

## 3. The Microscopic Connection

### 3.1 Classical Mechanics

In nonequilibrium statistical mechanics we seek to model transport processes beginning with an understanding of the motion and interactions of individual atoms or molecules. The laws of classical mechanics govern the motion of atoms and molecules so in this chapter we begin with a brief description of the mechanics of Newton, Lagrange and Hamilton. It is often useful to be able to treat *constrained* mechanical systems. We will use a Principle due to Gauss to treat many different types of constraint - from simple bond length constraints, to constraints on kinetic energy. As we shall see, kinetic energy constraints are useful for constructing various constant temperature ensembles. We will then discuss the Liouville equation and its formal solution. This equation is the central vehicle of nonequilibrium statistical mechanics. We will then need to establish the link between the microscopic dynamics of individual atoms and molecules and the macroscopic hydrodynamical description discussed in the last chapter. We will discuss two procedures for making this connection. The Irving and Kirkwood procedure relates hydrodynamic variables to nonequilibrium *ensemble averages* of microscopic quantities. A more direct procedure we will describe, succeeds in deriving *instantaneous* expressions for the hydrodynamic field variables.

#### Newtonian Mechanics

Classical mechanics (Goldstein, 1980) is based on Newton's three laws of motion. This theory introduced the concepts of a force and an acceleration. Prior to Newton's work, the connection had been made between forces and velocities. Newton's laws of motion were supplemented by the notion of a force acting at a distance. With the identification of the force of gravity and an appropriate initial condition - initial coordinates and velocities - trajectories could be computed. Philosophers of science have debated the content of Newton's laws but when augmented with a force which is expressible as a function of time, position or possibly of velocity, those laws lead to the equation,

$$m\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t) \quad (3.1)$$

which is well-posed and possesses a unique solution.

#### Lagrange's equations

After Newton, scientists discovered different sets of equivalent laws or axioms upon which classical mechanics could be based. More *elegant* formulations are due to Lagrange and Hamilton. Newton's laws are less general than they might seem. For instance the position  $\mathbf{r}$ , that appears in Newton's equation must be a

Cartesian vector in a Euclidean space. One does not have the freedom of say, using angles as measures of position. Lagrange solved the problem of formulating the laws of mechanics in a form which is valid for *generalised* coordinates.

Let us consider a system with generalised coordinates  $q$ . These coordinates may be Cartesian positions, angles or any other convenient parameters that can be found to uniquely specify the configuration of the system. The kinetic energy  $T$ , will in general be a function of the coordinates and their time derivatives  $\dot{q}$ . If  $V(q)$  is the potential energy, we define the Lagrangian to be  $L(q, \dot{q}, t) \equiv T(q, \dot{q}, t) - V(q)$ . The fundamental dynamical postulate states that the motion of a system is such that the *action*,  $S$ , is an extremum

$$\delta S = \delta \int_{t_0}^{t_1} L(q, \dot{q}, t) dt = 0 \quad (3.2)$$

Let  $q(t)$  be the coordinate trajectory that satisfies this condition and let  $q(t) + \delta q(t)$  where  $\delta q(t)$  is an arbitrary variation in  $q(t)$ , be an arbitrary trajectory. The varied motion must be consistent with the initial and final positions. So that,  $\delta q(t_1) = \delta q(t_0) = 0$ . We consider the change in the action due to this variation.

$$\delta S = \int_{t_0}^{t_1} L(q + \delta q, \dot{q} + \delta \dot{q}, t) dt - \int_{t_0}^{t_1} L(q, \dot{q}, t) dt = \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right) dt \quad (3.3)$$

Integrating the second term by parts gives

$$\delta S = \left[ \frac{\partial L}{\partial \dot{q}} \delta q \right]_{t_0}^{t_1} + \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial q} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) \right) \delta q dt \quad (3.4)$$

The first term vanishes because  $\delta q$  is zero at both limits. Since for  $t_0 < t < t_1$ ,  $\delta q(t)$  is arbitrary, the only way that the variation in the action  $\delta S$ , can vanish is if the equation,

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) = 0 \quad (3.5)$$

holds for all time. This is Lagrange's equation of motion. If the coordinates are taken to be Cartesian, it is easy to see that Lagrange's equation reduces to Newton's.

## Hamiltonian mechanics

Although Lagrange's equation has removed the special status attached to Cartesian coordinates, it has introduced a new difficulty. The Lagrangian is a function of generalised coordinates, their time derivatives and possibly of time. The equation is not symmetric with respect to the interchange of coordinates and velocities. Hamilton derived an equivalent set of equations in which the roles played by

coordinates and *velocities* can be interchanged. Hamilton defined the *canonical momentum*  $p$ ,

$$p \equiv \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}} \quad (3.6)$$

and introduced the function

$$H(q, p, t) \equiv \dot{q} \frac{\partial L}{\partial \dot{q}} - L = \dot{q}p - L \quad (3.7)$$

This function is of course now known as the Hamiltonian. Consider a change in the Hamiltonian which can be written as

$$dH = \dot{q}dp + p d\dot{q} - dL \quad (3.8)$$

The Lagrangian is a function of  $q$ ,  $\dot{q}$  and  $t$  so that the change  $dL$ , can be written as

$$dL = \frac{\partial L}{\partial q} dq + \frac{\partial L}{\partial \dot{q}} d\dot{q} + \frac{\partial L}{\partial t} dt \quad (3.9)$$

Using the definition of the canonical momentum  $p$ , and substituting for  $dL$ , the expression for  $dH$  becomes

$$dH = \dot{q}dp - \frac{\partial L}{\partial q} dq - \frac{\partial L}{\partial t} dt \quad (3.10)$$

Lagrange's equation of motion (3.5), rewritten in terms of the canonical momenta is

$$\dot{p} = \frac{\partial L}{\partial q} \quad (3.11)$$

so that the change in  $H$  is

$$dH = \dot{q}dp - \dot{p}dq - \frac{\partial L}{\partial t} dt \quad (3.12)$$

Since the Hamiltonian is a function of  $q, p$  and  $t$ , it is easy to see that Hamilton equations of motion are

$$\dot{q} = \frac{\partial H}{\partial p} \quad \text{and} \quad \dot{p} = -\frac{\partial H}{\partial q} \quad (3.13)$$

As mentioned above these equations are symmetric with respect to coordinates and momenta. Each has equal status in Hamilton's equations of motion. If  $H$  has no explicit time dependence, its value is a constant of the motion. Other formulations of classical mechanics such as the Hamilton-Jacobi equations will not concern us in this book.

## Gauss' Principle of Least Constraint

Apart from relativistic or quantum corrections, classical mechanics is thought to give an exact description of motion. In this section our point of view will change somewhat. Newtonian or Hamiltonian mechanics imply a certain set of constants of the motion: energy, and linear and angular momentum. In thermodynamically interesting systems the natural fixed quantities are the thermodynamic state variables; the number of molecules  $N$ , the volume  $V$  and the temperature  $T$ . Often the pressure rather than the volume may be preferred. Thermodynamically interesting systems usually exchange energy, momentum and mass with their surroundings. This means that within thermodynamic systems none of the classical constants of the motion are actually constant.

Typical thermodynamic systems are characterised by fixed values of thermodynamic variables: temperature, pressure, chemical potential, density, enthalpy or internal energy. The system is maintained at a fixed thermodynamic state (say temperature) by placing it in contact with a reservoir, with which it exchanges energy (heat) in such a manner as to keep the temperature of the system of interest fixed. The heat capacity of the reservoir must be much larger than that of the system, so that the heat exchanged from the reservoir does not affect the reservoir temperature.

Classical mechanics is an awkward vehicle for describing this type of system. The only way that thermodynamic systems can be treated in Newtonian or Hamiltonian mechanics is by explicitly modelling the system, the reservoir and the exchange processes. This is complex, tedious and as we will see below, it is also unnecessary. We will now describe a little known principle of classical mechanics which is extremely useful for *designing* equations of motion which are more useful from a macroscopic or thermodynamic viewpoint. This principle does indeed allow us to modify classical mechanics so that thermodynamic variables may be made constants of the motion.

Just over 150 years ago Gauss formulated a mechanics more general than Newton's. This mechanics has as its foundation Gauss' principle of least constraint. Gauss (1829) referred to this as *the most fundamental dynamical principle* (Whittacker 1937, Pars 1965). Suppose that the **cartesian** coordinates and velocities of a system are given at time  $t$ . Consider the function  $C$ , referred to by Hertz as the square of the curvature, where

$$C = \frac{1}{2} \sum_{i=1}^N m_i \left( \ddot{\mathbf{r}}_i - \frac{\mathbf{F}_i}{m_i} \right)^2 \quad (3.14)$$

$C$  is a function of the set of accelerations  $\{\ddot{\mathbf{r}}\}$ . Gauss' principle states that the actual physical acceleration corresponds to the minimum value of  $C$ . Clearly if the system is not subject to a constraint then  $C = 0$  and the system evolves under



Newton's equations of motion. For a constrained system it is convenient to change variables from  $\mathbf{r}_i$  to  $\mathbf{w}_i$  where

$$\begin{aligned}\mathbf{w}_i &= \sqrt{m_i} \mathbf{r}_i \\ \zeta_i &= \mathbf{F}_i / \sqrt{m_i}\end{aligned}\tag{3.15}$$

Because the  $\{\mathbf{w}_i\}$ , are related to the Jacobi metric, we will refer to this coordinate system as the Jacobi frame.

The types of constraints which might be applied to a system fall naturally into two types, **holonomic** and **nonholonomic**. A holonomic constraint is one which can be integrated out of the equations of motion. For instance, if a certain generalised coordinate is fixed, its conjugate momentum is zero for all time, so we can simply consider the problem in the reduced set of unconstrained variables. We need not be conscious of the fact that a force of constraint is acting upon the system to fix the coordinate and the momentum. An analysis of the two dimensional motion of an ice skater need not refer to the fact that the gravitational force is exactly resisted by the stress on the ice surface fixing the vertical coordinate and velocity of the ice skater. We can ignore these degrees of freedom.

Nonholonomic constraints usually involve velocities. These constraints are not integrable. In general a nonholonomic constraint will do work on a system. Thermodynamic constraints are invariably nonholonomic. It is known that the Action Principle cannot be used to describe motion under nonholonomic constraints (Evans and Morriss, 1984).

We can write a general constraint in the Jacobi frame in the form

$$g(\mathbf{w}, \dot{\mathbf{w}}, t) = 0\tag{3.16}$$

where  $g$  is a function of Jacobi positions, velocities and possibly time. Either type of constraint function, holonomic or nonholonomic, can be written in this form. If this equation is differentiated with respect to time, once for nonholonomic constraints and twice for holonomic constraints we see that,

$$\mathbf{n}(\mathbf{w}, \dot{\mathbf{w}}, t) \cdot \ddot{\mathbf{w}} = s(\mathbf{w}, \dot{\mathbf{w}}, t)\tag{3.17}$$

We refer to this equation as the *differential constraint equation* and it plays a fundamental role in Gauss' Principle of Least Constraint. It is the equation for a plane which we refer to as the *constraint plane*.  $\mathbf{n}$  is the vector normal to the constraint plane.

Our problem is to solve Newton's equation subject to the constraint. Newton's equation gives us the acceleration in terms of the unconstrained forces. The differential constraint equation places a condition on the acceleration vector for the system. The differential constraint equation says that the constrained

acceleration vector must terminate on a hyper-plane in the  $3N$ -dimensional Jacobi acceleration space (equation 3.17).

Imagine for the moment that at some initial time the system satisfies the constraint equation  $g = 0$ . In the absence of the constraint the system would evolve according to Newton's equations of motion where the acceleration is given by

$$\ddot{\mathbf{w}}_i^u = \zeta_i \quad (3.18)$$

This trajectory would in general not satisfy the constraint. Further, the constraint function  $g$  tells us that the only accelerations which do continuously satisfy the constraint, are those which terminate on the constraint plane. To obtain the constrained acceleration we must project the unconstrained acceleration back into the constraint plane.

Gauss' principle of least constraint gives us a prescription for constructing this projection. *Gauss' principle states that the trajectories actually followed are those which deviate as little as possible, in a least squares sense, from the unconstrained Newtonian trajectories.* The projection which the system actually follows is the one which minimises the magnitude of the Jacobi frame constraint force. This means that the force of constraint must be parallel to the normal of the constraint surface. The Gaussian equations of motion are then

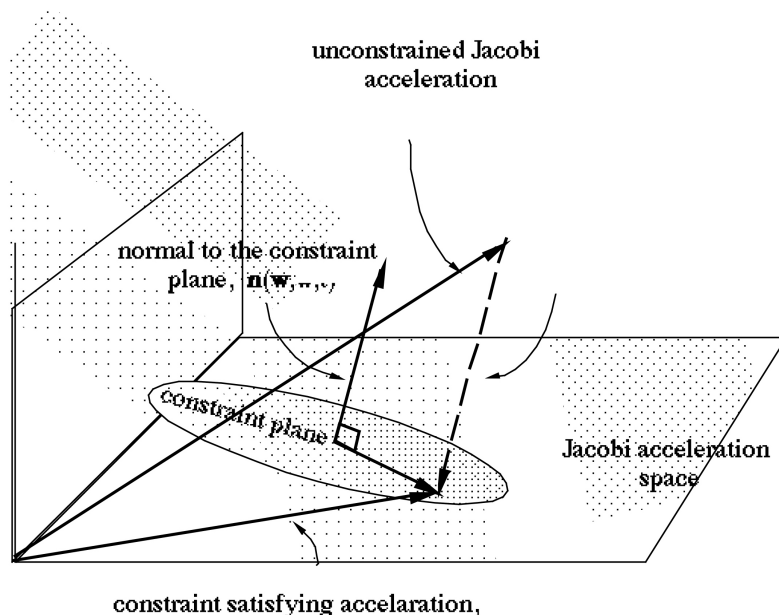
$$\ddot{\mathbf{w}}_i = \zeta_i - \lambda \mathbf{n} \quad (3.19)$$

where  $\lambda$  is a Gaussian multiplier which is a function of position, velocity and time.

To calculate the multiplier we use the differential form of the constraint function. Substituting for the acceleration we obtain

$$\lambda = \frac{\mathbf{n} \cdot \dot{\zeta} - s}{\mathbf{n} \cdot \mathbf{n}} \quad (3.20)$$

It is worthwhile at this stage to make a few comments about the procedure outlined above. First, notice that the original constraint equation is never used explicitly. Gauss' principle only refers to the **differential** form of the constraint equation. This means that the precise *value* of the constrained quantity is undetermined. The constraint acts only to stop its value changing. In the holonomic case Gauss' principle and the principle of least action are of course completely equivalent. In the nonholonomic case the equations resulting from the application of Gauss' Principle cannot be derived from a Hamiltonian and the principle of least action **cannot** be used to derive constraint satisfying equations. In the nonholonomic case, Gauss' principle does **not** yield equations of motion for which the work done by the constraint forces is a minimum.



**Figure 3.1 Gauss' Principle of Least Constraint**

The derivation of constrained equations of motion given above is geometric, and is done in the transformed coordinates which we have termed the Jacobi frame. It is not always convenient to write a constraint function in the Jacobi frame, and from an operational point of view a much simpler derivation of constrained equations of motion is possible using Lagrange multipliers. The square of the curvature  $C$  is a function of accelerations only (the Cartesian coordinates and velocities are considered to be given parameters). Gauss' principle reduces to finding the minimum of  $C$ , subject to the constraint. The constraint function must also be written as a function of accelerations, but this is easily achieved by differentiating with respect to time. If  $G$  is the acceleration dependent form of the constraint, then the constrained equations of motion are obtained from

$$\frac{\partial}{\partial \ddot{\mathbf{r}}}(C - \lambda G) = 0 \quad (3.21)$$

It is easy to see that the Lagrange multiplier  $\lambda$ , is (apart from the sign) equal to the Gaussian multiplier. We will illustrate Gauss' principle by considering some useful examples.

### Gauss' Principle for Holonomic Constraints

The most common type of holonomic constraint in statistical mechanics is probably that of fixing bond lengths and bond angles in molecular systems. The

vibrational degrees of freedom typically have a relaxation timescale which is orders of magnitude faster than translational degrees of freedom, and are therefore often irrelevant to the processes under study. As an example of the application of Gauss' principle of least constraint for holonomic constraints we consider a diatomic molecule with a fixed bond length. The generalisation of this method to more than one bond length is straightforward (see Edberg, Evans and Morriss, 1986) and the application to bond angles is trivial since they can be formulated as second nearest neighbour distance constraints. The constraint function for a diatomic molecule is that the distance between sites one and two be equal to  $d_{12}$ , that is

$$g(\mathbf{r}, \dot{\mathbf{r}}, t) = \mathbf{r}_{12}^2 - d_{12}^2 = 0, \quad (3.22)$$

where we define  $\mathbf{r}_{12}$  to be the vector from  $\mathbf{r}_1$  to  $\mathbf{r}_2$ , ( $\mathbf{r}_{12} \equiv \mathbf{r}_2 - \mathbf{r}_1$ ). Differentiating twice with respect to time gives the acceleration dependent constraint equation,

$$\mathbf{r}_{12} \cdot \ddot{\mathbf{r}}_{12} + (\dot{\mathbf{r}}_{12})^2 = 0 \quad (3.23)$$

To obtain the constrained equations of motion we minimise the function  $C$  subject to the constraint equation (3.23). That is

$$\frac{\partial}{\partial \ddot{\mathbf{r}}_i} \left\{ \frac{m_1}{2} \left( \ddot{\mathbf{r}}_1 - \frac{\mathbf{F}_1}{m_1} \right)^2 + \frac{m_2}{2} \left( \ddot{\mathbf{r}}_2 - \frac{\mathbf{F}_2}{m_2} \right)^2 - \lambda \left( \mathbf{r}_{12} \cdot \ddot{\mathbf{r}}_{12} + (\dot{\mathbf{r}}_{12})^2 \right) \right\} = 0 \quad (3.24)$$

For  $i$  equal to 1 and 2 this gives

$$\begin{aligned} m_1 \ddot{\mathbf{r}}_1 &= \mathbf{F}_1 - \lambda \mathbf{r}_{12} \\ m_2 \ddot{\mathbf{r}}_2 &= \mathbf{F}_2 + \lambda \mathbf{r}_{12} \end{aligned} \quad (3.25)$$

Notice that the extra terms in these equations of motion have opposite signs. This is because the coefficients of the  $\mathbf{r}_1$  and  $\mathbf{r}_2$  accelerations have opposite signs. The total constraint force on the molecule is zero so there is no change in the total momentum of the molecule. To obtain an expression for the multiplier  $\lambda$  we combine these two equations to give an equation of motion for the bond vector  $\mathbf{r}_{12}$ ,

$$\ddot{\mathbf{r}}_{12} = \left( \frac{\mathbf{F}_2}{m_2} - \frac{\mathbf{F}_1}{m_1} \right) + \lambda \left( \frac{1}{m_2} + \frac{1}{m_1} \right) \mathbf{r}_{12} \quad (3.26)$$

Substituting this into the differential form of the constraint function (3.23), gives

$$\lambda = - \frac{\mathbf{r}_{12} \cdot (m_1 \mathbf{F}_2 - m_2 \mathbf{F}_1) + m_1 m_2 \dot{\mathbf{r}}_{12}^2}{(m_1 + m_2) \mathbf{r}_{12}^2} \quad (3.27)$$

It is very easy to implement these constrained equations of motion as the multiplier is a simple explicit function of the positions, velocities and Newtonian forces. For more complicated systems with multiple bond length and bond angle

constraints (all written as distance constraints) we obtain a set of coupled linear equations to solve for the multipliers.

## Gauss' Principle for Nonholonomic Constraints

One of the simplest and most useful applications of Gauss' Principle is to derive equations of motion for which the ideal gas temperature (ie. the kinetic energy) is a constant of the motion (Evans et. al. 1983). Here the constraint function is

$$g(\mathbf{r}, \dot{\mathbf{r}}, t) = \sum_{i=1}^N \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - \frac{3}{2} N k_B T = 0 \quad (3.28)$$

Differentiating once with respect to time gives the equation for the constraint plane

$$\sum_{i=1}^N m_i \dot{\mathbf{r}}_i \cdot \ddot{\mathbf{r}}_i = 0 \quad (3.29)$$

Therefore to obtain the constrained Gaussian equations we minimise  $C$  subject to the constraint equation (3.29). That is

$$\frac{\partial}{\partial \ddot{\mathbf{r}}_i} \left( \frac{1}{2} \sum_{j=1}^N m_j \left( \dot{\mathbf{r}}_j - \frac{\mathbf{F}_j}{m_j} \right)^2 + \lambda \sum_{j=1}^N m_j \dot{\mathbf{r}}_j \cdot \ddot{\mathbf{r}}_j \right) = 0 \quad (3.30)$$

This gives

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i - \lambda m_i \dot{\mathbf{r}}_i \quad (3.31)$$

Substituting the equations of motion into the differential form of the constraint equation, we find that the multiplier is given by

$$\lambda = \frac{\sum_{i=1}^N \mathbf{F}_i \cdot \dot{\mathbf{r}}_i}{\sum_{i=1}^N m_i \dot{\mathbf{r}}_i^2} \quad (3.32)$$

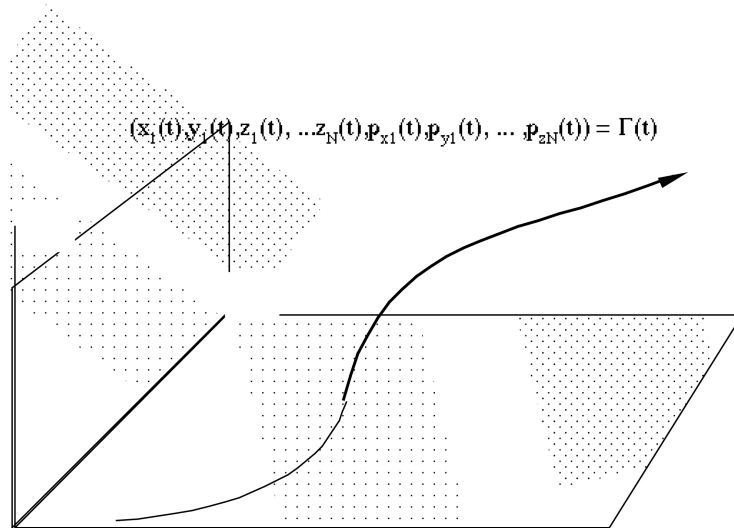
As before,  $\lambda$  is a simple function of the forces and velocities so that the implementation of the constant kinetic energy constraint in a molecular dynamics computer program only requires a trivial modification of the equations of motion in a standard program. Equations (3.31 & 32) constitute what have become known as the *Gaussian isokinetic equations of motion*. These equations were first proposed simultaneously and independently by Hoover et. al. (1982) and Evans (1983). In these original papers Gauss' principle was however not referred to. It was a year before the connection with Gauss' principle was made.

With regard to the general application of Gauss' principle of least constraint one should always examine the statistical mechanical properties of the resulting dynamics. If one applies Gauss' principle to the problem of maintaining a constant heat flow, then a comparison with linear response theory shows that the Gaussian

equations of motion **cannot** be used to calculate thermal conductivity (Hoover 1986). The correct application of Gauss' principle is limited to arbitrary holonomic constraints and apparently, to nonholonomic constraint functions which are **homogeneous** functions of the momenta.

### 3.2 Phase Space

To give a complete description of the state of a 3-dimensional  $N$ -particle system at any given time it is necessary to specify the  $3N$  coordinates and  $3N$  momenta. The  $6N$  dimensional space of coordinates and momenta is called *phase space* (or  $\Gamma$ -space). As time progresses the phase point  $\Gamma$ , traces out a path which we call the phase space trajectory of the system. As the equations of motion for  $\Gamma$  are  $6N$  first order differential equations, there are  $6N$  constants of integration (they may be for example the  $6N$  initial conditions  $\Gamma(0)$ ). Rewriting the equations of motion in terms of these constants shows that the trajectory of  $\Gamma$  is completely determined by specifying these  $6N$  constants. An alternate description of the time evolution of the system is given by the trajectory in the extended  $\Gamma'$ -space, where  $\Gamma' = (\Gamma, t)$ . As the  $6N$  initial conditions uniquely determine the trajectory, two points in phase space with different initial conditions form distinct non-intersecting trajectories in  $\Gamma'$ -space.



**Figure 3.2 Phase Space Trajectory 6N-dimensional  $\Gamma$ -space. As time evolves the system traces out a trajectory in 6N-dimensional  $\Gamma$ -space.**

To illustrate the ideas of  $\Gamma$ -space and  $\Gamma'$ -space it is useful to consider one of the simplest mechanical systems, the harmonic oscillator. The Hamiltonian for the

harmonic oscillator is  $H = \frac{1}{2}(kx^2 + p^2/m)$  where  $m$  is the mass of the oscillator and  $k$  is the spring constant. The equations of motion are

$$\begin{aligned}\dot{x} &= \frac{\partial H}{\partial p} = \frac{p}{m} \\ \dot{p} &= -\frac{\partial H}{\partial x} = -kx\end{aligned}\tag{3.33}$$

and the energy (or the Hamiltonian) is a constant of the motion. The  $\Gamma$ -space for this system is 2-dimensional  $(x, p)$  and the  $\Gamma$ -space trajectory is given by

$$(x(t), p(t)) = \left( x_0 \cos \omega t + \frac{p_0}{m\omega} \sin \omega t, p_0 \cos \omega t - m\omega x_0 \sin \omega t \right)\tag{3.34}$$

The constants  $x_0$  and  $p_0$  are the 2 integration constants written in this case as an initial condition. The frequency  $\omega$  is related to the spring constant and mass by  $\omega^2 = k/m$ . The  $\Gamma$ -space trajectory is an ellipse,

$$m^2 \omega^2 x(t)^2 + p(t)^2 = m^2 \omega^2 x_0^2 + p_0^2\tag{3.35}$$

which intercepts the  $x$ -axis at  $\pm(x_0^2 + p_0^2/m^2\omega^2)^{1/2}$  and the  $p$ -axis at  $\pm(p_0^2 + m^2\omega^2 x_0^2)^{1/2}$ . The period of the motion is  $T = 2\pi/\omega = 2\pi(m/k)^{1/2}$ . This is the surface of constant energy for the harmonic oscillator. Any oscillator with the same energy must traverse the same  $\Gamma$ -space trajectory, that is another oscillator with the same energy, but different initial starting points  $(x_0, p_0)$  will follow the same ellipse but with a different initial phase angle.

The trajectory in  $\Gamma'$ -space is a elliptical coil, and the constant energy surface in  $\Gamma'$ -space is a elliptical cylinder, and oscillators with the same energy start from different points on the ellipse at time zero (corresponding to different initial phase angles), and wind around the elliptical cylinder. The trajectories in  $\Gamma'$ -space are non-intersecting. If two trajectories in  $\Gamma'$ -space meet at time  $t$ , then the two trajectories must have had the same initial condition. As the choice of time origin is arbitrary, the trajectories must be the same for all time.

In  $\Gamma$ -space the situation is somewhat different. The trajectory for the harmonic oscillator winds around the ellipse, returning to its initial phase point  $(x_0, p_0)$  after a time  $T$ . The period of time taken for a system to return to (or to within an  $\varepsilon$ -neighbourhood of) its initial starting phase is called the Poincaré recurrence time. For a simple system such as the harmonic oscillator the recurrence time is trivial to calculate, but for higher dimensional systems the recurrence time quickly exceeds the estimated age of the universe.

### 3.3 Distribution Functions and the Liouville Equation

In the first few sections of this chapter we have given a description of the mechanics of individual  $N$ -particle systems. The development which follows

describes an *ensemble* of such systems; that is an essentially infinite number of systems characterised by identical dynamics and identical state variables ( $N, V, E$  or  $T$  etc.) but different initial conditions, ( $\Gamma(0)$ ). We wish to consider the average behaviour of a collection of macroscopically identical systems distributed over a range of initial states (microstates). In generating the ensemble we make the usual assumptions of classical mechanics. We assume that it is possible to know all the positions and momenta of an  $N$  particle system to arbitrary precision at some initial time, and that the motion can be calculated exactly from the equations of motion.

The ensemble contains an infinite number of individual systems so that the number of systems in a particular state may be considered to change continuously as we pass to neighbouring states. This assumption allows us to define a density function  $f(\Gamma, t)$ , which assigns a probability to points in phase space. Implicit in this assumption is the requirement that  $f(\Gamma, t)$ , has continuous partial derivatives with respect to all its variables, otherwise the phase density will not change *continuously* as we move to neighbouring states. If the system is Hamiltonian and all trajectories are confined to the energy surface then  $f(\Gamma, t)$  will not have a continuous partial derivatives with respect to energy. Problems associated with this particular source of discontinuity can obviously be avoided by eliminating the energy as a variable, and considering  $f(\Gamma, t)$  to be a density function defined on a surface of constant energy (effectively reducing the dimensionality of the system). However it is worth pointing out that other sources of discontinuity in the phase space density, may not be so easily removed.

To define a distribution function for a particular system we consider an ensemble of identical systems whose initial conditions *span* the phase space specified by the macroscopic constraints. We consider an infinitesimal element of phase space located at  $\Gamma \equiv (\mathbf{q}, \mathbf{p})$ . The fraction of systems  $\delta N$ , which at time  $t$  have coordinates and momenta within  $\delta \mathbf{q}$ ,  $\delta \mathbf{p}$  of  $\mathbf{q}$ ,  $\mathbf{p}$  is used to define the phase space distribution function  $f(\mathbf{q}, \mathbf{p}, t)$ , by

$$\delta N = f(\mathbf{q}, \mathbf{p}, t) \delta \mathbf{q} \delta \mathbf{p} \quad (3.36)$$

The total number of systems in the ensemble is fixed, so integrating over the whole phase space we can normalise the distribution function,

$$1 = \int f(\mathbf{q}, \mathbf{p}, t) \delta \mathbf{q} \delta \mathbf{p} \quad (3.37)$$

If we consider a small volume element of phase space, the number of trajectories entering the rectangular volume element  $\delta \mathbf{q} \delta \mathbf{p}$  through some face will in general be different from the number which leave through an opposite face. For the faces normal to the  $q_1$ -axis, located at  $q_1$ , and  $q_1 + \delta q_1$ , the fraction of ensemble members entering the first face is



$$f(q_1, \dots, t) \dot{q}_1(q_1, \dots, t) \delta q_2 \dots \delta q_{3N} \delta \mathbf{p}$$

Similarly the fraction of points leaving through the second face is

$$\begin{aligned} & f(q_1 + \delta q_1, \dots, t) \dot{q}_1(q_1 + \delta q_1, \dots, t) \delta q_2 \dots \delta q_{3N} \delta \mathbf{p} \\ & \approx \left( f(q_1, \dots, t) + \frac{\partial f}{\partial q_1} \delta q_1 \right) \left( \dot{q}_1(q_1 + \delta q_1, \dots, t) + \frac{\partial \dot{q}_1}{\partial q_1} \delta q_1 \right) \delta q_2 \dots \delta q_{3N} \delta \mathbf{p} \end{aligned}$$

Combining these expressions gives the change in  $\delta N$  due to fluxes in the  $q_1$  direction

$$\frac{d}{dt} \delta N_{q_1} = - \left( \dot{q}_1 \frac{\partial f}{\partial q_1} + f \frac{\partial \dot{q}_1}{\partial q_1} \right) \delta \mathbf{q} \delta \mathbf{p} \quad (3.38)$$

Summing over all coordinate (and momentum) directions gives the total fractional change  $\delta N$  as

$$\frac{d}{dt} \delta N = - \sum_{i=1}^N \left[ f \left( \frac{\partial}{\partial \mathbf{q}_i} \cdot \dot{\mathbf{q}}_i + \frac{\partial}{\partial \mathbf{p}_i} \cdot \dot{\mathbf{p}}_i \right) + \dot{\mathbf{q}}_i \cdot \frac{\partial f}{\partial \mathbf{q}_i} + \dot{\mathbf{p}}_i \cdot \frac{\partial f}{\partial \mathbf{p}_i} \right] \delta \mathbf{q} \delta \mathbf{p} \quad (3.39)$$

Dividing through by the phase space volume element  $\delta \mathbf{q} \delta \mathbf{p}$  we obtain the rate of change in density  $f(\mathbf{q}, \mathbf{p})$ , at the point  $(\mathbf{q}, \mathbf{p})$ ,

$$\frac{1}{\delta \mathbf{q} \delta \mathbf{p}} \frac{d}{dt} \delta N = \frac{\partial}{\partial t} \left( \frac{\delta N}{\delta \mathbf{q} \delta \mathbf{p}} \right) = \frac{\partial f}{\partial t} \Big|_{\mathbf{q}, \mathbf{p}} \quad (3.40)$$

Using the notation,  $\Gamma = (\mathbf{q}, \mathbf{p}) = (q_1, q_2, \dots, q_{3N}, p_1, p_2, \dots, p_{3N})$  for the  $6N$ -dimensional phase point, this may be written as

$$\frac{\partial f}{\partial t} \Big|_{\Gamma} = -f \frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} - \dot{\Gamma} \cdot \frac{\partial f}{\partial \Gamma} = -\frac{\partial}{\partial \Gamma} \cdot (f \dot{\Gamma}) \quad (3.41)$$

This is the Liouville equation for the phase space distribution function. Using the streaming or total time derivative of the distribution function, we can rewrite the Liouville equation in an equivalent form as,

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \dot{\Gamma} \cdot \frac{\partial f}{\partial \Gamma} = -f \frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} = -f \Lambda(\Gamma) \quad (3.42)$$

This equation has been obtained without reference to the equations of motion. Its correctness does not require the existence of a Hamiltonian to generate the equations of motion. The equation rests on two conditions: that ensemble members cannot be created or destroyed and that the distribution function is sufficiently smooth that the appropriate derivatives exist.  $\Lambda(\Gamma)$  is called the *phase space compression factor* since it is equal to the negative time derivative of the logarithm of the phase space distribution function.

$$\frac{d}{dt} \ln[f(\Gamma, t)] = -\Lambda(\Gamma) \quad (3.43)$$

The Liouville equation is usually written in a slightly simpler form. If the equations of motion can be generated from a Hamiltonian, then it is a simple matter to show that  $\Lambda(\Gamma) = 0$ . This is so even in the presence of external fields which may be driving the system away from equilibrium by performing work on the system.

$$\Lambda(\Gamma) = \sum_{i=1}^N \left( \frac{\partial}{\partial \mathbf{q}_i} \cdot \dot{\mathbf{q}}_i + \frac{\partial}{\partial \mathbf{p}_i} \cdot \dot{\mathbf{p}}_i \right) = \sum_{i=1}^N \left( \frac{\partial}{\partial \mathbf{q}_i} \cdot \frac{\partial H}{\partial \mathbf{p}_i} - \frac{\partial}{\partial \mathbf{p}_i} \cdot \frac{\partial H}{\partial \mathbf{q}_i} \right) = 0 \quad (3.44)$$

The existence of a Hamiltonian is a sufficient, but not necessary condition for the phase space compression factor to vanish. If phase space is incompressible then the Liouville equation takes on its simplest form,

$$\frac{df}{dt} = 0. \quad (3.45)$$

## Time Evolution of the distribution function

The following sections will be devoted to developing a formal operator algebra for manipulating the distribution function and averages of mechanical phase variables. This development is an extension of the treatment given by Berne (1977) which is applicable to Hamiltonian systems only. We will use the compact operator notation

$$\frac{\partial f}{\partial t} = -iL f = - \left( \frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} + \dot{\Gamma} \cdot \frac{\partial}{\partial \Gamma} \right) f \quad (3.46)$$

for the Liouville equation, equation (3.41). The operator  $iL$  is called the distribution function (or  $f$ -) Liouvillean. Both the distribution function  $f$ , and the  $f$ -Liouvillean are functions of the initial phase  $\Gamma$ . We assume that there is no explicit time dependence in the equations of motion (time varying external fields will be treated in Chapter 8). Using this notation we can write the formal solution of the Liouville equation for the time dependent  $N$ -particle distribution function  $f(\Gamma, t)$  as

$$f(\Gamma, t) = \exp[-iLt] f(\Gamma, 0) \quad (3.47)$$

where  $f(\Gamma, 0)$ , is the initial distribution function. This representation for the distribution function contains the exponential of an operator, which is a symbolic representation for the infinite series of operators. The  $f$ -propagator is defined as,

$$\exp[-iLt] = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} (iL)^n \quad (3.48)$$

The formal solution given above can therefore be written as

$$f(t) = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} (iL)^n f(0) = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} \frac{\partial^n}{\partial t^n} f(0) \quad (3.49)$$

This form makes it clear that the formal solution derived above is the Taylor series expansion of the *explicit* time dependence of  $f(\Gamma, t)$ , about  $f(\Gamma, 0)$ .

## Time Evolution of phase variables

We will need to consider the time evolution of functions of the phase of the system. Such functions are called phase variables. An example would be the phase variable for the internal energy of a system,  $H_0(\Gamma) = \sum_i p_i^2 / 2m_i + \Phi(q_1, \dots, q_n)$ . Phase variables by definition, do not depend on time explicitly, their time dependence comes solely from the time dependence of the phase  $\Gamma$ . Using the chain rule, the equation of motion for an arbitrary phase variable  $B(\Gamma)$  can be written as

$$\dot{B}(\Gamma) = \dot{\Gamma} \cdot \frac{\partial}{\partial \Gamma} B = \sum_{i=1}^N \left( \dot{q}_i \cdot \frac{\partial}{\partial q_i} + \dot{p}_i \cdot \frac{\partial}{\partial p_i} \right) B(\Gamma) \equiv iL(\Gamma) B(\Gamma) \quad (3.50)$$

The operator associated with the time derivative of a phase variable  $iL(\Gamma)$  is referred to as the phase variable (or  $P$ -) Liouvillean. The formal solution of this equation can be written in terms of the  $P$ -propagator,  $e^{iL t}$ . This gives the value of the phase variable as a function of time

$$B(t) = \exp[iL t] B(0) \quad (3.51)$$

This expression is very similar in form to that for the distribution function. It is the Taylor series expansion of the total time dependence of  $B(t)$ , expanded about  $B(0)$ . If the phase space compression factor  $\Lambda(\Gamma)$  is identically zero then the  $P$ -Liouvillean is equal to the  $f$ -Liouvillean, and the  $P$ -propagator is simply the adjoint or Hermitian conjugate of the  $f$ -propagator. In general this is not the case.

## Properties of Liouville Operators

In this section we will derive some of the more important properties of the Liouville operators. These will lead us naturally to a discussion of various *representations* of the properties of classical systems. The first property we shall discuss relates the  $P$ -Liouvillean to the  $f$ -Liouvillean as follows,

$$\int d\Gamma f(0) iL B(\Gamma) = - \int d\Gamma B(\Gamma) iL f(0) \quad (3.52)$$

This is true for an arbitrary distribution function  $f(0)$ . To prove this identity the LHS can be written as

$$\begin{aligned}
 \int d\Gamma f(\Gamma) \dot{\Gamma} \cdot \frac{\partial}{\partial \Gamma} B(\Gamma) &= \left[ f(\Gamma) \dot{\Gamma} B(\Gamma) \right]_S - \int d\Gamma B(\Gamma) \frac{\partial}{\partial \Gamma} \cdot (f(\Gamma) \dot{\Gamma}) \\
 &= - \int d\Gamma B(\Gamma) \left( \dot{\Gamma} \cdot \frac{\partial}{\partial \Gamma} + \frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} \right) f(\Gamma) \\
 &= - \int d\Gamma B(\Gamma) iL f(\Gamma)
 \end{aligned} \tag{3.53}$$

The boundary term (or surface integral) is zero because  $f(S) \rightarrow 0$  as any component of the momentum goes to infinity, and  $f$  can be taken to be periodic in all coordinates. If the coordinate space for the system is bounded then the surface  $S$  is the system boundary, and the surface integral is again zero as there can be no flow through the boundary.

Equations (3.52 & 53) show that  $L$ ,  $L$  are adjoint operators. If the equations of motion are such that the phase space compression factor, (3.43), is identically zero, then obviously  $L=L$  and the Liouville operator is self-adjoint, or *Hermitian*.

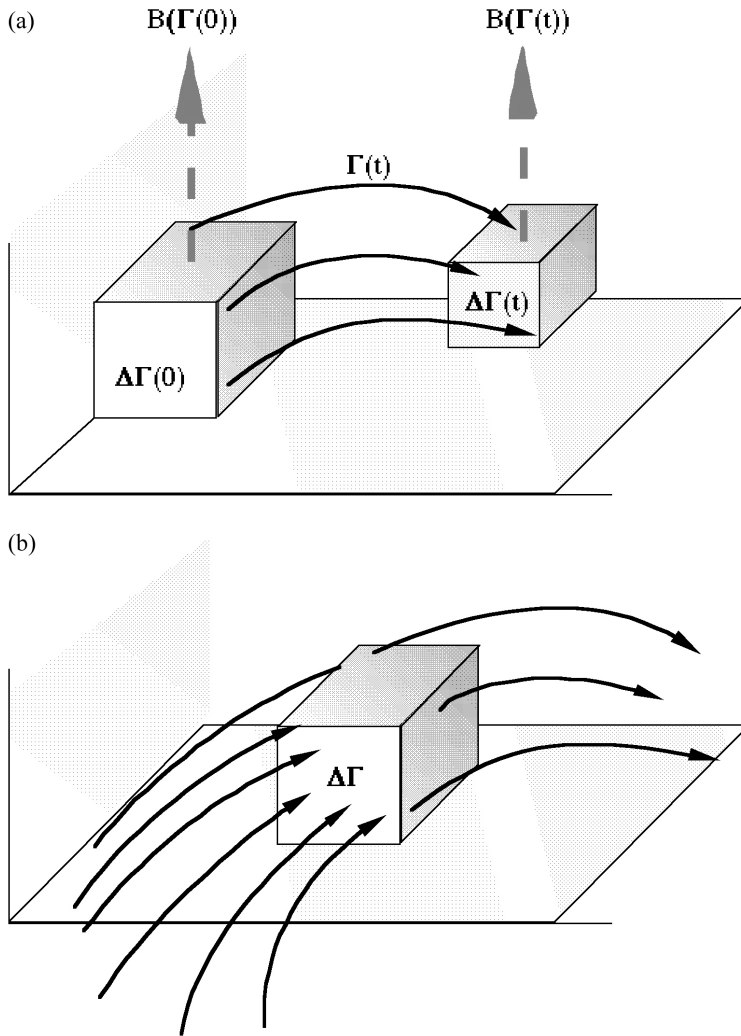
## Schrödinger and Heisenberg Representations

We can calculate the value of a phase variable  $B(t)$  at time  $t$  by following  $B$  as it changes along a single trajectory in phase space. The average  $\langle B(\Gamma(t)) \rangle$  can then be calculated by summing the values of  $B(t)$  with a weighting factor determined by the probability of starting from each initial phase  $\Gamma$ . These probabilities are chosen from an initial distribution function  $f(\Gamma, 0)$ . This is the Heisenberg picture of phase space averages.

$$\langle B(t) \rangle = \int d\Gamma B(t) f(\Gamma) = \int d\Gamma f(\Gamma) \exp[iL t] B(\Gamma) \tag{3.54}$$

The Heisenberg picture is exactly analogous to the Lagrangian formulation of fluid mechanics; we can imagine that the phase space *mass point* has a differential box  $d\Gamma$  surrounding it which changes shape (and volume for a compressible fluid) with time as the phase point follows its trajectory. The probability of the differential element, or mass  $f(\Gamma)d\Gamma$  remains constant, but the value of the observable changes implicitly in time.

The second view is the Schrödinger, or distribution based picture (Fig. 3.3).



**Figure 3.3 The Schrödinger-Heisenberg Equivalence**

(a) The Heisenberg picture:  $\dot{B} = \dot{\Gamma} \cdot \frac{\partial}{\partial \Gamma} B \equiv iL(\Gamma) B(\Gamma)$

(b) The Schrödinger picture:  $\frac{\partial f(\Gamma, t)}{\partial t} = -\frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} f(\Gamma, t) \equiv -iL f(\Gamma, t); f(\Gamma, t) \sim \frac{\Delta N(t)}{\Delta \Gamma}$  = the local density of ensemble representatives.

In this case we note that  $\langle B(t) \rangle$  can be calculated by sitting at a particular point in phase space and calculating the density of ensemble points as a function of time. This will give us the time dependent  $N$ -particle distribution function  $f(\Gamma, t)$ . The average of  $B$  can now be calculated by summing the values of  $B(\Gamma)$  but weighting these values by the current value of the distribution function at that place in phase space. Just as in the Eulerian formulation of fluid mechanics, the observable takes on a fixed value  $B(\Gamma)$  for all time, while mass points with different probability flow through the box.

$$\langle B(t) \rangle = \int d\Gamma B(\Gamma) f(\Gamma, t) = \int d\Gamma B(\Gamma) \exp[-iLt] f(\Gamma, 0) \quad (3.55)$$

The average value of  $B$  changes with time as the distribution function changes. The average of  $B$  is computed by multiplying the value of  $B(\Gamma)$ , by the probability of find the phase point  $\Gamma$  at time  $t$ , that is  $f(\Gamma, t)$ .

As we have just seen these two pictures are of course equivalent. One can also prove their equivalence using the Liouville equation. This proof is obtained by successive integrations by parts, or equivalently by repeated applications of equation (3.52). Consider

$$\begin{aligned} \int d\Gamma f(\Gamma) B(t) &= \int d\Gamma f(\Gamma) \exp[iLt] B(\Gamma) = \sum_{n=0}^{\infty} \frac{1}{n!} \int d\Gamma f(\Gamma) (iLt)^n B(\Gamma) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int d\Gamma f(\Gamma) \left( i\dot{\Gamma} \cdot \frac{\partial}{\partial \Gamma} \right)^n B(\Gamma) \end{aligned} \quad (3.56)$$

One can *unroll* each  $P$ -Liouvillean in turn from the phase variable to the distribution function (for the first transfer we consider  $(iL)^{n-1} B$  to be a composite phase variable) so that equation (3.56) becomes,

$$\int d\Gamma f(\Gamma) B(t) = \sum_{n=0}^{\infty} \frac{1}{n!} \int d\Gamma \left( -t \frac{\partial}{\partial \Gamma} \cdot (\dot{\Gamma} f(\Gamma)) \right) \left( i\dot{\Gamma} \cdot \frac{\partial}{\partial \Gamma} \right)^{n-1} B(\Gamma)$$

This is essentially the property of phase and distribution function Liouvillians which we have already proved, applied to  $n^{\text{th}}$  Liouvillean. Repeated application of this result leads to

$$\int d\Gamma f(\Gamma) B(t) = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} \int d\Gamma \left( \left( \frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} \right)^n f(\Gamma) \right) B(\Gamma) = \int d\Gamma B(\Gamma) \exp[-iLt] f(\Gamma)$$

So finally we have the result,

$$\int d\Gamma f(\Gamma) B(t) = \int d\Gamma B(\Gamma) f(\Gamma, t) \quad (3.57)$$

The derivation we have used assumes that the Liouvillean for the system has no explicit time dependence. (In Chapter 8 we will extend the derivation of these and other results to the time dependent case.) Our present derivation make no

other references to the details of either the initial distribution function, or the equations of motion for the system. This means that these results are valid for systems subject to time independent external fields, whether or not those equations are derivable from an Hamiltonian. These results are also independent of whether or not the phase space compression factor vanishes identically.

A final point that can be made concerning the Schrödinger and Heisenberg pictures is that these two ways of computing phase averages by no means exhaust the range of possibilities. The Schrödinger and Heisenberg pictures differ in terms of the time chosen to calculate the distribution function,  $f(\Gamma, t)$ . In the Heisenberg picture that time is zero while in the Schrödinger picture the time is  $t$ . One can of course develop intermediate representations corresponding any time between 0 and  $t$  (eg. the interaction representation).

### 3.4 Ergodicity, Mixing and Lyapunov Exponents

For many systems it is apparent that after possible initial transients lasting a time  $t_0$ , the N particle distribution function  $f(\Gamma, t)$ , becomes essentially time independent. This is evidenced by the fact that the macroscopic properties of the system relax to fixed average values. This obviously happens for equilibrium systems. It also occurs in some nonequilibrium systems, so-called nonequilibrium steady states. We will call all such systems stationary.

For a stationary system, we may define the ensemble average of a phase variable  $B(\Gamma)$ , using the stationary distribution function  $f(\Gamma)$ , so that

$$\langle B \rangle = \int d\Gamma f(\Gamma) B(\Gamma) \quad (3.58)$$

On the other hand we may define a time average of the same phase variable as,

$$\langle B \rangle_t = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} dt B(t) \quad (3.59)$$

where  $t_0$  is the relaxation time required for the establishment of the stationary state. An ergodic system is a stationary system for which the ensemble and time averages of *usual* phase variables, exist and are equal. By *usual* we mean phase variable representations of the common macroscopic thermodynamic variables (see §3.7).

It is commonly believed that all realistic nonlinear many body systems are ergodic.

#### Example

We can give a simple example of ergodic flow if we take the energy surface to be the two-dimensional unit square  $0 < p < 1$  and  $0 < q < 1$ . We shall assume that the equations of motion are given by

$$\begin{aligned}\dot{p} &= \alpha \\ \dot{q} &= 1\end{aligned}\tag{3.60}$$

and we impose periodic boundary conditions on the system. These equations of motion can be solved to give

$$\begin{aligned}p(t) &= p_0 + \alpha t \\ q(t) &= q_0 + t\end{aligned}\tag{3.61}$$

The phase space trajectory on the energy surface is given by eliminating  $t$  from these two equations

$$p = p_0 + \alpha(q - q_0)\tag{3.62}$$

If  $\alpha$  is a rational number,  $\alpha = m/n$ , then the trajectory will be periodic and will repeat after a period  $T = n$ . If  $\alpha$  is irrational, then the trajectory will be dense on the unit square but will not fill it. When  $\alpha$  is irrational the system is ergodic. To show this explicitly consider the Fourier series expansion of an arbitrary phase function  $A(q, p)$ ,

$$A(q, p) = \sum_{j, k=-\infty}^{\infty} A_{jk} \exp(2\pi i(jq + kp))\tag{3.63}$$

We wish to show that the time average and phase average of  $A(q, p)$  are equal for  $\alpha$  irrational. The time average is given by

$$\begin{aligned}\langle A \rangle_t &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} dt \sum_{j, k=-\infty}^{\infty} A_{jk} \exp(2\pi i[j(q_0 + t) + k(p_0 + \alpha t)]) \\ &= A_{00} + \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{j, k \neq 0}^{\infty} A_{jk} \exp(2\pi i[j(q_0 + t_0) + k(p_0 + \alpha t_0)]) \frac{e^{2\pi i(j+\alpha k)T} - 1}{2\pi i(j + \alpha k)}\end{aligned}\tag{3.64}$$

For irrational  $\alpha$ , the denominator can never be equal to zero, therefore

$$\langle A \rangle_t = A_{00}\tag{3.65}$$

Similarly we can show that the phase space average of  $A$  is

$$\langle A \rangle_{qp} = \int_0^1 dq \int_0^1 dp A(q, p) = A_{00}\tag{3.66}$$

and hence the system is ergodic. For rational  $\alpha$  the denominator in (3.64) does become singular for a particular  $jk$ -mode. The system is in the pure state labelled by  $jk$ . There is no mixing.

Ergodicity does **not** guarantee the relaxation of a system toward a stationary state. Consider a probability density which is not constant over the unit square, for example let  $f(q, p, t=0)$  be given by

$$f(q, p, 0) = \sin(\pi p_0) \sin(\pi q_0)\tag{3.67}$$



then at time  $t$ , under the above dynamics (with irrational  $\alpha$ ), it will be

$$f(q, p, t) = \sin(\pi(p_0 - \alpha t)) \sin(\pi(q_0 - t)) \quad (3.68)$$

The probability distribution is not changed in shape, it is only displaced. It has also **not** relaxed to a time independent equilibrium distribution function. However after an infinite length of time it will have wandered uniformly over the entire energy surface. It is therefore ergodic but it is termed *non mixing*.

It is often easier to show that a system is not ergodic, rather than to show that it is ergodic. For example the phase space of a system must be metrically transitive for it to be ergodic. That is, all of phase space, except possibly a set of measure zero, must be accessible to *almost all* the trajectories of the system. The reference to *almost all*, is because of the possibility that a set of initial starting states of measure zero, may remain forever within a subspace of phase space which is itself of measure zero. Ignoring the more pathological cases, if it is possible to divide phase space into two (or more) finite regions of nonzero measure, so that trajectories initially in a particular region remain there forever, then the system is not ergodic. A typical example would be a system in which a particle was trapped in a certain region of configuration space. Later we shall see examples of this specific type.

## Lyapunov Exponents

If we consider two harmonic oscillators (see §3.2) which have the same frequency  $\omega$  but different initial conditions  $(x_1, p_1)$  and  $(x_2, p_2)$ , we can define the *distance* between the two phase points by

$$d = \|\Gamma\| = (\Gamma \cdot \Gamma)^{\frac{1}{2}} = \sqrt{(x_2 - x_1)^2 + \frac{(p_2 - p_1)^2}{m^2 \omega^2}} \quad (3.69)$$

Using the equation for the trajectory of the harmonic oscillator (3.34), we see that as a function of time this distance is given by

$$d(t) = \sqrt{(x_2(t) - x_1(t))^2 + \frac{(p_2(t) - p_1(t))^2}{m^2 \omega^2}} = d(0) \quad (3.70)$$

where  $x_i(t)$  and  $p_i(t)$  are the position and momenta of oscillator  $i$ , at time  $t$ . This means that the trajectories of two independent harmonic oscillators always remain the same distance apart in  $\Gamma$ -space.

This is not the typical behaviour of **nonlinear** systems. The neighbouring trajectories of most  $N$ -body nonlinear systems tend to drift apart with time. Indeed it is clear that if a system is to be mixing then the separation of neighbouring trajectories is a precondition. Weakly coupled harmonic oscillators are an exceptions to the generally observed trajectory separation. This was a

cause of some concern in the earliest dynamical simulations (Fermi, Pasta & Ulam, 1955).

As the separation between neighbouring trajectories can be easily calculated in a classical mechanical simulation, this has been used to obtain quantitative measures of the mixing properties of nonlinear many-body systems. If we consider two  $N$ -body systems composed of particles which interact via identical sets of interparticle forces, but whose initial conditions differ by a small amount, then the phase space separation is observed change exponentially as

$$d(t) \equiv \sqrt{(\Gamma_1(t) - \Gamma_2(t))^2} \approx c \exp(\lambda t) \quad (3.71)$$

At intermediate times the exponential growth of  $d(t)$  will be dominated by the fastest growing direction in phase space (which in general will change continuously with time). This equation defines the largest *Lyapunov exponent*  $\lambda$  for the system (by convention  $\lambda$  is defined to be real, so any oscillating part of the trajectory separation is ignored). For the harmonic oscillator the phase separation is a constant of the motion and therefore the Lyapunov exponent  $\lambda$ , is zero. In practical applications this exponential separation for an  $N$  particle system continues until it approaches a limit imposed by the externally imposed boundary conditions - the container walls, or the energy, or other thermodynamic constraints on the system (§ 7.8). If the system has energy as a constant of the motion then the maximum separation is the maximum distance between two points on the energy hypersphere. This depends upon the value of the energy and the dimension of the phase space.

The largest Lyapunov exponent indicates the rate of growth of trajectory separation in phase space. If we consider a third phase point  $\Gamma_3(t)$ , which is constrained such that the vector between  $\Gamma_1$  and  $\Gamma_3$  is always orthogonal to the vector between  $\Gamma_1$  and  $\Gamma_2$ , then we can follow the rate of change of a two dimensional area in phase space. We can use these two directions to define an area element  $V_2(t)$ , and rate of change of the volume element is given by

$$V_2(t) = V_2(0) \exp([\lambda_1 + \lambda_2]t) \quad (3.72)$$

As we already know the value of  $\lambda_1$ , this defines the second largest Lyapunov exponent  $\lambda_2$ . In a similar way, if we construct a third phase space vector  $\Gamma_{14}(t)$  which is constrained to be orthogonal to both  $\Gamma_{12}(t)$  and  $\Gamma_{13}(t)$ , then we can follow the rate of change of a three dimensional volume element  $V_3(t)$  and calculate the third largest exponent  $\lambda_3$ ;

$$V_3(t) = V_3(0) \exp([\lambda_1 + \lambda_2 + \lambda_3]t) \quad (3.73)$$

This construction can be generalised to calculate the full spectrum of Lyapunov exponents for an  $N$  particle system. We consider the trajectory  $\Gamma(t)$  of a

dynamical system in phase space and study the convergence or divergence of neighbouring trajectories by taking a set of basis vectors (tangent vectors) in phase space  $\{\delta_1, \delta_2, \delta_3, \dots\}$ , where  $\delta_i = \Gamma_i - \Gamma_0$ . Some care must be exercised in forming the set of basis vectors to ensure that the full dimension of phase space is spanned by the basis set, and that the basis set is minimal. This simply means that constants of the motion must be considered when calculating the dimension of accessible phase space. If the equation of motion for a trajectory is of the form

$$\dot{\Gamma} = G(\Gamma) \quad (3.74)$$

then the equation of motion for the tangent vector  $\delta_i$  is

$$\dot{\delta}_i = \mathbf{F}_i(\Gamma) = \mathbf{T}(\Gamma) \cdot \delta_i + O(\delta_i^2) \quad (3.75)$$

Here  $\mathbf{T}(\Gamma)$  is the Jacobian matrix (or stability matrix  $\partial G/\partial \Gamma$ ) for the system. If the magnitude of the tangent vector is small enough the nonlinear terms in equation (3.75) can be neglected. The formal solution of this equation is

$$\delta_i(t) = \exp\left[\int_0^t ds \mathbf{T}(s)\right] \delta_i(0) \quad (3.76)$$

The mean exponential rate of growth of the  $i^{\text{th}}$  tangent vector, gives the  $i^{\text{th}}$  Lyapunov exponent

$$\lambda_i(\Gamma(0), \delta_i(0)) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{\|\delta_i(t)\|}{\|\delta_i(0)\|} \quad (3.77)$$

The existence of the limit is ensured by the multiplicative ergodic theorem of Oseledec [1968] (see also Eckmann and Ruelle [1985]). The Lyapunov exponents can be ordered  $\lambda_1 > \lambda_2 > \dots > \lambda_M$  and if the system is ergodic, the exponents are independent of the initial phase  $\Gamma(0)$  and the initial phase space separation  $\delta_i(0)$ .

If we consider the volume element  $V_N$  where  $N$  is the dimension of phase space then we can show that the phase space compression factor gives the rate of change of phase space volume, and that this is simply related to the sum of the Lyapunov exponents by

$$\dot{V}_N = \left\langle \frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} \right\rangle V_N = \left( \sum_{i=1}^N \lambda_i \right) V_N \quad (3.78)$$

For a Hamiltonian system, the phase space compression factor is identically zero, so the phase space volume is conserved. This is a simple consequence of Liouville's theorem. From equation (3.78) it follows that the sum of the Lyapunov exponents is also equal to zero. If the system is time reversible then the Lyapunov exponents occur in pairs  $(-\lambda_i, \lambda_i)$ . This ensures that  $d(t)$ ,  $V_2(t)$ ,  $V_3(t)$ , etc. change at the same rate with both forward and backward time evolution. It is generally believed that it is necessary to have at least one positive Lyapunov exponent

for the system to be mixing. In chapters 7 and 10 we will return to consider Lyapunov exponents in both equilibrium and nonequilibrium systems.

### 3.5 Equilibrium Time Correlation Functions

We shall often refer to averages over equilibrium distribution functions  $f_0$  (we use the subscript zero to denote equilibrium, which should not be confused with  $f(0)$ , a distribution function at  $t = 0$ ). Distribution functions are called equilibrium if they pertain to steady, unperturbed equations of motion and they have no explicit time dependence. An equilibrium distribution function satisfies a Liouville equation of the form

$$\frac{\partial}{\partial t} f_0 = -iL f_0 = 0 \quad (3.79)$$

This implies that the equilibrium average of any phase variable is a stationary quantity. That is, for an arbitrary phase variable  $B$ ,

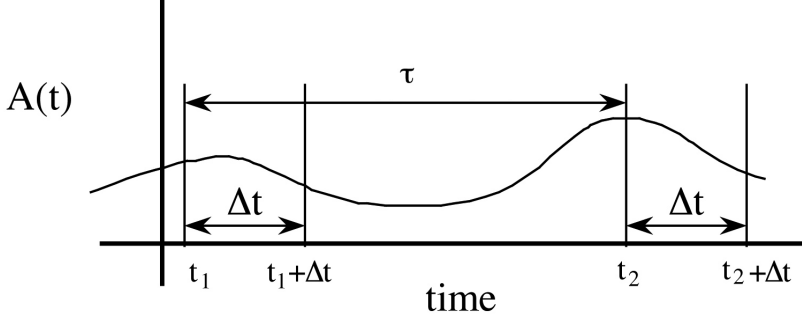
$$\begin{aligned} \frac{d}{dt} \langle B(t) \rangle_0 &= \frac{d}{dt} \int d\Gamma f_0(\Gamma) \exp(iLt) B(\Gamma) = \int d\Gamma f_0(\Gamma) \frac{\partial}{\partial t} \exp(iLt) B(\Gamma) \\ &= \int d\Gamma f_0(\Gamma) iL \exp(iLt) B(\Gamma) = - \int d\Gamma (iL f_0(\Gamma)) \exp(iLt) B(\Gamma) = 0 \end{aligned} \quad (3.80)$$

We will often need to calculate the equilibrium time correlation function of a phase variable  $A$  with another phase variable  $B$  at some other time. We define the equilibrium time correlation function of  $A$  and  $B$  by

$$C_{AB}(t) \equiv \int d\Gamma f_0 B^* e^{iLt} A = \langle A(t) B^* \rangle_0 \quad (3.81)$$

where  $B^*$  denotes the complex conjugate of the phase variable  $B$ . Sometimes we will refer to the autocorrelation function of a phase variable  $A$ . If this variable is real, one can form a simple graphical representation of how such functions are calculated (see Fig. 3.4).

Because the averages are to be taken over a stationary equilibrium distribution function, time correlation functions are only sensitive to time difference between which  $A$  and  $B$  are evaluated.  $C_{AB}(t)$  is independent of the particular choice of the time origin. If  $iL$  generates the distribution function  $f_0$ , then the propagator  $\exp(-iLt)$  preserves  $f_0$ . (The converse is not necessarily true.) To be more explicit  $f_0(t_1) = \exp(-iLt_1) f_0 = f_0$ , so that  $C_{AB}(t)$  becomes



**Figure 3.4** Equilibrium time autocorrelation function of real variable **A**.

$\langle A(0)A(\Delta t) \rangle = \frac{1}{N_\tau} \sum_{i=1}^{N_\tau} A(t_i)A(t_i + \Delta t)$  ,  $t_i = 0, \tau, 2\tau, \dots$  for samples in the sum to be independent,  $\tau$  should be chosen so that  $\langle A(0)A(\tau) \rangle \ll \langle A(0)^2 \rangle$

$$\begin{aligned}
 C_{AB}(t) &= \int d\Gamma f_0 B^* e^{iL t} A = \int d\Gamma f_0(t_1) B^* e^{iL t} A \\
 &= \int d\Gamma \left( e^{-iL t_1} f_0 \right) B^* e^{iL t} A \\
 &= \int d\Gamma f_0 \left( e^{iL t_1} B^* \right) e^{iL(t_1 + t)} A \\
 &= \int d\Gamma f_0 A(t_1 + t) B^*(t_1)
 \end{aligned} \tag{3.82}$$

In deriving the last form of (3.82) we have used the important fact that since  $iL = d\Gamma/dt \cdot \partial/\partial\Gamma$  and the equations of motion are real it follows that  $L$  is pure imaginary. Thus,  $(iL)^* = iL$  and  $(e^{iL t})^* = e^{iL t}$ . Comparing (3.82) with the definition of  $C_{AB}(t)$ , above we see that the equilibrium time correlation function is independent of the choice of time origin. It is solely a function of the difference in time of the two arguments,  $A$  and  $B$ . A further identity follows from this result if we choose  $t_1 = -t$ . We find that

$$C_{AB}(t) = \langle A(t) B^*(0) \rangle_0 = \langle A(0) B^*(-t) \rangle_0 \tag{3.83}$$

So that,

$$C_{AB}^*(t) = \langle A^* B(-t) \rangle = C_{BA}(-t) \tag{3.84}$$

or using the notation of section 3.3,

$$\begin{aligned}
 \left[ \int d\Gamma f_0 B^* e^{iL t} A \right]^* &= \left[ \int d\Gamma A e^{-iL t} (f_0 B^*) \right]^* = \left[ \int d\Gamma f_0 A e^{-iL t} B^* \right]^* \\
 &= \int d\Gamma f_0 A^* e^{-iL t} B
 \end{aligned} \tag{3.85}$$

The second equality in equation (3.85) follows by expanding the operator  $\exp(-iL t)$  and repeatedly applying the identity

$$\begin{aligned}
 iL(f_0 B^*) &= \frac{\partial}{\partial \Gamma} \cdot (\dot{\Gamma} f_0 B^*) = B^* \frac{\partial}{\partial \Gamma} \cdot (\dot{\Gamma} f_0) + f_0 \dot{\Gamma} \cdot \frac{\partial}{\partial \Gamma} B^* \\
 &= B^* iL f_0 + f_0 iL B^* = f_0 iL B^*
 \end{aligned}$$

The term  $iL f_0$  is zero from equation (3.79).

Over the scalar product defined by equation (3.81),  $L$  is an Hermitian operator. The Hermitian adjoint of  $L$  denoted,  $L^\dagger$  can be defined by the equation,

$$\left[ \int d\Gamma f_0 B^* e^{iL t} A \right]^* \equiv \int d\Gamma f_0 A^* \exp[-iL^\dagger t] B \quad (3.86)$$

Comparing (3.86) with (3.85) we see two things: we see that the Liouville operator  $L$  is self adjoint or **Hermitian** ( $L = L^\dagger$ ); and therefore the propagator  $e^{iL t}$ , is **unitary**. This result stands in contrast to those of §3.3, for arbitrary distribution functions.

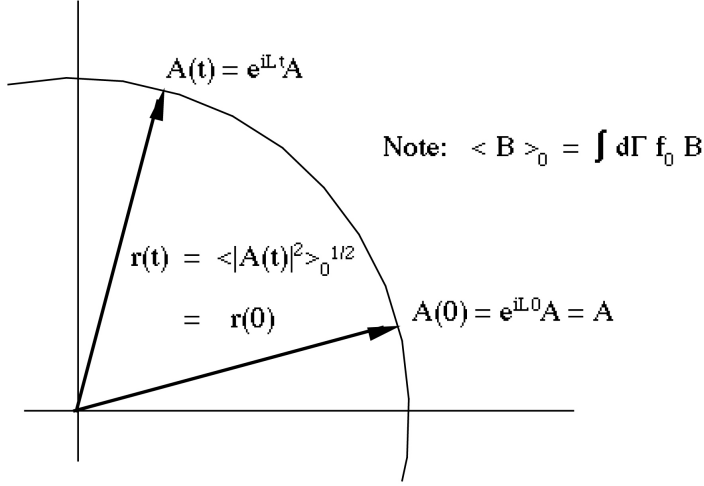
We can use the autocorrelation function of  $A$  to define a norm in Liouville space. This length or norm of a phase variable  $A$ , is defined by the equation,

$$\|A\|^2 = \int d\Gamma f_0 A(\Gamma) A^*(\Gamma) = \int d\Gamma f_0 |A(\Gamma)|^2 = \langle |A(\Gamma)|^2 \rangle_0 \geq 0 \quad (3.87)$$

We can see immediately that the norm of any phase variable is time independent because

$$\begin{aligned}
 \|A(t)\|^2 &= \int d\Gamma f_0 A(t) A^*(t) = \int d\Gamma f_0 [\exp(iL t) A(\Gamma)] [\exp(iL t) A^*(\Gamma)] \\
 &= \int d\Gamma f_0 \exp(iL t) (A(\Gamma) A^*(\Gamma)) = \int d\Gamma (\exp(-iL t) f_0) |A|^2 = \|A(0)\|^2
 \end{aligned} \quad (3.88)$$

The propagator is said to be norm preserving (Fig. 3.5). This is a direct result of the fact that the propagator is a unitary operator. The propagator can be thought of as a rotation operator in Liouville space.



**Figure 3.5 The propagator is norm preserving**

A phase variable whose norm is unity is said to be normalised. The scalar product,  $(A, B)$  of two phase variables  $A, B$  is simply the equilibrium average of  $A$  and  $B$ , namely  $\langle AB \rangle_0$ . The norm of a phase variable is simply the scalar product of the variable with itself. The autocorrelation function  $C_{AA}(t)$  has a zero time value which is equal to the norm of  $A$ . The propagator increases the angle between  $A$  and  $A(t)$ , and the scalar product which is the projection of  $A(t)$  along  $A$ , therefore decreases. The autocorrelation function of a given phase variable therefore measures the rate at which the  $6N$ -dimensional rotation occurs.

We will now derive some relations for the time derivatives of time correlation functions. It is easy to see that

$$\begin{aligned}
 \frac{d}{dt} C_{AB}(t) &= \frac{d}{dt} \int d\Gamma f_0 (\exp[iLt] A(\Gamma)) B^*(\Gamma) \\
 &= \int d\Gamma f_0 (iL \exp[iLt] A(\Gamma)) B^*(\Gamma) = C_{AB}(t) \\
 &= - \int d\Gamma (e^{iLt} A(\Gamma)) iL (f_0 B^*(\Gamma)) = - \int d\Gamma A(t) \frac{\partial}{\partial \Gamma} (\Gamma f_0 B^*(\Gamma)) \\
 &= - \int d\Gamma f_0 (e^{iLt} A(\Gamma)) iL B^*(\Gamma) \\
 &= - \int d\Gamma f_0 (e^{iLt} A(\Gamma)) \dot{B}^*(\Gamma) = -C_{AB}(t)
 \end{aligned} \tag{3.90}$$

### 3.6 Operator Identities

In this section we develop the operator algebra that we will need to manipulate expressions containing Liouvillians and their associated propagators. Most of the identities which we obtain are valid for arbitrary time independent operators. Thus far we have been dealing with propagators in the time domain. For many

problems it is more useful to consider their frequency dependent Laplace, or Fourier-Laplace, transforms. A useful mathematical object is the Laplace transform of the propagator. This is called the resolvent. The resolvent is an operator in the domain of the Laplace transform variable  $s$ ,

$$G(s) = \int_0^{\infty} dt e^{-st} e^{-iLt} \quad (3.91)$$

Our first operator identity is obtained by considering two arbitrary operators  $A$  and  $B$ ,

$$(A + B)^{-1} = A^{-1} - A^{-1}B(A + B)^{-1} \quad (3.92)$$

This identity is easily verified by operating from the right-hand side of this equation with  $(A + B)$ , so

$$\begin{aligned} (A + B)^{-1}(A + B) &= [A^{-1} - A^{-1}B(A + B)^{-1}](A + B) \\ &= A^{-1}(A + B) - A^{-1}B \\ &= A^{-1}A + A^{-1}B - A^{-1}B = I \end{aligned} \quad (3.93)$$

The operator expression  $(A + B)^{-1}$  is the inverse of the operator  $(A + B)$ . To interpret an operator inverse of  $(A + B)^{-1}$ , we use the series expansion

$$(I + A)^{-1} = \sum_{n=0}^{\infty} (-A)^n \quad (3.94)$$

First we prove that the right-hand side of this expression is indeed the inverse of the operator  $(I + A)$ . To do this consider

$$\sum_{n=0}^{\infty} (-A)^n (I + A) = \sum_{n=0}^{\infty} (-A)^n - \sum_{n=1}^{\infty} (-A)^n = I \quad (3.95)$$

so that this series expansion allows us to represent the inverse of  $(I + A)$  in terms of an infinite series of products of the operator  $A$ .

## The Dyson Decomposition of Propagators

Now we can investigate the Laplace transform (or resolvent) of the exponential of an operator in more detail. We use the expansion of the exponential to show that

$$\begin{aligned} \int_0^{\infty} dt e^{-st} e^{-At} &= \int_0^{\infty} dt e^{-st} \sum_{n=0}^{\infty} \frac{(-)^n}{n!} A^n t^n = \sum_{n=0}^{\infty} \frac{(-)^n}{n!} A^n \int_0^{\infty} dt e^{-st} t^n \\ &= \sum_{n=0}^{\infty} (-)^n \frac{A^n}{s^{n+1}} = \frac{1}{s} \left( I + \frac{A}{s} \right)^{-1} = (s + A)^{-1} \end{aligned} \quad (3.96)$$



This means that the resolvent of the operator,  $e^{-At}$ , is simply  $(s + A)^{-1}$ . We can now consider the resolvent derived from the operator  $(A + B)$ , and using the first identity above, relate this resolvent to the resolvent of  $A$ . We can write

$$(s + A + B)^{-1} = (s + A)^{-1} - (s + A)^{-1} B (s + A + B)^{-1} \quad (3.97)$$

Substituting the Laplace integrals for the operators  $(s + A)^{-1}$  and  $(s + A + B)^{-1}$  into this equation gives

$$\begin{aligned} \int_0^\infty dt e^{-st} e^{-(A+B)t} &= \int_0^\infty dt e^{-st} e^{-At} - \int_0^\infty dt_1 e^{-st_1} e^{-At_1} B \int_0^\infty dt_2 e^{-st_2} e^{-(A+B)t_2} \\ &= \int_0^\infty dt e^{-st} \left\{ e^{-At} - \int_0^t dt_1 e^{-At_1} B e^{-(A+B)(t-t_1)} \right\} \end{aligned} \quad (3.98)$$

As the equality holds for all values of  $s$ , the integrands must be equal, so

$$e^{-(A+B)t} = e^{-At} - \int_0^t dt_1 e^{-At_1} B e^{-(A+B)(t-t_1)} \quad (3.99)$$

This result is a very important step towards understanding the relationship between different propagators and is referred to as the *Dyson decomposition* when applied to propagators (Dyson, 1949). The derivation that we have used here is only valid if both of the operators  $A$  and  $B$  have no explicit time dependence. (We consider the time dependent case in Chapter 8.) If we consider the propagators  $\exp((A + B)t)$  and  $\exp(At)$ , then a second Dyson decomposition can be obtained:

$$e^{(A+B)t} = e^{At} + \int_0^t dt_1 e^{At_1} B e^{(A+B)(t-t_1)} \quad (3.100)$$

It is handy to use a graphical shorthand for the Dyson equation. Using this shorthand notation these two equations become,

$$\Leftarrow = \Leftarrow - \Leftarrow (\blacklozenge - \circ) \Leftarrow \quad (3.101)$$

and

$$\Rightarrow = \Rightarrow + \Rightarrow (\blacklozenge - \circ) \Rightarrow \quad (3.102)$$

The diamond  $\blacklozenge$  denotes the  $(A + B)$ -Liouvillean and the circle  $\circ$  denotes the  $A$ -Liouvillean; the arrows  $\Leftarrow$  and  $\Rightarrow$  denote the propagators  $\exp(-(A + B)t)$  and  $\exp((A + B)t)$  respectively, while  $\Leftarrow$  and  $\Rightarrow$  denote  $\exp(-At)$  and  $\exp(At)$  respectively. An  $n - 1$  fold convolution is implied by a chain of  $n$  arrows.

As an example of the application of this result, consider the case where  $B$  is a small perturbation to the operator  $A$ . In this case the Dyson decomposition gives the full  $(A + B)$ -propagator as the sum of the unperturbed  $A$ -propagator plus a

correction term. One often faces the situation where we want to compare the operation of different propagators on either a phase variable or a distribution function. For example one might like to know the difference between the value of a phase variable  $A(\Gamma)$  propagated under the combined influence of the  $N$ -particle interactions **and** an applied external field  $F_e$ , with the value the phase variable might take at the same time in the absence of the external field. In that case (Evans and Morriss, 1984)

$$\begin{aligned} A(t, F_e) &= \Rightarrow A(\Gamma) \\ &= [\rightarrow + \rightarrow (\blacklozenge - \circ) \rightarrow + \rightarrow (\blacklozenge - \circ) \rightarrow (\blacklozenge - \circ) \rightarrow \\ &+ \rightarrow (\blacklozenge - \circ) \rightarrow (\blacklozenge - \circ) \rightarrow (\blacklozenge - \circ) \rightarrow \\ &+ \rightarrow (\blacklozenge - \circ) \rightarrow (\blacklozenge - \circ) \rightarrow (\blacklozenge - \circ) \rightarrow (\blacklozenge - \circ) \rightarrow \\ &+ \dots \dots \dots ] A(\Gamma) \end{aligned}$$

Therefore we can write,

$$\Rightarrow A(\Gamma) = \sum_{n=0}^{\infty} \rightarrow [(\blacklozenge - \circ) \rightarrow]^n A(\Gamma) \quad (3.103)$$

This equation is of limited usefulness because in general,  $\blacklozenge$  and  $\rightarrow$ , do not commute. This means that the Liouvillean frequently represented by  $\blacklozenge$ , is locked inside a convolution chain of propagators with which it does not commute. A more useful expression can be derived from (3.102) by realising that  $\blacklozenge$  commutes with its own propagator namely,  $\Rightarrow$ . Similarly  $\circ$  commutes with its own propagator,  $\rightarrow$ . We can 'unlock' the respective Liouvilleans from the chain in (3.102) by writing,

$$\Rightarrow = \rightarrow + \blacklozenge \Rightarrow \rightarrow - \Rightarrow \rightarrow \circ \quad (3.104)$$

We can recursively substitute for  $\Rightarrow$ , yielding,

$$\begin{aligned} \Rightarrow &= \rightarrow + \blacklozenge \rightarrow \rightarrow - \rightarrow \rightarrow \circ \\ &+ \blacklozenge \blacklozenge \rightarrow \rightarrow \rightarrow - 2 \blacklozenge \rightarrow \rightarrow \rightarrow \circ + \rightarrow \rightarrow \rightarrow \circ \circ \\ &+ \dots \dots \dots \end{aligned} \quad (3.105)$$

Now it is easy to show that,

$$(\rightarrow)^{n+1} = \left( \frac{t^n}{n!} \right) \rightarrow \quad (3.106)$$

Thus (3.105) can be written as,

$$\begin{aligned} \Rightarrow & \{1 + t(\diamond - \circ) + (t^2/2!)(\diamond\diamond - 2\diamond\circ + \circ\circ) \\ & + (t^3/3!)(\diamond\diamond\diamond - 3\diamond\diamond\circ + 3\diamond\circ\circ - \circ\circ\circ) \\ & + \dots\} \rightarrow \end{aligned} \quad (3.107)$$

This equation was first derived by Evans and Morriss (1984). Its utility arises from the fact that by 'unrolling' the Liouville operators to the left and the propagator to the right, **explicit** formulae for the expansion can usually be derived. A limitation of the formula is that successive terms on the right hand side do **not** constitute a power series expansion of the difference in the two propagators in powers of the difference between the respective Liouvillians. To be more explicit, the term,  $\frac{t^3}{3!}(\diamond\diamond\diamond - 3\diamond\diamond\circ + 3\diamond\circ\circ - \circ\circ\circ)$  is **not** in general of order  $(\diamond - \circ)^3$ .

### Campbell-Baker-Hausdorff Theorem

If  $A$  and  $B$  are non commuting operators then the operator expression  $\exp(A)\exp(B)$  can be written in the form  $\exp(C)$  where  $C$  is given by

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}\{[[A, B], B] + [[B, A], A]\} + \dots \quad (3.108)$$

The notation  $[.,.]$  is the usual Quantum Mechanical commutator. A rearrangement of this expansion, known as the Magnus expansion is well known to quantum theorists (Magnus, 1954). Any finite truncation of the Magnus expansion for the time displacement operator, gives a unitary time displacement operator approximation (Pechukas and Light, 1966). This result has not proved as useful for nonequilibrium statistical mechanics as it is for quantum theory. We give it here mainly for the sake of completeness.

### 3.7 The Irving-Kirkwood Procedure

In Chapter 2 we gave a brief outline of the structure of macroscopic hydrodynamics. We saw that given appropriate boundary conditions, it is possible to use the Navier-Stokes equations to describe the resulting macroscopic flow patterns. In this chapter we began the microscopic description of nonequilibrium systems using the Liouville equation. We will now follow a procedure first outlined by Irving and Kirkwood (1950), to derive microscopic expressions for the thermodynamic forces and fluxes appearing in the phenomenological equations of hydrodynamics.

In our treatment of the macroscopic equations we stressed the role played by the densities of conserved quantities. Our first task here will be to define microscopic expressions for the local densities of mass, momentum and energy. If the mass of the individual atoms in our system is  $m$  then the mass per unit

volume at a position  $\mathbf{r}$  and time  $t$  can be obtained by taking an appropriate average over the normalised  $N$ -particle distribution function  $f(\Gamma, t)$ . To specify that the particles should be at the macroscopic position  $\mathbf{r}$ , we will use a *macroscopic* delta function,  $\delta(\mathbf{r} - \mathbf{r}_i)$ . This *macroscopic* delta function is zero if atom  $i$  is outside some microscopic volume  $\delta V$ ; it is a constant if atom  $i$  is inside this volume ( $\delta$  is a smeared out version of the usual point delta function). We will assume that particle dynamics are given by field-free Newtonian equations of motion. The value of the constant is determined from the normalisation condition,

$$\int_V d\mathbf{r} \delta(\mathbf{r}) = 1 \quad (3.109)$$

The volume  $V$  is envisioned to be infinitesimal on a macroscopic scale.

The mass density  $\rho(\mathbf{r}, t)$  can be calculated from the following average,

$$\begin{aligned} \rho(\mathbf{r}, t) &= \int d\Gamma f(\Gamma, t) \sum_i m \delta(\mathbf{r} - \mathbf{r}_i) \\ &= \int d\Gamma f(\Gamma, 0) \sum_i m \delta(\mathbf{r} - \mathbf{r}_i(t)) = \left\langle \sum_i m \delta(\mathbf{r}_i(t) - \mathbf{r}) \right\rangle \end{aligned} \quad (3.110)$$

The first line of this equation is a Schrödinger representation of the density while the second and third lines are written in the Heisenberg representation. The equivalence of these two representations is easily seen by 'unrolling' the propagator from the distribution function onto the phase variable. Since  $\mathbf{r}$ , is a constant, a nominated position it is unchanged by this 'unrolling' procedure.

The momentum density,  $\rho(\mathbf{r}, t)\mathbf{u}(\mathbf{r}, t)$ , and total energy density,  $\rho(\mathbf{r}, t)e(\mathbf{r}, t)$ , are defined in an analogous manner.

$$\begin{aligned} \rho(\mathbf{r}, t)\mathbf{u}(\mathbf{r}, t) &= \int d\Gamma f(\Gamma, t) \sum_i m \mathbf{v}_i \delta(\mathbf{r} - \mathbf{r}_i) \\ &= \int d\Gamma f(\Gamma, t) \sum_i \mathbf{p}_i \delta(\mathbf{r} - \mathbf{r}_i) = \left\langle \sum_i \mathbf{p}_i(t) \delta(\mathbf{r}_i(t) - \mathbf{r}) \right\rangle \end{aligned} \quad (3.111)$$

$$\begin{aligned} \rho(\mathbf{r}, t)e(\mathbf{r}, t) &= \int d\Gamma f(\Gamma, t) \left[ \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_{ij} \phi_{ij} \right] \delta(\mathbf{r} - \mathbf{r}_i) \\ &= \left\langle \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_{ij} \phi_{ij} \right|_{\mathbf{r}_i(t) = \mathbf{r}} \end{aligned} \quad (3.112)$$

In these equations  $\mathbf{v}_i$  is the velocity of particle  $i$ ,  $\mathbf{p}_i$  is its momentum,  $\mathbf{r}_{ij} \equiv \mathbf{r}_j - \mathbf{r}_i$ , and we assume that the total potential energy of the system,  $\Phi$  is pair-wise additive and can be written as,

$$\Phi = \frac{1}{2} \sum_{i,j} \phi_{ij} \quad (3.113)$$

We arbitrarily assign one half of the potential energy to each of the two particles which contribute  $\Phi_{ij}$  to the total potential energy of the system.

The conservation equations involve time derivatives of the averages of the densities of conserved quantities. To begin, we will calculate the time derivative of the mass density.

$$\begin{aligned}
 \frac{\partial}{\partial t} \rho(\mathbf{r}, t) &= \int d\Gamma \frac{\partial f(\Gamma, t)}{\partial t} \sum_i m \delta(\mathbf{r} - \mathbf{r}_i) \\
 &= - \int d\Gamma \sum_i m \delta(\mathbf{r} - \mathbf{r}_i) iL f(\Gamma, t) \\
 &= \int d\Gamma f(\Gamma, t) iL \sum_i m \delta(\mathbf{r} - \mathbf{r}_i) \\
 &= \int d\Gamma f(\Gamma, t) \sum_i m \mathbf{v}_i \cdot \frac{\partial \delta(\mathbf{r} - \mathbf{r}_i)}{\partial \mathbf{r}_i} \\
 &= - \int d\Gamma f(\Gamma, t) \sum_i m \mathbf{v}_i \cdot \frac{\partial \delta(\mathbf{r} - \mathbf{r}_i)}{\partial \mathbf{r}} \\
 &= - \nabla \cdot \int d\Gamma f(\Gamma, t) \sum_i m \mathbf{v}_i \delta(\mathbf{r} - \mathbf{r}_i) \\
 &= - \nabla \cdot [\rho(\mathbf{r}, t) \mathbf{u}(\mathbf{r}, t)]
 \end{aligned} \tag{3.114}$$

The fifth equality follows using the delta function identity,

$$\frac{\partial}{\partial \mathbf{r}_i} \delta(\mathbf{r} - \mathbf{r}_i) \equiv - \frac{\partial}{\partial \mathbf{r}} \delta(\mathbf{r} - \mathbf{r}_i)$$

We have shown that the time derivative of the mass density yields the mass continuity equation (2.4) as expected. Strictly speaking therefore, we did not really need to define the momentum density in equation (3.111), as given the mass density definition, the mass continuity equation yields the momentum density expression. We will now use exactly the same procedure to differentiate the momentum density.

$$\begin{aligned}
 \frac{\partial}{\partial t} [\rho(\mathbf{r}, t) \mathbf{u}(\mathbf{r}, t)] &= \int d\Gamma \frac{\partial f(\Gamma, t)}{\partial t} \sum_i m_i \mathbf{v}_i \delta(\mathbf{r} - \mathbf{r}_i) \\
 &= - \int d\Gamma \sum_i m_i \mathbf{v}_i \delta(\mathbf{r} - \mathbf{r}_i) iL f(\Gamma, t) \\
 &= \int d\Gamma f(\Gamma, t) iL \sum_i m_i \mathbf{v}_i \delta(\mathbf{r} - \mathbf{r}_i) \\
 &= \int d\Gamma f(\Gamma, t) \sum_i \left[ \delta(\mathbf{r} - \mathbf{r}_i) iL m_i \mathbf{v}_i + m_i \mathbf{v}_i iL \delta(\mathbf{r} - \mathbf{r}_i) \right] \\
 &= \int d\Gamma f(\Gamma, t) \sum_i \left[ \delta(\mathbf{r} - \mathbf{r}_i) \mathbf{F}_i + m_i \mathbf{v}_i \cdot \frac{\partial \delta(\mathbf{r} - \mathbf{r}_i)}{\partial \mathbf{r}_i} \right]
 \end{aligned} \tag{3.115}$$

We have used Newtonian equations of motion for the Liouvillean  $iL$ .

$$\begin{aligned}\frac{\partial}{\partial t}[\rho(\mathbf{r},t)\mathbf{u}(\mathbf{r},t)] &= \int d\Gamma f(\Gamma,t) \sum_i \left[ \delta(\mathbf{r} - \mathbf{r}_i) \mathbf{F}_i + m_i \mathbf{v}_i \mathbf{v}_i \cdot \frac{\partial \delta(\mathbf{r} - \mathbf{r}_i)}{\partial \mathbf{r}_i} \right] \\ &= \int d\Gamma f(\Gamma,t) \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \mathbf{F}_i - \frac{\partial}{\partial \mathbf{r}} \cdot \int d\Gamma f(\Gamma,t) \sum_i m_i \mathbf{v}_i \mathbf{v}_i \delta(\mathbf{r} - \mathbf{r}_i)\end{aligned}$$

If we consider the second term on the right-hand side then

$$\begin{aligned}& \frac{\partial}{\partial \mathbf{r}} \cdot \int d\Gamma f(\Gamma,t) \sum_i m_i \mathbf{v}_i \mathbf{v}_i \delta(\mathbf{r} - \mathbf{r}_i) \\ &= \frac{\partial}{\partial \mathbf{r}} \cdot \left\langle \sum_i m_i (\mathbf{v}_i(t) - \mathbf{u}(\mathbf{r},t)) (\mathbf{v}_i(t) - \mathbf{u}(\mathbf{r},t)) \Big|_{\mathbf{r}_i(t)=\mathbf{r}} \right\rangle \\ &+ \frac{\partial}{\partial \mathbf{r}} \cdot \left\langle \sum_i m_i \mathbf{u}(\mathbf{r},t) \mathbf{u}(\mathbf{r},t) \Big|_{\mathbf{r}_i(t)=\mathbf{r}} \right\rangle\end{aligned}\tag{3.116}$$

In the final term in equation (3.116),  $\mathbf{u}(\mathbf{r},t)$  is independent of the particle index and can be factored outside the summation. The remaining summation is, using equation (3.110), simply equal to the mass density  $\rho(\mathbf{r},t)$ . Combining these results it follows that

$$\begin{aligned}\frac{\partial}{\partial t}[\rho(\mathbf{r},t)\mathbf{u}(\mathbf{r},t)] + \frac{\partial}{\partial \mathbf{r}} \cdot [\rho(\mathbf{r},t)\mathbf{u}(\mathbf{r},t)\mathbf{u}(\mathbf{r},t)] &= \int d\Gamma f(\Gamma,t) \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \mathbf{F}_i \\ &- \frac{\partial}{\partial \mathbf{r}} \cdot \left\langle \sum_i m_i (\mathbf{v}_i - \mathbf{u}(\mathbf{r},t)) (\mathbf{v}_i - \mathbf{u}(\mathbf{r},t)) \Big|_{\mathbf{r}_i(t)=\mathbf{r}} \right\rangle\end{aligned}\tag{3.117}$$

We will now consider the first term on the right hand side of this equation in some detail.

$$\begin{aligned}& \int d\Gamma f(\Gamma,t) \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \mathbf{F}_i \\ &= - \int d\Gamma f(\Gamma,t) \sum_{i,j} \delta(\mathbf{r} - \mathbf{r}_i) \frac{\partial \phi_{ij}}{\partial \mathbf{r}_i} \\ &= -\frac{1}{2} \int d\Gamma f(\Gamma,t) \sum_{i,j} \left[ \delta(\mathbf{r} - \mathbf{r}_i) \frac{\partial \phi_{ij}}{\partial \mathbf{r}_i} + \delta(\mathbf{r} - \mathbf{r}_j) \frac{\partial \phi_{ji}}{\partial \mathbf{r}_j} \right] \\ &= -\frac{1}{2} \int d\Gamma f(\Gamma,t) \sum_{i,j} \left[ \delta(\mathbf{r} - \mathbf{r}_i) \frac{\partial \phi_{ij}}{\partial \mathbf{r}_i} - \delta(\mathbf{r} - \mathbf{r}_j) \frac{\partial \phi_{ij}}{\partial \mathbf{r}_i} \right] \\ &= -\frac{1}{2} \int d\Gamma f(\Gamma,t) \sum_{i,j} \left[ \delta(\mathbf{r} - \mathbf{r}_i) - \delta(\mathbf{r} - \mathbf{r}_j) \right] \frac{\partial \phi_{ij}}{\partial \mathbf{r}_i}\end{aligned}\tag{3.118}$$

Treating the macroscopic delta function as an analytic function, we may expand  $\delta(\mathbf{r} - \mathbf{r}_j)$  as a Taylor series about  $\delta(\mathbf{r} - \mathbf{r}_i)$ . This gives

$$\begin{aligned}
\delta(\mathbf{r} - \mathbf{r}_j) &= \delta(\mathbf{r} - \mathbf{r}_i) + \mathbf{r}_{ij} \cdot \frac{\partial}{\partial \mathbf{r}_i} \delta(\mathbf{r} - \mathbf{r}_i) + \frac{1}{2!} \mathbf{r}_{ij} \mathbf{r}_{ij} \cdot \frac{\partial^2}{\partial \mathbf{r}_i^2} \delta(\mathbf{r} - \mathbf{r}_i) + \dots \\
&= \delta(\mathbf{r} - \mathbf{r}_i) - \mathbf{r}_{ij} \cdot \frac{\partial}{\partial \mathbf{r}} \delta(\mathbf{r} - \mathbf{r}_i) + \frac{1}{2!} \mathbf{r}_{ij} \mathbf{r}_{ij} \cdot \frac{\partial^2}{\partial \mathbf{r}^2} \delta(\mathbf{r} - \mathbf{r}_i) - \dots
\end{aligned} \tag{3.119}$$

Thus the difference between the two delta functions is

$$\begin{aligned}
\delta(\mathbf{r} - \mathbf{r}_i) - \delta(\mathbf{r} - \mathbf{r}_j) &= \mathbf{r}_{ij} \cdot \frac{\partial}{\partial \mathbf{r}} \delta(\mathbf{r} - \mathbf{r}_i) - \frac{1}{2!} \mathbf{r}_{ij} \mathbf{r}_{ij} \cdot \frac{\partial^2}{\partial \mathbf{r}^2} \delta(\mathbf{r} - \mathbf{r}_i) + \dots \\
&= \frac{\partial}{\partial \mathbf{r}} \cdot \mathbf{r}_{ij} O_{ij} \delta(\mathbf{r} - \mathbf{r}_i)
\end{aligned} \tag{3.120}$$

where the operator  $O_{ij}$  is given by,

$$O_{ij} = 1 - \frac{1}{2!} \mathbf{r}_{ij} \cdot \frac{\partial}{\partial \mathbf{r}} + \dots + \frac{1}{n!} \left( -\mathbf{r}_{ij} \cdot \frac{\partial}{\partial \mathbf{r}} \right)^{n-1} + \dots \tag{3.121}$$

Using this equation for the difference of the two delta functions  $\delta(\mathbf{r} - \mathbf{r}_i)$  and  $\delta(\mathbf{r} - \mathbf{r}_j)$  leads to

$$\begin{aligned}
&\frac{\partial}{\partial t} [\rho(\mathbf{r}, t) \mathbf{u}(\mathbf{r}, t)] + \frac{\partial}{\partial \mathbf{r}} \cdot [\rho(\mathbf{r}, t) \mathbf{u}(\mathbf{r}, t) \mathbf{u}(\mathbf{r}, t)] \\
&= -\frac{\partial}{\partial \mathbf{r}} \cdot \left[ \frac{1}{2} \int d\Gamma f(\Gamma, t) \sum_{i,j} \mathbf{r}_{ij} O_{ij} \delta(\mathbf{r} - \mathbf{r}_i) \frac{\partial \phi_{ij}}{\partial \mathbf{r}_i} + \left\langle \sum_i m_i (\mathbf{v}_i - \mathbf{u}(\mathbf{r}, t)) (\mathbf{v}_i - \mathbf{u}(\mathbf{r}, t)) \right|_{\mathbf{r}_i(t)=\mathbf{r}} \right]
\end{aligned} \tag{3.122}$$

Comparing this equation with the momentum conservation equation (2.12) we see that the pressure tensor is,

$$\mathbf{P}(\mathbf{r}, t) = \left\langle \sum_i m_i (\mathbf{v}_i - \mathbf{u}(\mathbf{r}, t)) (\mathbf{v}_i - \mathbf{u}(\mathbf{r}, t)) - \frac{1}{2} \sum_{i,j} \mathbf{r}_{ij} O_{ij} \delta(\mathbf{r} - \mathbf{r}_i) \mathbf{F}_{ij} \right|_{\mathbf{r}_i(t)=\mathbf{r}} \right\rangle \tag{3.123}$$

where  $\mathbf{F}_{ij} = -\partial \phi_{ij} / \partial \mathbf{r}_i$  is the force on particle  $i$  due to particle  $j$ .

We will now use the same technique to calculate the microscopic expression for the heat flux vector. The partial time derivative of the energy density is (from equation (3.112))

$$\begin{aligned}
\frac{\partial}{\partial t} [\rho(\mathbf{r}, t) e(\mathbf{r}, t)] &= \int d\Gamma \frac{\partial f(\Gamma, t)}{\partial t} \left( \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_i \phi_{ij} \right) \delta(\mathbf{r} - \mathbf{r}_i) \\
&= \int d\Gamma f(\Gamma, t) \left\{ iL \left( \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_i \phi_{ij} \right) \right\} \delta(\mathbf{r} - \mathbf{r}_i) \\
&\quad + \int d\Gamma f(\Gamma, t) \left( \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_i \phi_{ij} \right) iL \delta(\mathbf{r} - \mathbf{r}_i)
\end{aligned}$$

$$\begin{aligned}
 &= \int d\Gamma f(\Gamma, t) \left\{ \sum_i \mathbf{v}_i \cdot \mathbf{F}_i - \frac{1}{2} \sum_{i,j} (\mathbf{v}_i \cdot \mathbf{F}_{ij} + \mathbf{v}_j \cdot \mathbf{F}_{ji}) \right\} \delta(\mathbf{r} - \mathbf{r}_i) \\
 &+ \int d\Gamma f(\Gamma, t) \left( \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_{i,j} \phi_{ij} \right) \mathbf{v}_i \cdot \frac{\partial \delta(\mathbf{r} - \mathbf{r}_i)}{\partial \mathbf{r}_i} \\
 &= \int d\Gamma f(\Gamma, t) \frac{1}{2} \sum_{i,j} \mathbf{v}_i \cdot \mathbf{F}_{ij} [\delta(\mathbf{r} - \mathbf{r}_i) - \delta(\mathbf{r} - \mathbf{r}_j)] \\
 &- \int d\Gamma f(\Gamma, t) \frac{\partial}{\partial \mathbf{r}} \cdot \left( \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_{i,j} \phi_{ij} \right) \mathbf{v}_i \delta(\mathbf{r} - \mathbf{r}_i)
 \end{aligned} \tag{3.124}$$

In the second term, the gradient operator  $\partial/\partial \mathbf{r}$  is contracted into  $\mathbf{v}_i$ . Using our previous result for the difference of two delta functions, equation (3.120), gives

$$\begin{aligned}
 &\frac{\partial}{\partial t} [\rho(\mathbf{r}, t) e(\mathbf{r}, t)] \\
 &= - \frac{\partial}{\partial \mathbf{r}} \cdot \int d\Gamma f(\Gamma, t) \left[ \frac{1}{2} \sum_{i,j} \mathbf{r}_{ij} O_{ij} \mathbf{v}_i \cdot \mathbf{F}_{ij} - \left( \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_{i,j} \phi_{ij} \right) \mathbf{v}_i \right] \delta(\mathbf{r} - \mathbf{r}_i)
 \end{aligned} \tag{3.125}$$

From equation (2.24) we conclude that,

$$\begin{aligned}
 &\mathbf{J}_Q(\mathbf{r}, t) + \mathbf{P}(\mathbf{r}, t) \cdot \mathbf{u}(\mathbf{r}, t) + \rho(\mathbf{r}, t) e(\mathbf{r}, t) \mathbf{u}(\mathbf{r}, t) \\
 &= \int d\Gamma f(\Gamma, t) \left[ - \sum_{i,j} \mathbf{r}_{ij} O_{ij} \mathbf{v}_i \cdot \mathbf{F}_{ij} + \left( \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_{i,j} \phi_{ij} \right) \mathbf{v}_i \right] \delta(\mathbf{r} - \mathbf{r}_i)
 \end{aligned} \tag{3.126}$$

Now the definition of the energy density, equation (3.112) gives

$$\rho(\mathbf{r}, t) e(\mathbf{r}, t) \mathbf{u}(\mathbf{r}, t) = \int d\Gamma f(\Gamma, t) \left( \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_{i,j} \phi_{ij} \right) \delta(\mathbf{r} - \mathbf{r}_i) \mathbf{u}(\mathbf{r}, t) \tag{3.127}$$

so that,

$$\begin{aligned}
 &\mathbf{J}_Q(\mathbf{r}, t) + \mathbf{P}(\mathbf{r}, t) \cdot \mathbf{u}(\mathbf{r}, t) \\
 &= \int d\Gamma f(\Gamma, t) \left[ - \frac{1}{2} \sum_{i,j} \mathbf{r}_{ij} O_{ij} \mathbf{v}_i \cdot \mathbf{F}_{ij} + \left( \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_{i,j} \phi_{ij} \right) (\mathbf{v}_i - \mathbf{u}(\mathbf{r}, t)) \right] \delta(\mathbf{r} - \mathbf{r}_i)
 \end{aligned} \tag{3.128}$$

Similarly, from the definition of the pressure tensor  $\mathbf{P}(\mathbf{r}, t)$  (see equation (3.123)), we know that

$$\mathbf{P}(\mathbf{r}, t) \cdot \mathbf{u}(\mathbf{r}) = \int d\Gamma f(\Gamma, t) \left\{ \sum_i m_i (\mathbf{v}_i - \mathbf{u})(\mathbf{v}_i - \mathbf{u}) - \frac{1}{2} \sum_j \mathbf{r}_{ij} O_{ij} \mathbf{F}_{ij} \right\} \cdot \mathbf{u} \delta(\mathbf{r} - \mathbf{r}_i) \tag{3.129}$$

thus we identify the heat flux vector as,



$$\begin{aligned}
\mathbf{J}_Q(\mathbf{r}, t) &= \int d\Gamma f(\Gamma, t) \left[ \left( \frac{1}{2} \sum_i m \mathbf{v}_i^2 + \frac{1}{2} \sum_{i,j} \phi_{ij} \right) (\mathbf{v}_i - \mathbf{u}) - \sum_i m (\mathbf{v}_i - \mathbf{u}) (\mathbf{v}_i - \mathbf{u}) \cdot \mathbf{u} \right. \\
&\quad \left. - \frac{1}{2} \sum_{i,j} \mathbf{r}_{ij} (\mathbf{v}_i - \mathbf{u}) \cdot \mathbf{F}_{ij} O_{ij} \right] \delta(\mathbf{r} - \mathbf{r}_i) \\
&= \int d\Gamma f(\Gamma, t) \left[ \frac{1}{2} \sum_i m \{ (\mathbf{v}_i - \mathbf{u})^2 + \mathbf{u}^2 \} (\mathbf{v}_i - \mathbf{u}) + \frac{1}{2} \sum_{i,j} \phi_{ij} (\mathbf{v}_i - \mathbf{u}) \right. \\
&\quad \left. - \frac{1}{2} \sum_{i,j} \mathbf{r}_{ij} (\mathbf{v}_i - \mathbf{u}) \cdot \mathbf{F}_{ij} O_{ij} \right] \delta(\mathbf{r} - \mathbf{r}_i)
\end{aligned} \tag{3.130}$$

From the definitions of the mass density and momentum density (equations (3.110) and (3.111)) we find that

$$\int d\Gamma f(\Gamma, t) \sum_i (\mathbf{v}_i - \mathbf{u}) m \mathbf{u}^2 \delta(\mathbf{r} - \mathbf{r}_i) = 0 \tag{3.131}$$

so there is no contribution from the  $\mathbf{u}^2$  term. Further, if we define the peculiar energy of particle  $i$  to be

$$e_i = \frac{1}{2} m_i (\mathbf{v}_i - \mathbf{u})^2 + \frac{1}{2} \sum_j \phi_{ij} \tag{3.132}$$

then the heat flux vector can be written as

$$\mathbf{J}_Q(\mathbf{r}, t) = \int d\Gamma f(\Gamma, t) \left[ \sum_i (\mathbf{v}_i - \mathbf{u}) e_i - \frac{1}{2} \sum_{i,j} \mathbf{r}_{ij} (\mathbf{v}_i - \mathbf{u}) \cdot \mathbf{F}_{ij} O_{ij} \right] \delta(\mathbf{r} - \mathbf{r}_i)$$

or,

$$\mathbf{J}_Q(\mathbf{r}, t) = \left\langle \sum_i (\mathbf{v}_i - \mathbf{u}) e_i - \frac{1}{2} \sum_{i,j} \mathbf{r}_{ij} (\mathbf{v}_i - \mathbf{u}) \cdot \mathbf{F}_{ij} O_{ij} \right|_{\mathbf{r}_i(t) = \mathbf{r}} \rangle \tag{3.133}$$

### 3.8 Instantaneous Microscopic Representation of Fluxes

The Irving-Kirkwood procedure has given us microscopic expressions for the thermodynamic fluxes in terms of ensemble averages. At equilibrium in a uniform fluid, the Irving-Kirkwood expression for the pressure tensor is the same expression as that derived using Gibbs' ensemble theory for equilibrium statistical mechanics. If the fluid density is uniform in space, the  $O_{ij}$  operator appearing in the above expressions reduces to unity. This is easier to see if we calculate microscopic expressions for the fluxes in  $\mathbf{k}$ -space rather than real space. In the process we will better understand the nature of the Irving-Kirkwood expressions.

In this section we derive *instantaneous* expressions for the fluxes rather than the ensemble based, Irving-Kirkwood expressions. The reason for considering instantaneous expressions is two-fold. The fluxes are based upon conservation laws and these laws are valid instantaneously for every member of the ensemble. They do not require ensemble averaging to be true. Secondly, most computer simulation involves calculating system properties from a single system trajectory.

Ensemble averaging is almost never used because it is relatively expensive in computer time. The ergodic hypothesis, that the result obtained by ensemble averaging is equal to that obtained by time averaging the same property along a single phase space trajectory, implies that one should be able to develop expressions for the fluxes which do not require ensemble averaging. For this to be practically realisable it is clear that the mass, momentum and energy densities must be definable at each instant along the trajectory.

We define the Fourier transform pair by

$$f(\mathbf{k}) = \int d\mathbf{r} e^{i\mathbf{k}\cdot\mathbf{r}} f(\mathbf{r}) \quad f(\mathbf{r}) = \frac{1}{(2\pi)^3} \int d\mathbf{k} e^{-i\mathbf{k}\cdot\mathbf{r}} f(\mathbf{k}) \quad (3.134)$$

In the spirit of the Irving-Kirkwood procedure we define the instantaneous  $\mathbf{r}$ -space mass density to be,

$$\rho(\mathbf{r}, t) = \sum_{i=1}^N m \delta(\mathbf{r} - \mathbf{r}_i(t)) \quad (3.135)$$

where the explicit time dependence of  $\rho(\mathbf{r}, t)$  (that is the time dependence differentiated by the hydrodynamic derivative  $\partial/\partial t$ , with  $\mathbf{r}$  fixed) is through the time dependence of  $\mathbf{r}_i(t)$ . The  $\mathbf{k}$ -space instantaneous mass density is then

$$\rho(\mathbf{k}, t) = \int d\mathbf{r} \sum_{i=1}^N m \delta(\mathbf{r} - \mathbf{r}_i(t)) e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{i=1}^N m e^{i\mathbf{k}\cdot\mathbf{r}_i(t)} \quad (3.136)$$

We will usually allow the context to distinguish whether we are using ensemble averages or instantaneous expressions. The time dependence of the mass density is solely through the time dependence of  $\mathbf{r}_i$ , so that

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} = i\mathbf{k} \cdot \sum_{i=1}^N m \mathbf{v}_i e^{i\mathbf{k}\cdot\mathbf{r}_i(t)} \quad (3.137)$$

Comparing this with the Fourier transform of (2.4) (noting that  $d/dt|_{\mathbf{k}}$  in (3.137) corresponds to  $\partial/\partial t|_{\mathbf{r}}$  in (2.4)) we see that if we let  $\mathbf{J}(\mathbf{r}, t) = \rho(\mathbf{r}, t)\mathbf{u}(\mathbf{r}, t)$  then,

$$\mathbf{J}(\mathbf{k}, t) = \sum_{i=1}^N m \mathbf{v}_i e^{i\mathbf{k}\cdot\mathbf{r}_i(t)} \quad (3.138)$$

This equation is clearly the instantaneous analogue of the Fourier transform of the Irving-Kirkwood expression for the momentum density. There is no ensemble average required in (3.137). To look at the *instantaneous pressure tensor* we only need to differentiate equation (3.138) in time.

$$\frac{\partial}{\partial t} \mathbf{J}(\mathbf{k}, t) = \sum_{i=1}^N \left( i\mathbf{k} \cdot m \mathbf{v}_i(t) \mathbf{v}_i(t) e^{i\mathbf{k}\cdot\mathbf{r}_i(t)} + \mathbf{F}_i e^{i\mathbf{k}\cdot\mathbf{r}_i(t)} \right) \quad (3.139)$$

We can write the second term on the right hand side of this equation in the form of the Fourier transform of a divergence by noting that,

$$\begin{aligned}\sum_{i=1}^N \mathbf{F}_i e^{i\mathbf{k}\cdot\mathbf{r}_i} &= \frac{1}{2} \sum_{i,j=1}^N \left( \mathbf{F}_{ij} e^{i\mathbf{k}\cdot\mathbf{r}_i} + \mathbf{F}_{ji} e^{i\mathbf{k}\cdot\mathbf{r}_j} \right) = \frac{1}{2} \sum_{i,j=1}^N \mathbf{F}_{ij} \left( e^{i\mathbf{k}\cdot\mathbf{r}_i} - e^{i\mathbf{k}\cdot\mathbf{r}_j} \right) \\ &= \frac{1}{2} \sum_{i,j=1}^N \mathbf{F}_{ij} \left( e^{-i\mathbf{k}\cdot\mathbf{r}_{ij}} - 1 \right) e^{i\mathbf{k}\cdot\mathbf{r}_j} = -i\mathbf{k} \cdot \frac{1}{2} \sum_{i,j=1}^N \mathbf{r}_{ij} \mathbf{F}_{ij} \frac{e^{-i\mathbf{k}\cdot\mathbf{r}_{ij}} - 1}{-i\mathbf{k} \cdot \mathbf{r}_{ij}} e^{i\mathbf{k}\cdot\mathbf{r}_j}\end{aligned}\quad (3.140)$$

Combining (3.139) and (3.140) and performing an inverse Fourier transform we obtain the instantaneous analogue of equation (3.123). We could of course continue the analysis of §3.7 to remove the streaming contribution from the pressure tensor but this is more laborious in  $\mathbf{k}$ -space than in real space and we will not give this here. We can use our instantaneous expression for the pressure tensor to describe fluctuations in an equilibrium system. In this case the streaming velocity is of course zero, and

$$\mathbf{P}(\mathbf{k}, t) = \sum_{i=1}^N m_i \mathbf{v}_i(t) \mathbf{v}_i(t) e^{i\mathbf{k}\cdot\mathbf{r}_i(t)} - \frac{1}{2} \sum_{i,j} \mathbf{r}_{ij}(t) \mathbf{F}_{ij}(t) \frac{e^{-i\mathbf{k}\cdot\mathbf{r}_{ij}(t)} - 1}{-i\mathbf{k} \cdot \mathbf{r}_{ij}(t)} e^{i\mathbf{k}\cdot\mathbf{r}_j(t)} \quad (3.141)$$

The  $\mathbf{k}$ -space analysis given provided a better understanding of the Irving-Kirkwood operator  $O_{ij}$ . In  $\mathbf{k}$ -space it is not necessary to perform the apparently difficult operation of Taylor expanding delta functions.

Before we close this section we will try to make the equation for the momentum density,  $\mathbf{J}(\mathbf{r}, t) = \rho(\mathbf{r}, t) \mathbf{u}(\mathbf{r}, t)$ , a little clearer. In  $\mathbf{k}$ -space this equation is a convolution,

$$\mathbf{J}(\mathbf{k}, t) = \int d\mathbf{k}' \rho(\mathbf{k} - \mathbf{k}', t) \mathbf{u}(\mathbf{k}', t) \quad (3.142)$$

Does this definition of the streaming velocity  $\mathbf{u}$ , make good physical sense? One sensible definition for the streaming velocity  $\mathbf{u}$ , would be that velocity which minimises the sum of squares of deviations from the particle velocities  $\mathbf{v}_i$ . For simplicity we set  $t = 0$ , and let  $R$ , be that sum of squares,

$$R = \sum_{i=1}^N (\mathbf{v}_i - \mathbf{u}(\mathbf{r}_i))^2 = \sum_{i=1}^N (\mathbf{v}_i - \sum_n \mathbf{u}(\mathbf{k}_n) e^{-i\mathbf{k}_n \cdot \mathbf{r}_i})^2 \quad (3.143)$$

If  $\mathbf{u}(\mathbf{r})$  minimises this sum of squares then the derivative of  $R$  with respect to each of the Fourier components  $\mathbf{u}(\mathbf{k}_m)$ , must be zero. Differentiating (3.143) we obtain,

$$\frac{\partial R}{\partial \mathbf{u}(\mathbf{k}_m)} = 2 \sum_{i=1}^N (\mathbf{v}_i - \sum_n \mathbf{u}(\mathbf{k}_n) e^{-i\mathbf{k}_n \cdot \mathbf{r}_i}) e^{-i\mathbf{k}_m \cdot \mathbf{r}_i} = 0 \quad (3.144)$$

This implies that

$$\sum_{i=1}^N m \mathbf{v}_i e^{-i\mathbf{k}_m \cdot \mathbf{r}_i} = \sum_{i=1}^N \sum_n m \mathbf{u}(\mathbf{k}_n) e^{-i\mathbf{k}_n \cdot \mathbf{r}_i} e^{-i\mathbf{k}_m \cdot \mathbf{r}_i} \quad (3.145)$$

Both sides of this equation can be identified as  $\mathbf{k}$ -space variables,

$$\mathbf{J}(-\mathbf{k}_m) = \sum_n \mathbf{u}(\mathbf{k}_n) \rho(-\mathbf{k}_m - \mathbf{k}_n) \quad (3.146)$$

So that

$$\mathbf{J}(\mathbf{k}_m) = \sum_n \mathbf{u}(\mathbf{k}_n) \rho(\mathbf{k}_m - \mathbf{k}_n) \quad (3.147)$$

This is the Fourier series version of equation (3.142).

We can use the same procedure to calculate an expression for the heat flux vector. As we will see this procedure is very much simpler than the Irving-Kirkwood method described in §3.7. We begin by identifying the instantaneous expression for the instantaneous wavevector dependent energy density in a fluid at equilibrium,

$$\rho e(\mathbf{k}, t) = \sum_i \left( \frac{1}{2} m_i \dot{\mathbf{r}}_i^2(t) + \frac{1}{2} \sum_j \phi_{ij}(\mathbf{r}_{ij}(t)) \right) e^{i\mathbf{k} \cdot \mathbf{r}_i(t)} \quad (3.148)$$

This is instantaneous, wavevector dependent analogue of (3.112). To simplify notation in the following we will suppress the time argument for all phase variables. The time argument will always be  $t$ . If we calculate the rate of change of the energy density we find,

$$\begin{aligned} \frac{\partial}{\partial t} [\rho e(\mathbf{k}, t)] &= i\mathbf{k} \cdot \sum_i \dot{\mathbf{r}}_i(t) \left( \frac{1}{2} m_i \dot{\mathbf{r}}_i^2(t) + \frac{1}{2} \sum_j \phi_{ij}(\mathbf{r}_{ij}(t)) \right) e^{i\mathbf{k} \cdot \mathbf{r}_i(t)} + \sum_i m_i \dot{\mathbf{r}}_i \cdot \ddot{\mathbf{r}}_i e^{i\mathbf{k} \cdot \mathbf{r}_i(t)} \\ &\quad + \frac{1}{2} \sum_{i,j} \left( \dot{\mathbf{r}}_i \cdot \frac{\partial \phi_{ij}}{\partial \mathbf{r}_i} + \dot{\mathbf{r}}_j \cdot \frac{\partial \phi_{ij}}{\partial \mathbf{r}_j} \right) e^{i\mathbf{k} \cdot \mathbf{r}_i(t)} \end{aligned} \quad (3.149)$$

Where we use the notation  $\phi_i = \frac{1}{2} \sum_j \phi_{ij}$ . If we denote the energy of particle  $i$  as  $e_i$  and  $\mathbf{F}_{ij}$  as the force exerted on particle  $i$  due to  $j$  then (3.149) can be rewritten as,

$$\begin{aligned} \frac{\partial}{\partial t} [\rho e(\mathbf{k}, t)] &= i\mathbf{k} \cdot \sum_i \dot{\mathbf{r}}_i e_i e^{i\mathbf{k} \cdot \mathbf{r}_i} + \sum_i \dot{\mathbf{r}}_i \cdot \mathbf{F}_i e^{i\mathbf{k} \cdot \mathbf{r}_i} - \frac{1}{2} \sum_{i,j} (\dot{\mathbf{r}}_i \cdot \mathbf{F}_{ij} - \dot{\mathbf{r}}_j \cdot \mathbf{F}_{ij}) e^{i\mathbf{k} \cdot \mathbf{r}_i} \\ &= i\mathbf{k} \cdot \sum_i \dot{\mathbf{r}}_i e_i e^{i\mathbf{k} \cdot \mathbf{r}_i} + \frac{1}{2} \sum_{i,j} \dot{\mathbf{r}}_i \cdot \mathbf{F}_{ij} (e^{i\mathbf{k} \cdot \mathbf{r}_i} - e^{i\mathbf{k} \cdot \mathbf{r}_j}) \end{aligned} \quad (3.150)$$

This equation involves the same combination of exponents as we saw for the pressure tensor in (3.140). Expanding exponentials to first order in  $\mathbf{k}$ , and using equation (2.24) we find that the wavevector dependent heat flux vector can be written as

$$\mathbf{J}_Q(\mathbf{k} = 0) = \sum_i \dot{\mathbf{r}}_i e_i - \frac{1}{2} \sum_{i,j} \mathbf{r}_{ij} \mathbf{F}_{ij} \cdot \dot{\mathbf{r}}_i + O(k^2) \quad (3.151)$$

In  $\mathbf{r}$ -space rather than  $\mathbf{k}$ -space the expressions for the instantaneous pressure tensor and heat flux vector become,

$$\mathbf{P}(\mathbf{r}, t) = \sum_i \left[ m_i (\mathbf{v}_i(t) - \mathbf{u}(t)) (\mathbf{v}_i(t) - \mathbf{u}(t)) - \frac{1}{2} \sum_j \mathbf{r}_{ij}(t) \mathbf{F}_{ij}(t) \right] \delta(\mathbf{r} - \mathbf{r}_i(t)) \quad (3.152)$$

$$\mathbf{J}_Q(\mathbf{r}, t) = \sum_i \left[ (\mathbf{v}_i(t) - \mathbf{u}(t)) e_i(t) - \frac{1}{2} \sum_{i,j} \mathbf{r}_{ij}(t) (\mathbf{v}_i(t) - \mathbf{u}(t)) \cdot \mathbf{F}_{ij}(t) \right] \delta(\mathbf{r} - \mathbf{r}_i(t)) \quad (3.153)$$

Our procedure for calculating microscopic expressions for the hydrodynamic densities and fluxes relies upon establishing a correspondence between the microscopic and macroscopic forms of the continuity equations. These equations refer only to the divergence of the pressure tensor and heat flux. Strictly speaking therefore we can only determine the divergences of the flux tensors. We can add any divergence free quantity to our expressions for the flux tensors without affecting the identification process.

### 3.9 The Kinetic Temperature

We obtain an instantaneous expression for the temperature by analysing the expression for the pressure tensor (3.150) for the case of an ideal gas at equilibrium. Thus if  $n(\mathbf{r}, t)$  is the local instantaneous number density,

$$\frac{3}{2} n(\mathbf{r}, t) k_B T(\mathbf{r}, t) = \sum_{i=1}^N \frac{1}{2} m_i (\mathbf{v}_i(t) - \mathbf{u}(\mathbf{r}, t))^2 \delta(\mathbf{r}_i(t) - \mathbf{r}) \quad (3.154)$$

We will call this expression for the temperature, the kinetic temperature. In using this expression for the temperature we are employing a number of approximations. Firstly we are ignoring the number of degrees of freedom which are frozen by the instantaneous determination of  $\mathbf{u}(\mathbf{r}, t)$ . Secondly, and more importantly, we are assuming that in a nonequilibrium system the kinetic temperature is identical to the thermodynamic temperature  $T_T$ ,

$$T_T = \left. \frac{\partial E}{\partial S} \right|_{N,V} \quad (3.155)$$

This is undoubtedly an approximation. It would be true if the postulate of local thermodynamic equilibrium was exact. However we know that the energy, pressure, enthalpy etc. are all functions of the thermodynamic forces driving the system away from equilibrium. These are nonlinear effects which vanish in Newtonian fluids. Presumably the entropy is also a function of these driving forces. It is extremely unlikely that the field dependence of the entropy and the energy are precisely those required for the exact equivalence of the kinetic and

thermodynamic temperatures for all nonequilibrium systems. Recent calculations of the entropy of systems very far from equilibrium support the hypothesis that the kinetic and thermodynamic temperatures are in fact different (Evans, 1989). Outside the linear (Newtonian), regime the kinetic temperature is a convenient operational (as opposed to thermodynamic) state variable. If a nonequilibrium system is in a steady state both the kinetic and the thermodynamic temperatures must be constant in time. Furthermore we expect that outside the linear regime in systems with a unique nonequilibrium steady state, that the thermodynamic temperature should be a monotonic function of the kinetic temperature.

## References

- Berne, B. J., (1977). Chapter 5 in: "Statistical Mechanics Part B: Time Dependent Processes", ed (B. J. Berne ed.). Plenum Press, New York.
- Dyson, F.J. (1949). *Phys. Rev.* 75, 486; Feynman, R.P. (1951). *ibid.* 84, 108.
- Eckmann, J.-P. and Ruelle, D., (1985), *Rev. Mod. Phys.*, 57, 617.
- Eddberg, Evans and Morriss, 1986
- Evans, D.J., (1989). *J. Stat. Phys.*, Cohen Festschrift (to appear).
- Evans, D. J., Hoover, W. G., Failor, B. H., Moran, B. and Ladd, A. J. C., (1983). *Phys. Rev.* A28, 1016.
- Evans, D.J. and Morriss, G.P., (1984), *Chem. Phys.*, 87, 451.
- Evans, D.J. and Morriss, G.P. (1984). *Chem. Phys.*, 87, 451.
- Evans, D.J. and Morriss, G.P., (1984). *J. Chem. Phys.*, 81, 1984.
- Fermi, E., Pasta, J.G. and Ulam, S.M., (1955). *LASL Report LA-1940*, Los Alamos Science Laboratory New Mexico. see also "Collected Works of Enrico Fermi", 2,978, University of Chicago Press.
- Gauss, K. F. (1829). *J. Reine Angew. Math.*, IV, 232.
- Goldstein, H. (1980). "Classical Mechanics" (2<sup>nd</sup> Edition), Addison-Wesley, Reading MA.
- Holian, B.L., Hoover, W.G. and Posch, H.A. (1987). *Phys. Rev. Letts.*, 59, 10.
- Hoover, W. G., (1986). "Molecular Dynamics", Lecture Notes in Physics 258, Springer
- Irving, J. H. and Kirkwood, J. G., (1950). *J. Chem. Phys.*, 18, 817.
- Magnus, W. (1954). *Commun. Pure Appl. Math.* 7, 649.
- Oseledec, V.I., (1968), *Trudy Mosk. Mat. Obsc.* 19, 179 [Moscow Math. Soc.].
- Pars, L. A. (1979). "A Treatise on Analytical Dynamics", Ox Bow, Woodbridge Conn.

Pechukas, P. and Light, J.C. (1966). *J. Chem. Phys.* 44, 3897.

Whittaker, E.T., (1904). "A Treatise on Analytical Dynamics of Particles and Rigid Bodies", Macmillan.