_i} \dot{q}_i + \frac{\partial \tilde{\rho}}{\partial p_i} \dot{p}_i \right\} \quad (18)$$

Liouville 定理：系综的几率密度（代表点密度）在运动中不变，即

$$\frac{d\rho}{dt} = 0 \quad (19)$$

或者

$$\frac{d\tilde{\rho}}{dt} = 0 \quad (20)$$

d/dt : “跟着代表点一起运动” 去观察的时间变化率。Liouville 定理，在相空间运动中代表点的密度不变，既不会变得更密集，也不会变得更稀释。

统计权重和状态数密度

Liouville 定理成立的条件要求系统是保守系，即 H 不显含时间。并且还要求在所考察的时间内系统不受外界作用的干扰（若外界作用可以用势场的形式表达且满足绝热近似，则 Liouville 定理仍然适用。非平衡态的线性响应理论是以 Liouville 定理为基础的）。Liouville 定理对平衡态基本统计假设（等几率原理或微正则系综）的提出是一个重要的依据。Liouville 定理也是关于相互作用多粒子系统的 BBGKY 以及非平衡态的线性响应理论的出发点。

[1] Consider an arbitrary “volume” ω in the relevant region of the phase space and let the “surface” enclosing this volume be denoted by σ . The rate at which the number of representative points in this volume increases with time is

$$\frac{\partial}{\partial t} \int_{\omega} \rho d\omega , \quad (21)$$

where $d\omega \equiv d^{3N}q d^{3N}p$. The net rate at which the representative points “flow” out of ω (across the bounding surface σ) is given by

$$\int_{\sigma} \rho \mathbf{v} \cdot \hat{\mathbf{n}} d\sigma = \int_{\omega} \nabla \cdot (\rho \mathbf{v}) d\omega , \quad (22)$$

where \mathbf{v} is the velocity vector of the representative points in the region of the surface element $d\sigma$ while $\hat{\mathbf{n}}$ is the (outward) unit vector normal to this element.

$$\nabla \cdot (\rho \mathbf{v}) = \sum_{i=1}^{3N} \left\{ \frac{\partial(\rho \dot{q}_i)}{\partial q_i} + \frac{\partial(\rho \dot{p}_i)}{\partial p_i} \right\} .$$

Since there are no “sources” or “sinks” in the phase space and hence the total number of representative points remains conserved,

$$\frac{\partial}{\partial t} \int_{\omega} \rho d\omega = - \int_{\omega} \nabla \cdot (\rho \mathbf{v}) d\omega , \quad (23)$$

$$\int_{\omega} \left\{ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \right\} d\omega = 0 \quad (24)$$

The necessary and sufficient condition that integral vanish for all arbitrary volumes ω is that the integrand itself vanish everywhere in the relevant region of the phase space, thus

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 , \quad (25)$$

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left\{ \frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right\} + \rho \sum_{i=1}^{3N} \left\{ \frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right\} = 0 . \quad (26)$$

which is the equation of continuity for the swarm of the representative points.

$$\frac{\partial \dot{q}_i}{\partial q_i} = \frac{\partial^2 H(q_i, p_i)}{\partial q_i \partial p_i} = \frac{\partial^2 H(q_i, p_i)}{\partial p_i \partial q_i} = -\frac{\partial \dot{p}_i}{\partial p_i} .$$

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + [\rho, H] = 0 . \quad (27)$$

The “local” density of the representative points, as viewed by an observer moving with a representative point, stays constant in time. The swarm of the representative points moves in the phase space in essentially the same manner as an incompressible fluid moves in the physical space.

A distinction must be made, however, between $\frac{d\rho}{dt} = 0$ and $\frac{\partial \rho}{\partial t} = 0$. While the former derives from the basic mechanics of the particles and is therefore quite generally true, the latter is only a requirement for equilibrium which, in a given case, may or may not be satisfied. The condition that ensures simultaneous validity of the two equations is

$$[\rho, H] = \sum_{i=1}^{3N} \left\{ \frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right\} = 0 . \quad (28)$$

one possible way of satisfying above condition is to assume that ρ , which is already assumed to have no explicit dependence on time, is independent of the coordinates (q, p) as well, that is,

$$\rho(q, p) = \text{const.} \quad (29)$$

over the relevant region of the phase space (and, of course, is zero everywhere else). Physically, this choice corresponds to an ensemble of systems that at all times are uniformly distributed over all possible microstates. The ensemble average then reduces to

$$\langle f \rangle = \frac{\int f(q, p) \rho(q, p, t) d^{3N}q d^{3N}p}{\int \rho(q, p, t) d^{3N}q d^{3N}p} = \frac{1}{\omega} \int f(q, p) d\omega , \quad (30)$$

ω denotes the total “volume” of the relevant region of the phase space. In this case, any member of the ensemble is equally likely to be in any one of the various possible microstates, inasmuch as any representative point in the swarm is equally likely to be in the neighborhood of any phase point in the allowed region of the phase space. This statement is usually referred to as the postulate of “equal a priori probabilities” for the various possible microstates (or for the various volume elements in the allowed region of the phase space). The resulting ensemble is referred to as the **microcanonical ensemble**.

A more general way is to assume that the dependence of ρ on (q, p) comes only through an explicit dependence on the Hamiltonian $H(q, p)$,

$$\rho(q, p) = \rho[H(q, p)] . \quad (31)$$

It provides a class of density functions for which the corresponding ensemble is stationary.

The most natural choice in this class of ensembles is the one for which

$$\rho(q, p) \propto \exp \left[-\frac{H(q, p)}{kT} \right] . \quad (32)$$

The ensemble is referred to as the **canonical ensemble**.

[4]

[5] For classical systems that averaging over the energy surface¹ is consistent with time evolution. Liouville’s theorem will tell us that volume in phase space is conserved, so the

¹The energy surface should be thought of as the energy shell $E < \mathcal{H}(\mathbb{P}, \mathbb{Q}) < E + \delta E$ in the limit

trajectories only stir the energy surface around, they do not change the relative weights of different parts of the surface.

Treating all states in phase space with a given energy on an equal footing gave sensible predictions for the ideal gas. Liouville's theorem, true for all Hamiltonian systems, will tell us that all states are created equal. Systems of point particles obeying Newton's laws without dissipation are examples of Hamiltonian dynamical systems. These systems conserve the energy, given by the Hamiltonian $\mathcal{H}(\mathbb{P}, \mathbb{Q})$. The laws of motion are given from \mathcal{H} by Hamilton's equations:

$$\dot{q}_\alpha = \frac{\partial \mathcal{H}}{\partial p_\alpha} , \quad (33)$$

$$\dot{p}_\alpha = -\frac{\partial \mathcal{H}}{\partial q_\alpha} . \quad (34)$$

Consider a bunch of particles interacting with a potential energy U :

$$\mathcal{H}(\mathbb{P}, \mathbb{Q}) = \sum_{\alpha} \frac{p_{\alpha}^2}{2m_{\alpha}} + U(q_1, \dots, q_{3N}) . \quad (35)$$

The expected Newtonian equations of motion:

$$\dot{q}_\alpha = \frac{\partial \mathcal{H}}{\partial p_\alpha} = \frac{p_\alpha}{m_\alpha} , \quad (36)$$

$$\dot{p}_\alpha = -\frac{\partial \mathcal{H}}{\partial q_\alpha} = -\frac{\partial U}{\partial q_\alpha} = f_\alpha(q_1, \dots, q_{3N}) , \quad (37)$$

where f_α is the force on coordinate α . Hamiltonian systems have properties that are quite distinct from general systems of differential equations.

Consider the evolution law for a general probability density in phase space:

$$\rho(\mathbb{P}, \mathbb{Q}) = \rho(q_1, \dots, q_{3N}, p_1, \dots, p_{3N}) , \quad (38)$$

$\delta E \rightarrow 0$. We focus here on the energy surface rather than the energy shell because the different energies in the shell do not get intermingled by the time evolution.

The microcanonical ensemble has ρ equal to a constant in a thin range of energies, and zero outside that range. This probability density ρ is locally conserved: probability cannot be created or destroyed, it can only flow around in phase space. The probability density in $6N$ dimensions has a phase-space probability current $(\rho\dot{\mathbb{P}}, \rho\dot{\mathbb{Q}})$ and hence satisfies a continuity equation

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= - \sum_{\alpha=1}^{3N} \left(\frac{\partial(\rho \dot{q}_{\alpha})}{\partial q_{\alpha}} + \frac{\partial(\rho \dot{p}_{\alpha})}{\partial p_{\alpha}} \right) \\ &= - \sum_{\alpha=1}^{3N} \left(\frac{\partial \rho}{\partial q_{\alpha}} \dot{q}_{\alpha} + \rho \frac{\partial \dot{q}_{\alpha}}{\partial q_{\alpha}} + \frac{\partial \rho}{\partial p_{\alpha}} \dot{p}_{\alpha} + \rho \frac{\partial \dot{p}_{\alpha}}{\partial p_{\alpha}} \right) \\ \frac{\partial \rho}{\partial t} + \sum_{\alpha=1}^{3N} \left(\frac{\partial \rho}{\partial q_{\alpha}} \dot{q}_{\alpha} + \frac{\partial \rho}{\partial p_{\alpha}} \dot{p}_{\alpha} \right) &= \frac{d\rho}{dt} = 0 .\end{aligned}\tag{39}$$

which is Liouville's theorem. $\frac{d\rho}{dt}$ is called the total derivative of ρ with respect to time; it is the evolution of ρ seen by a particle moving with the flow. In a three-dimensional flow, $\frac{d\rho_{3N}}{dt} = \frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho = \frac{\partial \rho}{\partial t} + \sum_{i=1}^3 \dot{x}_i \left(\frac{\partial \rho}{\partial x_i} \right)$; the first term is the change in ρ due to the time evolution at fixed position, and the second is the change in ρ that a particle moving with velocity \mathbf{v} would see if the ρ field did not change in time.

Flows in phase space are incompressible. In fluid mechanics, if the density $\frac{d\rho_{3N}}{dt} = 0$ it means that the fluid is incompressible. The density of a small element of fluid does not change as it moves around in the fluid; hence the small element is not compressing or expanding. In Liouville's theorem, it means the same thing; a small volume in phase space will evolve into a new shape, perhaps stretched, twisted, or folded, but with exactly the same volume.

There are no attractors. In other dynamical systems, most states of the system are usually transient, and the system settles down onto a small set of states called the **attractor**. A

damped pendulum will stop moving; the attractor has zero velocity and vertical angle. A forced, damped pendulum will settle down to oscillate with a particular amplitude; the attractor is a circle in phase space. The decay of these transients in dissipative systems would seem closely related to equilibration in statistical mechanics, where at long times all initial states of a system will settle down into static equilibrium behavior. Perversely, we have just proven that equilibration in statistical mechanics happens by a completely different mechanism! In equilibrium statistical mechanics all states are created equal; transient states are temporary only insofar as they are very unusual, so as time evolves they disappear, to arise again only as rare fluctuations.

Microcanonical ensembles are time independent. An initial uniform density in phase space will stay uniform. More generally, since energy is conserved, a uniform density over a small shell of energies $(E, E + \delta E)$ will stay uniform.

Liouville's theorem tells us that the energy surface may get stirred around, but the relative weights of parts of the surface are given by their phase-space volumes and do not change. This property is a necessary condition for our microcanonical ensemble to describe the time-independent equilibrium state.

[3] Liouville's theorem states that the phase space density $\rho(\Gamma, t)$ behaves like an incompressible fluid.

Follow the evolution of $d\mathcal{N}$ pure states in an infinitesimal volume $d\Gamma = \prod_{i=1}^N d^3p d^3q$ around the point (\mathbf{p}, \mathbf{q}) . After an interval δt these states have moved to the vicinity of

another point $(\mathbf{p}', \mathbf{q}')$, where

$$q'_\alpha = q_\alpha + \dot{q}_\alpha \delta t + \mathcal{O}(\delta t^2) , \quad (40)$$

$$p'_\alpha = p_\alpha + \dot{p}_\alpha \delta t + \mathcal{O}(\delta t^2) . \quad (41)$$

In the above expression, the q_α and p_α refer to any of the $6N$ coordinates and momenta, and \dot{q}_α and \dot{p}_α are the corresponding velocities. The original volume element $d\Gamma$ is in the shape of a hypercube of sides dp_α and dq_α . In the time interval δt it gets distorted, and the **projected sides of the new volume element** are given by

$$dq'_\alpha = dq_\alpha + \frac{\partial \dot{q}_\alpha}{\partial q_\alpha} dq_\alpha \delta t + \mathcal{O}(\delta t^2) , \quad (42)$$

$$dp'_\alpha = dp_\alpha + \frac{\partial \dot{p}_\alpha}{\partial p_\alpha} dp_\alpha \delta t + \mathcal{O}(\delta t^2) . \quad (43)$$

To order of δt^2 , the new volume element is $d\Gamma' = \prod_{i=1}^N d^3 p' d^3 q'$. For each pair of conjugate coordinates

$$dq'_\alpha \cdot dp'_\alpha = dq_\alpha \cdot dp_\alpha \left[1 + \left(\frac{\partial \dot{q}_\alpha}{\partial q_\alpha} + \frac{\partial \dot{p}_\alpha}{\partial p_\alpha} \right) \delta t + \mathcal{O}(\delta t^2) \right] . \quad (44)$$

since the time evolution of coordinates and momenta are governed by the canonical,

$$\frac{\partial \dot{q}_\alpha}{\partial q_\alpha} = \frac{\partial}{\partial q_\alpha} \frac{\partial \mathcal{H}}{\partial p_\alpha} = \frac{\partial^2 \mathcal{H}}{\partial p_\alpha \partial q_\alpha} , \quad (45)$$

$$\frac{\partial \dot{p}_\alpha}{\partial p_\alpha} = - \frac{\partial}{\partial p_\alpha} \frac{\partial \mathcal{H}}{\partial q_\alpha} = - \frac{\partial^2 \mathcal{H}}{\partial p_\alpha \partial q_\alpha} . \quad (46)$$

The projected area in Eq. (44) is **unchanged** for any pair of coordinates, and hence the volume element is unaffected, $d\Gamma = d\Gamma'$. All the pure states $d\mathcal{N}$ originally in the vicinity of (\mathbf{p}, \mathbf{q}) are transported to the neighborhood of $(\mathbf{p}', \mathbf{q}')$, but occupy exactly the same volume. The ratio $d\mathcal{N}/d\Gamma$ is left unchanged, and ρ behaves like the density of an incompressible fluid.

The incompressibility condition $\rho(\mathbf{p}', \mathbf{q}', t + \delta t) = \rho(\mathbf{p}, \mathbf{q}, t)$ can be written as

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \sum_{\alpha=1}^{3N} \left(\frac{\partial\rho}{\partial p_{\alpha}} \cdot \frac{dp_{\alpha}}{dt} + \frac{\partial\rho}{\partial q_{\alpha}} \cdot \frac{dq_{\alpha}}{dt} \right) = 0 . \quad (47)$$

$$\frac{\partial\rho}{\partial t} = \sum_{\alpha=1}^{3N} \left(\frac{\partial\rho}{\partial p_{\alpha}} \cdot \frac{\partial\mathcal{H}}{\partial q_{\alpha}} - \frac{\partial\rho}{\partial q_{\alpha}} \cdot \frac{\partial\mathcal{H}}{\partial p_{\alpha}} \right) = -\{\rho, \mathcal{H}\} . \quad (48)$$

where the Poisson bracket of two functions in phase space is introduced as

$$\{A, B\} \equiv \sum_{\alpha=1}^{3N} \left(\frac{\partial A}{\partial q_{\alpha}} \cdot \frac{\partial B}{\partial p_{\alpha}} - \frac{\partial A}{\partial p_{\alpha}} \cdot \frac{\partial B}{\partial q_{\alpha}} \right) = -\{B, A\} . \quad (49)$$

Note the distinction between $\frac{\partial\rho}{\partial t}$ and $\frac{d\rho}{dt}$: the former partial derivative refers to the changes in ρ at a particular location in phase space, while the latter total derivative follows the evolution of a volume of fluid as it moves in phase space.

1. Under the action of time reversal, $(\mathbf{p}, \mathbf{q}, t) \rightarrow (-\mathbf{p}, \mathbf{q}, -t)$, the Poisson bracket $\{\rho, \mathcal{H}\}$ changes sign, and Eq. (48) implies that the density reverses its evolution, that is, $\rho(\mathbf{p}, \mathbf{q}, t) = \rho(-\mathbf{p}, \mathbf{q}, -t)$.

2. The time evolution of the ensemble average is given by

$$\frac{d\langle\mathcal{O}\rangle}{dt} = \int d\Gamma \frac{\partial\rho(\mathbf{p}, \mathbf{q}, t)}{\partial t} \mathcal{O}(\mathbf{p}, \mathbf{q}) = \sum_{\alpha=1}^{3N} \int d\Gamma \mathcal{O}(\mathbf{p}, \mathbf{q}) \left(\frac{\partial\rho}{\partial p_{\alpha}} \cdot \frac{\partial\mathcal{H}}{\partial q_{\alpha}} - \frac{\partial\rho}{\partial q_{\alpha}} \cdot \frac{\partial\mathcal{H}}{\partial p_{\alpha}} \right) . \quad (50)$$

The partial derivatives of ρ in the above equation can be removed by using the method of integration by parts, that is, $\int f\rho' = -\int \rho f'$, since ρ vanishes on the boundaries of the integrations, leading to

$$\begin{aligned} \frac{d\langle\mathcal{O}\rangle}{dt} &= - \sum_{\alpha=1}^{3N} \int d\Gamma \rho \left[\left(\frac{\partial\mathcal{O}}{\partial p_{\alpha}} \cdot \frac{\partial\mathcal{H}}{\partial q_{\alpha}} - \frac{\partial\mathcal{O}}{\partial q_{\alpha}} \cdot \frac{\partial\mathcal{H}}{\partial p_{\alpha}} \right) + \mathcal{O} \left(\frac{\partial^2\mathcal{H}}{\partial p_{\alpha}\partial q_{\alpha}} - \frac{\partial^2\mathcal{H}}{\partial q_{\alpha}\partial p_{\alpha}} \right) \right] \\ &= - \int d\Gamma \rho \{\mathcal{H}, \mathcal{O}\} = \langle \{\mathcal{O}, \mathcal{H}\} \rangle . \end{aligned} \quad (51)$$

Note that the total time derivative cannot be taken inside the integral sign, that is,

$$\frac{d\langle\mathcal{O}\rangle}{dt} \neq \int d\Gamma \frac{d\rho(\mathbf{p}, \mathbf{q}, t)}{dt} \mathcal{O}(\mathbf{p}, \mathbf{q}) . \quad (52)$$

This common mistake yields $d\langle\mathcal{O}\rangle/dt = 0$.

3. If the members of the ensemble correspond to an equilibrium macroscopic state, the ensemble averages must be independent of time. This can be achieved by a stationary density $\frac{\partial\rho_{\text{eq}}}{\partial t} = 0$, that is, by requiring

$$\{\rho_{\text{eq}}, \mathcal{H}\} = 0 . \quad (53)$$

A possible solution to the above equation is for ρ_{eq} to be a function of \mathcal{H} , that is, $\rho_{\text{eq}}(\mathbf{p}, \mathbf{q}) = \rho(\mathcal{H}(\mathbf{p}, \mathbf{q}))$. It is then easy to verify that $\{\rho(\mathcal{H}), \mathcal{H}\} = \rho'(\mathcal{H})\{\mathcal{H}, \mathcal{H}\} = 0$. This solution implies that the value of ρ is constant on surfaces of constant energy \mathcal{H} in phase space. This is indeed the basic assumption of statistical mechanics. For example, in the microcanonical ensemble, the total energy E of an isolated system is specified. All members of the ensemble must then be located on the surface $\mathcal{H}(\mathbf{p}, \mathbf{q}) = E$ in phase space. Equation (48) implies that a uniform density of points on this surface is stationary in time. The assumption of statistical mechanics is that the macrostate is indeed represented by such a uniform density of microstates. This is equivalent to replacing the objective measure of probability with a subjective one.

There may be additional conserved quantities associated with the Hamiltonian that satisfy $\{L_n, \mathcal{H}\} = 0$. In the presence of such quantities, a stationary density exists for any function of the form $\rho_{\text{eq}}(\mathbf{p}, \mathbf{q}) = \rho(\mathcal{H}(\mathbf{p}, \mathbf{q}), L_1(\mathbf{p}, \mathbf{q}), L_2(\mathbf{p}, \mathbf{q}), \dots)$. The value of L_n is

not changed during the evolution of the system, since

$$\begin{aligned}
\frac{dL_n(\mathbf{p}, \mathbf{q})}{dt} &\equiv \frac{L_n(\mathbf{p}(t+dt), \mathbf{q}(t+dt)) - L_n(\mathbf{p}(t), \mathbf{q}(t))}{dt} \\
&= \sum_{\alpha=1}^{3N} \left(\frac{\partial L_n}{\partial p_\alpha} \cdot \frac{\partial p_\alpha}{\partial t} + \frac{\partial L_n}{\partial q_\alpha} \cdot \frac{\partial q_\alpha}{\partial t} \right) \\
&= - \sum_{\alpha=1}^{3N} \left(\frac{\partial L_n}{\partial p_\alpha} \cdot \frac{\partial \mathcal{H}}{\partial q_\alpha} - \frac{\partial L_n}{\partial q_\alpha} \cdot \frac{\partial \mathcal{H}}{\partial p_\alpha} \right) = \{L_n, \mathcal{H}\} = 0 . \tag{54}
\end{aligned}$$

Hence, the functional dependence of ρ_{eq} on these quantities merely indicates that all accessible states, that is, those that can be connected without violating any conservation law, are equally likely.

4. Non-stationary densities need converge onto the stationary solution ρ_{eq} . This contradicts the time reversal symmetry noted in point (1) above: for any solution $\rho(t)$ converging to ρ_{eq} , there is a time-reversed solution that diverges from it. The best that can be hoped for is to show that the solutions $\rho(t)$ are in the neighborhood of ρ_{eq} the majority of the time, so that time averages are dominated by the stationary solution. This brings us to the problem of **ergodicity**, which is **whether it is justified to replace time averages with ensemble averages**. In measuring the properties of any system, we deal with only one representative of the equilibrium ensemble. However, most macroscopic properties do not have instantaneous values and require some form of averaging. For example, the pressure P exerted by a gas results from the impact of particles on the walls of the container. The number and momenta of these particles vary at different times and different locations. The measured pressure reflects an average over many characteristic microscopic times. If over this time scale the representative point of the system moves around and uniformly samples the accessible points in phase space, we may replace the time average with the ensemble average. For a few systems it is possible to prove an **ergodic theorem**,

which states that the representative point comes arbitrarily close to all accessible points in phase space after a sufficiently long time. However, the proof usually works for time intervals that grow exponentially with the number of particles N , and thus exceed by far any reasonable time scale over which the pressure of a gas is typically measured. As such the proofs of the ergodic theorem have so far little to do with the reality of macroscopic equilibrium.

2.1 Boltzmann Equation and Liouville's Theorem

[6] Define a function $f(\mathbf{r}, \mathbf{p}, t)$ as the density of particles in phase space. The number of particles in volume element $d\mathbf{r}$ at position \mathbf{r} , having momenta that lie in some momentum-space volume $d\mathbf{p}$ around momentum \mathbf{p} , is $f(\mathbf{r}, \mathbf{p}, t)d\mathbf{r}d\mathbf{p}$. We ask how the function f evolves with time. Since each particle in the assembly can be described in terms of three momentum and three spatial coordinates, the general form of the equation reads

$$\frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial \mathbf{r}_i} \frac{d\mathbf{r}_i}{dt} + \sum_i \frac{\partial f}{\partial \mathbf{p}_i} \frac{d\mathbf{p}_i}{dt} = \left. \frac{df}{dt} \right|_{\text{collisions}}. \quad (55)$$

The left side of this equation gives the time rate of change of particles in the volume element $d\mathbf{r}d\mathbf{p}$ as a function of the coordinates $\mathbf{r}_i, \mathbf{p}_i, i = 1, 2, 3, \dots, n$, for an n particle assembly. As the particles move, the surface enclosing them in phase space becomes distorted and the expression gives the rate of change of density through this distortion and through any other effects. The right side gives the loss or gain of particles through collisions. It is called the **Boltzmann equation**.

To see how the evolution proceeds for a collisionless process in which the right side of equation is zero, we draw a simple two-dimensional picture. We have an assembly of particles initially confined between positions \mathbf{r}_1 and \mathbf{r}_2 and between momentum values

\mathbf{p}_a and \mathbf{p}_b . Some time later, the momentum values are unchanged, but the particles have moved so that the higher momentum particles are now at positions between \mathbf{r}'_1 and \mathbf{r}'_2 while the lower momentum particles are at positions between \mathbf{r}''_1 and \mathbf{r}''_2 . However, because the base and height of the enclosing area has not changed, the number density of particles per unit area has remained constant.

A similar argument holds when forces are applied to the particles. In that case the momenta of particles are not constant and the parallelepiped will also be displaced in a vertical direction. However, a similar argument can then be applied to show that the area covered by the particles still remains constant and the density of particles in this two-dimensional representation is unchanged. This is particularly easy to see if the force is the same on all particles. In that case, $d\mathbf{p}/dt$ is uniform and the difference in values $\mathbf{p}_a - \mathbf{p}_b$ is maintained constant. When different forces are exerted on different constituents of the gas, the area occupied by each constituent remains constant. These results also hold when there are gradients in the force fields.

A further extension of the argument can be applied to the full six-dimensional distribution. Unless there are some means for creating or destroying particles in the assembly - through collisions or particle-antiparticle pair formation - the density of particles will be constant along the trajectory in the six-dimensional space.

This is the sense of **Liouville's theorem**: The **six-dimensional space density of particles in an assembly remains unchanged unless collisions occur**:

$$\frac{df}{dt} = 0 . \tag{56}$$

Liouville's theorem has interesting applications to cosmic-ray particles, which move through the Galaxy, guided by magnetic field lines. Many of these particles are so energetic that

they must be able to escape from the magnetic fields permeating the Galaxy. Their density in space outside the Galaxy could, therefore, be the same as the density that we measure in the vicinity of the Earth, provided the particles had enough time since their creation to traverse distances comparable to those of remote clusters of galaxies. Under these conditions, **the spatial density of cosmic rays in extragalactic space would be the same as that measured at the Earth.** This argument need not be true for low-energy particles if these particles can remain bottled up in local magnetic fields within our Galaxy.

The highest energy cosmic rays, whose energies range up to 10^{20} eV cannot be magnetically confined to the Galaxy. An extragalactic component could originate in gamma-ray bursts, active galactic nuclei, clusters of galaxies, or in as yet unidentified sources. The Liouville theorem tells us that **the local density of at least these particles should be an indication of their extragalactic density.** A slight caveat to this must, however, be kept in mind. Particles at such high energies cannot travel large distances across the Universe before they are destroyed in collisions with microwave background radiation photons. Liouville's theorem may then permit us to say only that the observed flux of these high-energy cosmic rays represents their prevalence just in our local part of the Universe, rather than everywhere in the cosmos.

Finally we should still mention the problem, where a swarm of particles moves through a gravitational field. There we were concerned with tidal disruption of globular clusters, but noted that while the clusters became extended along a direction pointing toward the Galactic center, the gravitational forces also tended to produce a compression lateral to that direction. This compression produces additional transverse velocities making the evo-

lutionary pattern quite complex. Liouville's theorem, however, gives us at least one solid guide toward understanding the overall development. It tells us that whatever detailed dynamical arguments we apply the results must always agree at least with Liouville's requirement of a constant phase space density.

2.2 Magnetic Mirrors, Magnetic Bottles, and Cosmic-Ray Particles

[6] A betatron accelerates charged particles. A different scheme for magnetically accelerating cosmic-ray particles was suggested by Enrico Fermi. In Fermi acceleration the cosmic-ray particles are thought to travel between cosmic gas clouds. Each cloud has an embedded magnetic field. When a particle approaches the cloud and enters its field perpendicular to the field direction, it is turned back by virtue of the magnetic force given by equation

$$\mathbf{F} = \frac{q\mathbf{v} \times \mathbf{B}}{c} . \quad (57)$$

For, after traveling in a semicircle, the particle once again finds itself at the edge of the cloud and headed into the direction from which it came.

A similar reflection can occur for particles approaching a cloud along the lines of force. If the particle impinges on a cloud that is receding from it, the particle's momentum after encounter is smaller than before. If the particle impinges on an approaching cloud, its final momentum is higher than before the collision. In general the probability for collision is greater for an approaching than for a receding cloud. (This corresponds to everyday experience. On a highway we pass more cars going in the opposite direction than going along the direction in which we are traveling.)

Statistically, therefore, particles will derive increased momentum from encounters with clouds and can be accelerated to high energies. The process is similar to the acceleration of a ping-pong ball between two slowly approaching paddles. After the ball has made many bounces off each paddle, it is going far faster than either of these reflecting surfaces. Generally, a charged particle moves along the lines of force of a magnetic field, spiraling as it goes. The pitch angle is given by equation

$$\tan \theta = \frac{v_c}{v_z} . \quad (58)$$

If the particle encounters a region of the magnetic field where the lines are more compressed, it experiences an increase in the field strength and by Faraday's law

$$\frac{1}{c} \frac{\partial}{\partial t} \oint \mathbf{B} \cdot d\mathbf{s} = - \oint \mathbf{E} \cdot d\mathbf{l} \quad (59)$$

its circular velocity v_c increases. However, because the field itself is not doing any work on the charge, the increase in v_c must be bought at the expense of kinetic energy initially resident in the longitudinal motion, that is, at the expense of a reduction in v_z . When the particle has advanced into the intense magnetic field to such a depth that all its kinetic energy is spent in circular motion, the pitch angle θ becomes $\pi/2$; the particle is reflected and spirals back out of the intense field.

As the particle first spirals into the intensifying magnetic field, its angular momentum about the axis of symmetry of the motion is conserved. Hence the magnetic moment \mathbf{M} :

$$\mathbf{M} = \frac{\mathbf{j} \times \mathbf{r}}{2c} , \quad \mathbf{j} = q\mathbf{v} , \quad (60)$$

is also conserved. Substituting the gyroradius, for r we find

$$M = \frac{v_c p_c}{2B} , \quad (61)$$

along the direction of the magnetic field. From this it follows that the transverse kinetic energy is directly proportional to the field B . If a particle has an initial pitch angle θ in a field B , it can only penetrate the field until it reaches a region where the field is B_0 and $\sin \theta = 1$:

$$B_0 = \frac{B}{\sin^2 \theta} . \quad (62)$$

Here it is reflected and spirals back out of the intense field.

A magnetic bottle consists of two such magnetic mirrors between which a particle is reflected going back and forth without possibility of escape. The Fermi mechanism ping-pong acceleration could involve a (shrinking) magnetic bottle in which the two magnetic mirrors approach.

We sometimes characterize cosmic-ray particles by a magnetic rigidity BR_L which equals pc/q for motion strictly perpendicular to the field. The rigidity has dimensions of energy per charge.

A more likely variant of the Fermi process has come to be widely accepted. The particles are accelerated by successive reflections off a rapidly expanding supernova ejection shell. After each impact the accelerated particle returns - deflected back toward the shell by an external magnetic field into which the shell is expanding. A succession of reflections off the expanding shell progressively accelerates the particle to cosmic-ray energies. This variant has also come to be referred to as **Fermi acceleration**.

A cosmic-ray particle must eventually suffer destruction due to one of several competing processes - inelastic impact on another particle, loss from the Galaxy, and hence loss from contact with the accelerating clouds, and so on. The number of accelerating reflections between destructive events is thus limited. To reach truly high energies, therefore, cosmic-

ray particles must be injected into the accelerating fields with rather high initial energies. Sufficiently energetic particles may be produced in supernova explosions. But the highest energy cosmic rays appear to reach the Galaxy from extragalactic sources.

The most energetic particles, having energies of order 10^{20} eV, have a gyroradius that exceeds the Galaxy's dimensions. These particles could therefore not be retained in the Galaxy for any length of time. If their origin were local, they should appear to come primarily from the central regions of the Galaxy or the Milky Way plane. Instead, many come from other directions and are therefore believed to be extragalactic.

We know that the heavy nuclei, which form an abundant part of the cosmic-ray flux, would suffer destructive collisions during the long stay in interstellar space required by the rather slow Fermi acceleration mechanism. Yet we find iron nuclei to be abundant at least up to energies of the order of 10^{12} to 10^{13} eV. We are driven toward a mechanism that could accelerate these particles rapidly.

Currently, quasars, the active nuclei of galaxies, gamma-ray bursts, pulsars, magnetars and supernova explosions slamming into the interstellar medium, or a fast stellar wind, are all considered to be likely sources for producing cosmic rays at various energies. Further observations should lead to a clarification of these hypotheses. We can also hope that the study of solar flares, which are responsible for the solar cosmic-ray component, will give us a better understanding of at least one mechanism for accelerating energetic particles.

[2] The phase-space density can vary with time, but, according to [Liouville's theorem](#), it is conserved in a frame of reference that travels in phase space with the particles, under certain conditions. The propagation of cosmic rays in the Galaxy generally satisfies these conditions.

Consider a distribution of particles moving in space and subject to “smooth” forces (no collisions). The particles flow through phase space as time proceeds. Let an **observer** travel alongside a selected small group of the particles as they propagate through phase space. Liouville’s theorem (not derived here) tells us that the **distribution function f in the observer’s frame of reference is conserved**. The phase-space density of particles will not change as they travel through phase space.

The **invariance of specific intensity I** in astronomical measurements is a **direct consequence of Liouville’s theorem**. The invariance of f and the equivalence of f and I tell us that **I is also conserved if one follows a group of photons of some frequency ν through space**. This is valid **given negligible absorption, scattering, or redshift**.

An important ramification of this is that the **emitted surface brightness $B(\nu, \theta, \phi)$ of an astronomical body** and the **detected specific intensity $I(\nu, \theta, \phi)$ from that same body** are, in effect, the **same quantity**; that is,

$$B(\nu, \theta, \phi) = I(\nu, \theta, \phi) \quad (\text{Equality of surface brightness and specific intensity}) \quad (63)$$

The **brightness B** is the **power emitted from unit-projected area of the surface into the unit-solid angle at ν, ϕ** whereas the **specific intensity I** is the **power received at the detector per unit-detector area and solid angle at ν, ϕ** . Both quantities have the same units ($\text{W m}^{-2} \text{ Hz}^{-1} \text{ sr}^{-1}$).

This **invariance of I with the distance to a diffuse celestial object** is a related example. The measured specific intensity (power per steradian) of the sun’s surface would not change if the sun were to be moved to twice its current distance. This is usually explained as the cancellation of the increased size (as distance squared) of the patch within the observer’s beam and the decreased radiation (inverse square of distance) received from each element

of the patch. Nevertheless, it is a consequence of Liouville's theorem.

The observer of Liouville's theorem follows the particles even as they are guided along curved paths by magnetic fields. This observer would thus be in a noninertial (i.e., accelerating) frame of reference to which the Lorentz transformations of special relativity do not apply. However, one can adopt an inertial (constant velocity) frame of reference that, at some moment, is instantaneously at rest with respect to the particles. In this case, the Lorentz transformations would apply.

One may then ask how phase-space density f transforms from this moving inertial frame to the laboratory (stationary) frame if the particles are moving at a speed approaching the speed of light c . We demonstrate f has the same value in two such frames; that is, f is a Lorentz invariant. Then I/ν^3 is also a Lorentz invariant.

Now, the frequency ν of photons is not a Lorentz invariant according to the relativistic Doppler shift. Hence, because I/ν^3 is a Lorentz invariant, the specific intensity I is not. The specific intensity emanating from a celestial object will be greatly modified if the source (e.g., a knot of radiation in a jet) is moving rapidly toward or away from the observer.

3 Ergodicity

[5] An ergodic system has an energy surface which is well stirred. Using Liouville's theorem and assuming ergodicity will allow us to show that the microcanonical ensemble average gives the long-time average behavior that we call equilibrium.

The energy surface is thoroughly stirred by the time evolution; it is not divided into some

kind of components that do not intermingle. A system which is thoroughly stirred is said to be **ergodic**².

Definition 1

In an ergodic system, the trajectory of almost every^a point in phase space eventually passes arbitrarily close^b to every other point (position and momentum) on the surface of constant energy.

^aWhat does almost every mean? Technically, it means all but a set of zero volume (measure zero). Basically, the qualification ‘almost’ is there to avoid problems with unusual, specific initial conditions like all the particles moving precisely at the same velocity in neat rows.

^bWhy not just assume that every point on the energy surface gets passed through? Boltzmann originally did assume this. However, it can be shown that a smooth curve (our time trajectory) cannot fill up a whole volume (the energy surface). In an ergodic system the trajectory covers the energy surface densely, but not completely.

The most important consequence of ergodicity is that time averages are equal to microcanonical averages³. Intuitively, since the trajectory $(\mathbb{P}(t), \mathbb{Q}(t))$ covers the whole energy surface, the average of any property $O(\mathbb{P}(t), \mathbb{Q}(t))$ over time is the same as the average of O over the energy surface.

The energy surface cannot be divided into components which do not intermingle. Define an ergodic component R of a set S ⁴ to be a subset that remains invariant under the flow (so $r(t) \in R$ for all $r(0) \in R$).

²Mathematicians distinguish between ergodic (stirred) and mixing (scrambled)

³If an ergodic system equilibrates (i.e., does not oscillate forever), the time average behavior will be determined by the equilibrium behavior; ergodicity then implies that the equilibrium properties are equal to the microcanonical averages.

⁴ S is the energy surface.

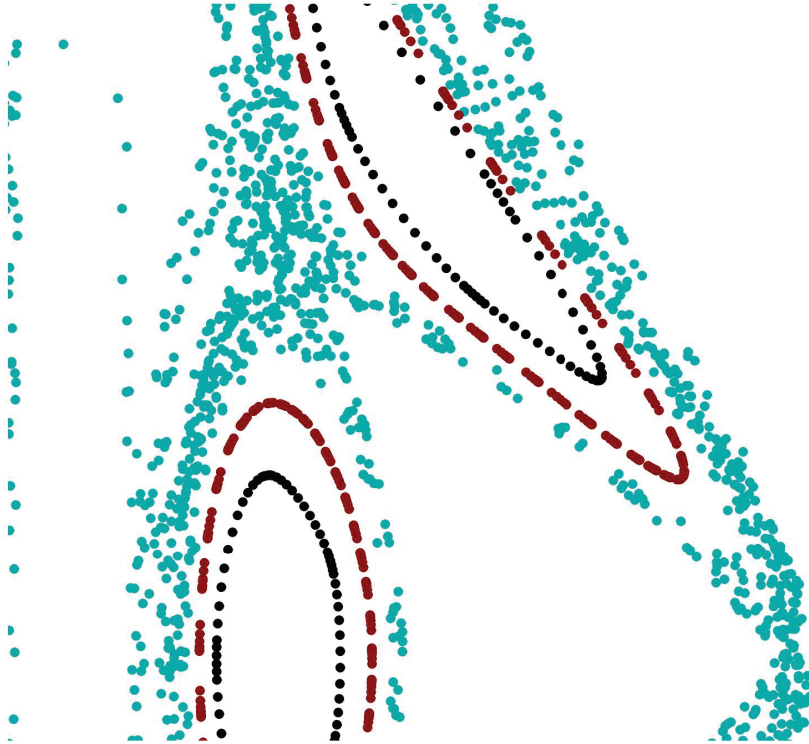


Fig. 1: KAM tori and nonergodic motion. This is a (Poincaré) cross-section of Earth's motion in the three-body problem, with Jupiter's mass set at almost 70 times its actual value. The closed loops correspond to trajectories that form tori in phase space, whose cross-sections look like deformed circles in our view. The complex filled region is a single trajectory exhibiting chaotic motion, and represents an ergodic component. The tori, each an ergodic component, can together be shown to occupy nonzero volume in phase space, for small Jovian masses. Note that this system is not ergodic according to either of our definitions. The trajectories on the tori never explore the rest of the energy surface. The region R formed by the chaotic domain is invariant under the time evolution; it has positive volume and the region outside R also has positive volume.

Definition 2

A time evolution in a set S is ergodic if and only if all the ergodic components R in S either have zero volume or have a volume equal to the volume of S .

A trajectory $r(t)$ must lie within a single ergodic component. If $r(t)$ covers the energy surface densely, then there is ‘no more room’ for a second ergodic component with non-zero volume. Conversely, if there is only one ergodic component R with volume equal to S , then any trajectory starting in R must get arbitrarily close to all points in R , otherwise the points in R ‘far’ from the trajectory (outside the closure of the trajectory) would be an invariant set of nonzero volume.

The time averages must equal microcanonical averages. Denote the microcanonical average of an observable O as $\langle O \rangle_S$, and the time average starting at initial condition (\mathbb{P}, \mathbb{Q}) as $\overline{O(\mathbb{P}(0), \mathbb{Q}(0))} = \lim_{T \rightarrow \infty} (1/T) \int_0^T O(\mathbb{P}(t), \mathbb{Q}(t)) dt$.

The time average \overline{O} equals the ensemble average $\langle O \rangle_S$ for an ergodic system.

Time averages are constant on trajectories. If O is a nice function (e.g. without any infinities on the energy surface), then

$$\overline{O(\mathbb{P}(0), \mathbb{Q}(0))} = \overline{O(\mathbb{P}(t), \mathbb{Q}(t))} , \quad (64)$$

the future time average does not depend on the values of O during the finite time interval $(0, t)$. Thus the time average \overline{O} is constant along the trajectory.⁵

⁵If we could show that \overline{O} had to be a continuous function, we would now be able to use the first definition of ergodicity to show that it was constant on the energy surface, since our trajectory comes close to every point on the surface. But it is not obvious that \overline{O} is continuous; for example, it is not continuous for Hamiltonian systems that are not ergodic. Consider two initial conditions at nearby points, one just inside a chaotic region and the other on a KAM torus. The infinite time averages on these two

Time averages are constant on the energy surface. Now consider the subset R_a of the energy surface where $\bar{O} < a$, for some value a . Since \bar{O} is constant along a trajectory, any point in R_a is sent under the time evolution to another point in R_a , so R_a is an ergodic component. If we have ergodic dynamics on the energy surface, that means the set R_a has either zero volume or the volume of the energy surface. This implies that \bar{O} is a constant on the energy surface (except on a set of zero volume); its value is a^* , the lowest value where R_{a^*} has the whole volume. Thus the equilibrium, time average value of our observable O is independent of initial condition.

Time averages equal microcanonical averages. Is this equilibrium value given by the microcanonical ensemble average over S ? We need to show that the trajectories do not linger in some regions of the energy surface more than they should (based on the thickness of the energy shell). Liouville's theorem told us that the microcanonical ensemble was time independent, so the ensemble average equals its time average, which equals the ensemble average of the time average. But the time average is constant (except on a set of zero volume), so in an ergodic system the ensemble average equals the time average trajectories for most quantities will be quite different; \bar{O} will typically have a jump at the boundary.

everywhere (except on a set of zero volume).⁶

Ergodicity has been proven for the collisions of hard spheres, and for geodesic motion on finite surfaces with constant negative curvature⁷, but not for many systems of immediate practical importance. Indeed, several fundamental problems precisely involve systems which are not ergodic.

References

- [1] R.K. Pathria and P.D. Beale. *Statistical Mechanics*. Elsevier Science, 1996.
- [2] H. Bradt. *Astrophysics Processes*. September 2008.
- [3] M. Kardar. *Statistical Physics of Particles*. Cambridge University Press, 2007.
- [4] K.S. Thorne and R.D. Blandford. *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics*. Princeton University Press, 2017.

⁶ $\langle O \rangle_S = \langle O(t) \rangle_S = \langle O(\mathbb{P}(t), \mathbb{Q}(t)) \rangle_S$, where the average $\langle \cdot \rangle_S$ integrates over initial conditions $(\mathbb{P}(0), \mathbb{Q}(0))$ but evaluates O at $(\mathbb{P}(t), \mathbb{Q}(t))$. Averaging over all time, and using the fact that the time average $\overline{O} = a^*$ (almost everywhere),

$$\begin{aligned} \langle O \rangle_S &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle O(\mathbb{P}(t), \mathbb{Q}(t)) \rangle_S dt \\ &= \left\langle \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T O(\mathbb{P}(t), \mathbb{Q}(t)) dt \right\rangle_S \\ &= \langle \overline{O(\mathbb{P}, \mathbb{Q})} \rangle_S = \langle a^* \rangle_S = a^* . \end{aligned}$$

⁷Geodesic motion on a sphere would be motion at a constant speed around great circles. Geodesics are the shortest paths between two points. In general relativity, falling bodies travel on geodesics in space-time.

- [5] J. Sethna. *Statistical Mechanics: Entropy, Order Parameters and Complexity*. Oxford Master Series in Physics. OUP Oxford, 2006.
- [6] M. Harwit. *Astrophysical Concepts*. Astronomy and Astrophysics Library. Springer New York, 2006.