## CHAPTER 2.

# Introduction to time reversible, thermostatted dynamical systems and statistical mechanical ensembles

"I have found it convenient, instead of considering one system of material particles, to consider a large number of systems similar to each other in all respects, except the initial circumstances of the motion, which are supposed to vary from system to system, the total energy being the same in all."
Maxwell 1879

### 2.1 TIME REVERSIBLITY IN DYNAMICAL SYSTEMS

Consider an isolated Hamiltonian system of interacting particles. The microscopic state of the system is represented by a phase space vector of the coordinates and canonical momenta of all the particles, in an exceedingly high dimensional space - phase space - $\{\mathbf{q}_1,\mathbf{q}_2,..\mathbf{q}_N,\mathbf{p}_1,..\mathbf{p}_N\} \equiv (\mathbf{q},\mathbf{p}) \equiv \mathbf{\Gamma}$ where $\mathbf{q}_i,\mathbf{p}_i$ are the position and conjugate momentum of particle, $i$. The equations of motion for the system with an autonomous Hamiltonian $H(\mathbf{q},\mathbf{p})$, are,

$$\begin{aligned}\dot{\mathbf{q}}_i &= \frac{\partial H(\mathbf{q},\mathbf{p})}{\partial \mathbf{p}_i} \\ \dot{\mathbf{p}}_i &= -\frac{\partial H(\mathbf{q},\mathbf{p})}{\partial \mathbf{q}_i}\end{aligned} \quad . \qquad (2.1.1)$$

**Definition**

We define the *time reversal mapping*, $M^T$ …, as

$$M^T\mathbf{\Gamma} \equiv M^T(\mathbf{q},\mathbf{p}) \equiv (\mathbf{q},-\mathbf{p}) . \qquad (2.1.2)$$

Then because the Hamiltonian for a system of interacting point particles, $H(\mathbf{q},\mathbf{p}) = \sum p_i^2 / 2m + \Phi(\mathbf{q})$, is even in the momenta, we see from the equations of motion that

$$\dot{\mathbf{\Gamma}} \equiv iL(\mathbf{\Gamma})\mathbf{\Gamma} = (\frac{\mathbf{p}}{m} \bullet \frac{\partial}{\partial \mathbf{q}}, -\frac{\partial \Phi}{\partial \mathbf{q}} \bullet \frac{\partial}{\partial \mathbf{p}}) \bullet (\mathbf{q},\mathbf{p}) = (\mathbf{p} / m, -\boldsymbol{\partial}\Phi / \boldsymbol{\partial}\mathbf{q}) \qquad (2.1.3)$$

where $iL... \equiv \dot{\mathbf{\Gamma}} \bullet \partial .../\partial \mathbf{\Gamma}$, is the time derivative operator for phase functions and $\dot{\mathbf{\Gamma}}$ is given by the equations of motion (*e.g.* (2.1.1)).

**Definition**

We refer to $L$ as the *p* or *phase-Liouvillean*.

The formal solution of (2.1.3) is

$$S^t\mathbf{\Gamma} \equiv \exp[iL(\mathbf{\Gamma})t]\mathbf{\Gamma} , \qquad (2.1.4)$$

and

$$\begin{aligned} dS^t\mathbf{\Gamma} / dt &= iL(\mathbf{\Gamma})\exp[iL(\mathbf{\Gamma})t]\mathbf{\Gamma} = iL(\mathbf{\Gamma})S^t\mathbf{\Gamma} \\ &= \exp[iL(\mathbf{\Gamma})t]iL(\mathbf{\Gamma})\mathbf{\Gamma} = S^t\dot{\mathbf{\Gamma}} \end{aligned} \qquad (2.1.5)$$

If we apply the time reversal map to the *p*-Liouvillean we see that,

$$\begin{aligned} M^T iL(\mathbf{\Gamma}) &\equiv M^T\dot{\mathbf{\Gamma}} \bullet \frac{\partial}{\partial \mathbf{\Gamma}} = \mathrm{M}^T(\frac{\mathbf{p}}{m} \bullet \frac{\partial}{\partial \mathbf{q}}, -\frac{\partial \Phi}{\partial \mathbf{q}} \bullet \frac{\partial}{\partial \mathbf{p}}) \\ &= (\frac{-\mathbf{p}}{m} \bullet \frac{\partial}{\partial \mathbf{q}}, -\frac{\partial \Phi}{\partial \mathbf{q}} \bullet \frac{\partial}{\partial - \mathbf{p}})\mathrm{M}^T \\ &= iL(\mathbf{\Gamma}^T)M^T = -iL(\mathbf{\Gamma})M^T \end{aligned} \qquad (2.1.6)$$

Using this result we can apply the time reversal map to a propagated phase,

$$\begin{aligned} M^T(S^t\mathbf{\Gamma}) &= \exp[-iL(\mathbf{\Gamma})t]M^T(\mathbf{\Gamma}) \\ &= \exp[iL(\mathbf{\Gamma}^T)t]M^T(\mathbf{\Gamma})\ . \\ &= S^t\mathbf{\Gamma}^T \end{aligned} \tag{2.1.7}$$

**Definition**

*Time reversible dynamics* satisfies the equation:

$$M^T\exp(iL(\mathbf{\Gamma})t)M^T\exp(iL(\mathbf{\Gamma})t)\mathbf{\Gamma} = \mathbf{\Gamma} \tag{2.1.8}$$

To prove this we can use successive applications of (2.1.6). We note:

$$\begin{aligned} &M^T\exp(iL(\mathbf{\Gamma})t)M^T\exp(iL(\mathbf{\Gamma})t)\mathbf{\Gamma} \\ &= M^T\exp(iL(\mathbf{\Gamma})t)\exp(-iL(\mathbf{\Gamma})t)\mathbf{\Gamma}^T \\ &= M^T\mathbf{\Gamma}^T = \mathbf{\Gamma} \end{aligned} \tag{2.1.9}$$

We will say in words what time reversibility entails. If we start at a point in phase space evolve that phase forward in time an amount *t*; reverse the signs of all the momenta leaving the coordinates fixed; go forward in time using the same equations of motion for a duration *t* and finally reverse all the moment once again, then we end up at the same point in phase space where we originally started.

We will add a few remarks about notation. The $S^t$ notation for the propagator hides much subtlety. We will only use it for the simplest problems. To compare the two notations we note that

$$\begin{aligned} M^T(e^{iL(\mathbf{\Gamma})t}\mathbf{\Gamma}) &\equiv M^T S^t\mathbf{\Gamma} = e^{-iL(\mathbf{\Gamma})t}M^T\mathbf{\Gamma} = e^{iL(\mathbf{\Gamma}^T)t}M^T\mathbf{\Gamma} \\ &= e^{iL(\mathbf{\Gamma}^T)t}\mathbf{\Gamma}^T \qquad . \\ &\equiv S^t\mathbf{\Gamma}^T \end{aligned} \tag{2.1.10}$$

The convention is that the $S^t$ notation absorbs into the p-propagator, the phase to its right. There is no way to express $e^{-iL(\boldsymbol{\Gamma})t}\boldsymbol{\Gamma}^T$ in the $S^t$ notation.

If we take the solar system and reverse all the momenta and angular momenta of the planets then the resulting dynamics is also a solution of the equations of motion. If you watch a movie of the planets going around the sum and then play that movie backwards, the resulting motion is still a solution of Hamilton's equations of motion.

However, if we do the same to the motion of a waterfall or a jet aircraft taking off, although the time reversed movie is in fact still a solution of the dynamical equations of motion, the time reversed movie of a waterfall violates the Second "Law" of Thermodynamics. The time reversed movie of the jet plane would constitute a perpetual motion machine of the second kind, thereby also violating the Second Law of Thermodynamics. This is the so-called time irreversibility paradox first pointed out by Kelvin and later by Loschmidt. The resolution of this paradox forms one of the main themes of this book.

If we watch the time reversed movie of the solar system, then observing the time reversed planetary orbits we would not see anything that would violate the Second "Law" of thermodynamics. Thus the Second "Law" of Thermodynamics is somehow coupled to the complexity of the system. Maxwell was the first to realise this point (see the quote for Chapter 3). As we will see, the Fluctuation Theorem proved in Chapter 3 will resolve these apparent paradoxes. In the process the Fluctuation Theorem obviates the need for the Second "Law" of Thermodynamics.

## 2.2 INTRODUCTION TO TIME REVERSIBLE, THERMOSTATTED DYNAMICAL SYSTEMS

Most systems of thermodynamic interest are not isolated. The vast majority of engines, devices and all biological organelles exchange heat back and forth with their surroundings. These surroundings can usually be viewed as being vastly larger than the engine, or the system of interest.

Think of the operation of an automobile engine. Ultimately the chemical energy in the fuel is on average dissipated as heat to the surrounding atmosphere and earth. Because of the huge size ratio of the surroundings to the automobile, we can regard the surroundings as being unperturbed by the operation of the automobile. In principle the size of the surroundings can be expanded virtually without limit. So sufficiently far away from the system of interest we can regard the surroundings as being unperturbed by the system of interest and as we shall see these unchanging surroundings can be regarded as being in a state of thermodynamic equilibrium. Later we will learn how to treat cases where the surroundings are the same size as the system of interest and therefore may also be out of equilibrium but for the moment we will consider a nonequilibrium system of interest in contact with a much larger equilibrium reservoir.

Because in Chapter 6 we will derive the Relaxation Theorem and through it the form of the equilibrium distribution of microstates for an N-particle system, we cannot employ any knowledge of that equilibrium distribution prior to Chapter 6.

We assume that classical mechanics gives an adequate description of the dynamics. We assume that the total momentum of, the system of interest and separately, of the surroundings is zero. These systems are not in motion relative to each other or the observer. The microscopic state of the system is represented by a phase space vector of the coordinates and momenta of all the particles in an exceedingly high dimensional space - phase space - $\{\mathbf{q}_1,\mathbf{q}_2,..\mathbf{q}_N,\mathbf{p}_1,..\mathbf{p}_N\} \equiv (\mathbf{q},\mathbf{p}) \equiv \mathbf{\Gamma}$ where $\mathbf{q}_i,\mathbf{p}_i$ are the position and conjugate momentum of particle, $i$.

Experimentally we can only control a small number of variables that specify the *macroscopic* state of the system. We might only be able to control the system energy, the average kinetic energy of all or some of the particles, the volume $V$ and the number of atoms in the system, $N$, which we assume to be constant. There is therefore an enormous range of microstates that are consistent with the small number of macroscopic constraints.

In writing the microscopic equations of motion for the system it will be convenient to decompose the total system into two subsystems: the system of interest and the surroundings. As we have said above, the surroundings may be regarded as being unperturbed by the system of interest. Conversely, provided the surroundings are not moving with respect to either the observer or the system of interest and provided they have an unchanging distribution of states, the precise details of the microscopic equations of motion or indeed the nature of the particles that constitute those surrounding systems, have no impact on the system of interest. The surrounding particles are too small and too far from the system of interest.

The operation of an automobile is unaffected by the microscopic details of the road and atmosphere. Only a few macroscopic properties are important: average chemical composition, temperature and pressure of the air *etc, etc*.

A typical experiment of interest is conveniently summarised by the following example. Consider an electrical conductor (a molten salt for example) subject, at say $t = 0$, to an applied electric field, $\mathbf{E}$. We wish to understand the behaviour of this system from an atomic or molecular point of view. As in a laboratory, the molten salt is contained in a solid electrically insulating conduction cell and this cell is allowed to exchange heat with the much larger surroundings so that the average kinetic energy of the particles is constant.

If we use purely Hamiltonian equations of motion the entire system will eventually heat up. We will need to supplement the Hamiltonian equations of motion with some time reversible non-Hamiltonian terms buried deep in the surroundings so that a true nonequilibrium steady state is possible. The work that is done on the system is on average converted into heat which is conducted through the system of interest and the walls eventually being removed on average by these non Hamiltonian terms in the remote boundaries. Because these non-Hamiltonian terms are physicaly remote from the system of interest there is no way that the system of interest can "know" how the heat is eventually removed.

The first time-reversible, deterministic thermostats and ergostats were invented simultaneously but independently by Hoover and Evans in the early 1980's.

Prior to this development there was no satisfactory mathematical way of modelled thermostatted, nonequilibrium steady states.

We could study the macroscopic behaviour of the macroscopic system by taking just one of the huge number of microstates that satisfy the macroscopic conditions, and then solving the equations of motion for this single microscopic trajectory. We could then compute time averages, $\overline{A}_t$, of a phase function, $A(\mathbf{\Gamma})$:

$$\overline{A}_t \equiv \frac{1}{t}\int_0^t ds\, A(S^s\mathbf{\Gamma}). \qquad (2.2.1)$$

However, we would have to take care that our microscopic trajectory $S^t\mathbf{\Gamma}$, was a <u>*typical*</u> trajectory and that it did not behave in an exceptional way. Perhaps a better way of understanding the macroscopic system would be to select a set of $\mathrm{N}_\Gamma$ initial phases (*i.e.* microstates) $\{\mathbf{\Gamma}_j, j=1,...N_\Gamma\}$ distributed according to the naturally occurring states that are consistent with the small number of macro-constraints and compute the time dependent properties of the macroscopic system by taking a time dependent *ensemble* average $\langle \mathrm{A(t)} \rangle$ of a phase function $\mathrm{A}(\mathbf{\Gamma})$ over the <u>*ensemble*</u> of time evolved phases $\langle A(t)\rangle = \lim_{N_\Gamma\to\infty}\sum_{j=1}^{N_\Gamma} A(S^t\mathbf{\Gamma}_j)/N_\Gamma$. Indeed repeating the experiment with initial states that are consistent with the specified initial conditions is often what an experimentalist attempts to do in the laboratory. One could then try to compute the initial phase space density normalized to unity $f(\mathbf{\Gamma};0)$ and try to understand the time dependent evolution of this density $f(\mathbf{\Gamma};t)$. Time dependent averages of phase functions could then be computed a time dependent ensemble average:

$$\langle A(t)\rangle = \int d\mathbf{\Gamma}\, A(\mathbf{\Gamma})f(\mathbf{\Gamma};t) = \int d\mathbf{\Gamma}\, A(S^t\mathbf{\Gamma})f(\mathbf{\Gamma};0). \qquad (2.2.2)$$

The equality of the average over the initial distribution with that taken over the time dependent distribution is guaranteed by the fact that the normalization of the distribution function is always unity, independent of time. Although the concept of ensemble averaging seems natural and intuitive to experimental scientists, the use of ensembles has caused some problems and misunderstandings from a more purely mathematical viewpoint.

**Definition**

A system is said to be *time stationary* if the time averages appearing in (2.2.2) become independent of time either for all times or in the long time limit.

**Definition**

A stationary system is said to be *ergodic* if the time average of a reasonably smooth phase function along a trajectory that starts almost any where in the ostensible phase space, is equal to the ensemble average taken over an ensemble of systems consistent with the small number of macroscopic constraints on the system.

$$\left\langle A \right\rangle = \lim_{t\to\infty} \overline{A}_t = \frac{1}{t}\int_0^t ds\; A(S^s\mathbf{\Gamma}), \quad \forall \mathbf{\Gamma}(0) \tag{2.2.3}$$

Experience shows that for an isolated Hamiltonian system of interacting particles with no applied dissipative fields, the system will usually relax to a time stationary state where time averages of almost all macroscopic variables such and pressure or density become time independent. That state is called the state of microcanonical equilibrium. Similarly if a Hamiltonian system free of applied dissipative fields (like electric field for electrically conductive systems) which is allowed to exchange heat with a vastly larger heat bath which itself can be considered to be at equilibrium, then at long times the Hamiltonian system will be expected to relax to the canonical equilibrium state. Later, in Chapter 5 of this book we will (subject to some fairly simply stated mathematical conditions) prove the Equilibrium Relaxation Theorem which shows that initial nonequilibrium systems will at long times relax, perhaps non-monotonically to an ergodic state of thermodynamic equilibrium. That same Theorem also gives precise mathematical expressions for the equilibrium phase space distributions, both canonical and microcanonical.

Ensembles are well known to equilibrium statistical mechanics, the concept being first introduced by Maxwell in 1879. The use of ensembles in nonequilibrium statistical mechanics is less widely known and understood.[1] For our experiment it will often be convenient to choose the initial ensemble that is represented by the set of

---

[1] For further background information on nonequilibrium statistical mechanics see [16]

phases $\{\mathbf{\Gamma}_j, j = 1,...N_\Gamma\}$, to be one of the standard ensembles of equilibrium statistical mechanics. However, sometimes we may wish to vary this somewhat. In any case, in all the examples we will consider, the initial ensemble of phase vectors will be characterised by a known initial N-particle distribution function, $f(\mathbf{\Gamma};0)$, which gives the probability, $f(\mathbf{\Gamma};0)d\mathbf{\Gamma}$, that a member of the ensemble is within some infinitesmal neighbourhood $\mathrm{d}\mathbf{\Gamma}$ of a phase $\mathbf{\Gamma}$ at time 0, when the experiment began. By construction the number of ensemble members is conserved.

Consider an electric field that on average, does work on an electrically conducting system causing an electric current, $\mathbf{I} \equiv \sum c_i \dot{\mathbf{q}}_i$, to flow ($c_i$ is the electric charge on particle *i*). To remove the complicating effects of space charge build up or surface electrolysis we employ periodic boundary conditions in the direction of the electric filed. This allows the current to flow forever and also allows for the possibility of establishing a nonequilibrium time stationary, or steady, state.

It is exceedingly important to remember that we are *expressly* excluding the case where the system is an insulator and the field induces a polarization rather than a current! The difference between an insulator and a conductor can *only* be determined by the physics of the situation. If we subject sodium chloride to an electric field at room temperature then the field induces a polarization that changes the internal energy of the system. Electrostatic potential energy is stored in the system. If we make the single change of increasing the temperature of sodium chloride to 1000K then sodium chloride melts and becomes an electrical conductor. The electric field does not change the internal energy of the system. So the difference between an insulator and a conductor cannot be determined from the equations of motion (*i.e.* from the Hamiltonian)! The difference can be in the initial conditions for the equations of motion – in this case the initial energy or temperature. We will treat the case where fields change the internal energy of the system in Chapter 8. Until then external fields are assumed to only be dissipative, and do not change the internal energy of systems.

### Definition

If external fields are applied to the system of particles and the external field does work on the system and if that work can be turned *completely* into heat that then can diffuse out of the system, the external field is termed a *dissipative field.* If the work can be completely stored in the system in the form of potential energy, the external field is termed *nondissipative*.

An example of the latter could be to slowly strain an elastic solid. The work is stored in the intermolecular potential energy of the solid's constituent molecules. If you change the temperature of a system so that a system undergoes a phase change from a solid to a liquid, the same strain field can change from being nondissipative to being dissipative.

We expect that at an arbitrary time $t$, after the field has been applied, the ensemble averaged electric current $\langle \mathbf{I}(t) \rangle$ will be in the direction of the field; that work performed on the system by the field will be transformed into (or dissipated as) heat – Ohmic heating, $\langle \mathbf{I}(t) \rangle \bullet \mathbf{E}$. It will frequently be the case that the electrical conductor will be in contact with an electrically insulating, heat reservoir that fixes the average energy of the system so that on average heat flows from the system of interest, the conduction cell, towards the much larger heat reservoir. Nonetheless all the particles in this system (conduction cell plus reservoir) constitute a time reversible dynamical system.

We are interested in a number of problems suggested by this experiment:

1. How do we reconcile the Ohmic heating, with the time reversibility of the microscopic equations of motion? Why isn't there the possibility of Ohmic cooling?
2. For a given initial phase $\mathbf{\Gamma}_j$ that generates some time dependent current $\mathbf{I}_j(S^t\mathbf{\Gamma}_j)$, can we generate Loschmidt's conjugate <u>*antitrajectory*</u> which has a time reversed electric current with time-reversed time ordering?
3. Is there anything we can say about the deviations of the behaviour of individual ensemble members, from the average behaviour?

In noting these three questions, question 2 is slightly different from what is usually mentioned in textbooks that treat reversibility. For the antitrajectory not only is the current the opposite sign to that for the conjugate trajectory but the time ordered fluctuations and transients must exhibit time-reversed time ordering. The last temporal fluctuations that occur on a particular trajectory, are in fact the first fluctuations on the conjugate antitrajectory. This, as we will see in Chapter 7, is hugely important.

Consider a classical system of N interacting particles in a volume V. Initially (at $t = 0$), the microstates of the system are distributed according to a given normalized probability distribution function $f(\boldsymbol{\Gamma};0)$. To apply our results to realistic systems, we separate the *N* particle system into a system of interest and a wall thermostatting region containing $N_W$ particles. Note: the system of interest may contain unthermostatted wall particles. Within the thermostat, $N_{th}$ particles are subject to a fictitious thermostat or ergostat. The thermostat employs a switch, $S_i = 1,0$, which controls how many and which particles are thermostatted, $S_i = 0; \quad 1 \le i \le (N - N_{th})$, $S_i = 1; (N - N_{th} + 1) \le i \le N, N_{th} \le N_W$. We define the thermostat kinetic energy as

$$K_{th} \equiv \sum_{i=1}^{N} S_i \frac{p_i^2}{2m_i}. \tag{2.2.4}$$

and write the equations of motion for the composite *N*-particle system as

$$\begin{aligned} \dot{\mathbf{q}}_i &= \mathbf{p}_i / m_i + \mathbf{C}_i(\boldsymbol{\Gamma}) \bullet \mathbf{F}_e \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i(\mathbf{q}) + \mathbf{D}_i(\boldsymbol{\Gamma}) \bullet \mathbf{F}_e - S_i(\alpha \mathbf{p}_i + \boldsymbol{\gamma}_{th}) \\ \dot{\alpha} &= \left[ \frac{2K_{th}}{3(N_{th} - 1)k_B T_{th}} - 1 \right] \frac{1}{\tau^2}, \end{aligned} \tag{2.2.5}$$

where $\mathbf{F}_i(\mathbf{q}) = -\partial \Phi(\mathbf{q}) / \partial \mathbf{q}_i$ is the interatomic force on particle *i*, $\Phi(\mathbf{q})$ is the interparticle potential energy, $\mathbf{C}_i(\boldsymbol{\Gamma}), \mathbf{D}_i(\boldsymbol{\Gamma})$ are tensorial phase functions that couple the dissipative field to the system of interest.

The term involving, $-S_i \alpha \mathbf{p}_i$ is a deterministic time reversible Nosé-Hoover thermostat [19] used to add or remove heat from the particles in the reservoir region through introduction of an extra degree of freedom described by $\alpha$, $T_{th}$ is a *target* parameter that controls the time averaged kinetic energy of the thermostatted particles, and $\tau$ is the time constant for the integral feedback mechanism of the Nosé-

Hoover thermostat. If $2K_{th} > 3(N_{th}-1)k_B T_{th}$ then $\dot{\alpha} > 0$ and the thermostat multiplier will decrease implying that in the future less energy will be removed from the thermostat particles. Conversely if $2K_{th} < 3(N_{th}-1)k_B T_{th}$ then $\dot{\alpha} < 0$ implying that more energy will be removed from the thermostat in the future. Thus the thermostat tends to stabilize the average thermostat kinetic energy at the value $\bar{K}_{th} = 3(N_{th}-1)k_B T_{th} / 2$.

It is a trivial matter to check that the Nosé-Hoover thermostatted equations of motion are time reversal symmetric. From the third equation in (2.2.5) we see that $\dot{\alpha}$ is even under time reversal. This means that $\alpha$ is odd implying that the whole thermostatting term in the $\dot{\mathbf{p}}$ equation of motion is even as of course is the force.

*Assuming* that the system comes to a nonequilibrium steady state where at long times time averages of smooth phase functions become time independent we also expect there will be a time independent value for the thermostat multiplier $\lim_{t\to\infty}\langle\alpha(t)\rangle = \langle\alpha\rangle_\infty \Rightarrow \lim_{t\to\infty}\langle\dot{\alpha}(t)\rangle = 0$. From (2.2.5) we see that in this steady state

$$\lim_{t\to\infty}\frac{\int_0^t ds\, K_{th}(s)}{t} = \bar{K}_{th} = 3(N_{th}-1)k_B T_{th} / 2 \qquad (2.2.6)$$

In Chapter 5 we will prove that if the dissipative field is in fact zero and the system is *T*-mixing, the system described by (2.2.5) eventually comes to thermodynamic equilibrium and the target temperature of the Nosé-Hoover thermostat $T_{th}$, is then identical to the equilibrium thermodynamic temperature of the system. When the dissipative field is nonzero the thermodynamic temperature of the system of interest is in fact undefined. However in this case if the thermal reservoirs are made arbitrarily large compared to the system of interest the thermal reservoirs will be hardly affected by the system of interest and in this case $T_{th}$ can be regarded as the equilibrium thermodynamic temperature of the thermal reservoir. For an in depth discussion of the Nosé-Hoover thermostat see §5.2 of Evans and Morriss.

One might wonder whether other mathematical forms are possible for the thermostat. Could one replace a thermostat of the form $\dot{\mathbf{p}} = \mathbf{F} - \alpha\mathbf{p}$ with a so-called $\mu$-thermostat: $\dot{\mathbf{p}} = \mathbf{F} - \alpha|p|^{\mu-1}\mathbf{p}$? It turns out that if $\mu \neq 1$, these systems can never relax to equilibrium. Consequently we do not consider these $\mu \neq 1$, thermostats any further.

In equations (2.2.5) the fluctuating force $\boldsymbol{\gamma}_{th} = \frac{1}{N_{th}} \sum_{i=1}^{N} S_i \mathbf{F}_i$ ensures that the macroscopic momentum of the thermostatted particles is a constant of the motion, which we set to zero. Normally $\mathbf{C}_i(\boldsymbol{\Gamma}), \mathbf{D}_i(\boldsymbol{\Gamma}) = \mathbf{0}$ when $S_i = 1$. This means the dissipative field cannot do work on the thermostatting particles. Note however, that the choice of thermostat is reasonably arbitrary, *e.g.* we could use some other choice of time reversible deterministic thermostat, such as one obtained by use of Gauss' Principle of Least Constraint [19] to fix $K_{th}$, and arrive at essentially the same results.

In order to simplify the notation we introduce an extended phase space vector $\boldsymbol{\Gamma}^* \equiv (\boldsymbol{\Gamma}, \alpha)$ and from here on represent this implicitly using $\boldsymbol{\Gamma}$.

**Definition**

In the absence of the thermostatting terms the (Newtonian) equations of motion preserve the phase space volume, $\Lambda \equiv (\partial/\partial \boldsymbol{\Gamma}) \bullet \dot{\boldsymbol{\Gamma}}^{ad} = 0$: a condition known as the *adiabatic incompressibility of phase space*, or AI$\boldsymbol{\Gamma}$ [14]. The equations of motion for the particles in the system of interest are quite natural. The equations of motion for the thermostatted particles are supplemented with unnatural thermostat and force terms. Equations (2.2.5) are time reversible and heat can be either absorbed or given out by the thermostat. Similar constructions have been applied in various studies (see, for example, [16, 20]). Of course, if $S_i = 1$ for all *i*, we obtain a homogeneously thermostatted system that is often studied [19].

The model system could be quite realistic with only some particles subject to the external field. For example, some particles might be charged in an electrical conduction experiment, while other particles may be chemically distinct being solid at the temperatures and densities under consideration. Furthermore these particles may form the thermal boundaries or walls which thermostat and "contain" the electrically charged particles fluid particles inside a conduction cell. In this case $\mathrm{S_i} = 1$ only for wall particles and $\mathrm{S_i} = 0$ for all the fluid particles. This would provide a realistic model of electrical conduction.

In other cases we might consider a homogeneous thermostat where $S_i = 1, \forall i$. It is worth pointing out that as described, equations (2.2.5) are time reversible and heat can be both absorbed and given out by the thermostat.

**Definition**

If we consider a group of atoms within some small volume $\delta V$ centred on a position $\mathbf{r}$, the *local mass density* $\rho(\mathbf{r})$ is defined as

$$\rho(\mathbf{r}) \equiv \sum_{i \in \delta V} m_i / \delta V \tag{2.2.7}$$

and the local streaming velocity **u** is defined by the equation

$$\sum_{i \in \delta V} \mathbf{p}_i \equiv \rho(\mathbf{r})\mathbf{u}(\mathbf{r}) \tag{2.2.8}$$

where $\mathbf{p}_i$ is the momentum of particle *i*, measured in the laboratory frame. Adjusting the physical size of the volume $\delta V$ adjusts the special resolution within which we measure local properties.

**Definition**

If the momenta and velocities are computed relative to the *local streaming velocity* (ie $\mathbf{p}_i - m_i\mathbf{u}(\mathbf{r}_i), \quad \mathbf{v}_i - \mathbf{u}(\mathbf{r}_i)$) they are termed *peculiar momenta and velocities* respectively.

The use of peculiar momenta in the expressions for the kinetic temperature and the internal energy is important. All thermodynamic variables must be independent of the velocity of the frame of reference from which they are measured. For instance if we consider a glass of water in a moving train, the total energy of the molecules comprising the glass of water is dependent on the velocity of the train. However the internal energy and the kinetic temperature are independent of the motion of the train. All thermodynamic quantities must be evaluated using momenta and velocities measured in the local streaming velocity frame of reference. In writing (2.2.5) we chose the fluctuating force $\boldsymbol{\gamma}_{th}$ to make the momentum of the thermostat a constant of the motion with the value zero. This means that the average momentum of the entire system is also zero ensuring that all the momenta appearing in (2.2.5) are in fact peculiar.

**Definition**

For systems with no applied external field or for which the external field is dissipative, the Hamiltonian $H_0 \equiv \sum_{i:S_i=1} [p_i^2 / 2m + 1/2 \sum_j \Phi(\mathbf{q})]$ expressed in peculiar momenta, $\mathbf{p}_i$, has an average value that is the thermodynamic internal energy of the system. The *internal energy* is just the energy of the system with the local streaming kinetic energy removed. This thermodynamic quantity is completely mechanical – as the so-called first "law" of thermodynamics shows. The definition of the internal energy is valid even far from equilibrium provided the streaming velocity (2.2.8) is well defined.

An alternative thermostatting mechanism is to choose the thermostat multiplier $\alpha$ to make either the internal energy of the thermostat $H_0 \equiv \sum_{i:S_i=1} [p_i^2 / 2m + 1/2 \sum_j \Phi(\mathbf{q})]$, or of the entire system a constant of the motion. For this *ergostatted dynamics*, the thermostat multiplier, α, is chosen as the instantaneous solution to the equation,

$$\begin{aligned} \dot{H}_0(\mathbf{\Gamma}) &\equiv -\mathbf{J}(\mathbf{\Gamma})V \bullet \mathbf{F}_e - 2K_W(\mathbf{\Gamma})\alpha(\mathbf{\Gamma}) \\ &\equiv -\mathbf{J}(\mathbf{\Gamma})V \bullet \mathbf{F}_e - \dot{Q} \\ &= 0 \end{aligned} \quad . \qquad (2.2.9)$$

**Definition**

The heat *added to the thermostat* per unit time $\dot{Q}(t)$ is defined in equation (2.2.9). [**Aside**: When the thermostat is overwhelmingly larger than the system of interest and when it is in thermodynamic equilibrium the thermostat increases its entropy at a rate $\dot{S}_{th} = \dot{Q} / T_{th}$.]

A third thermostatting mechanism is where we make the peculiar kinetic energy of the thermostat

$$K_{th} \equiv \sum_{S_i=1} p_i^2 / 2m = (3N_{th} - 4)k_B T_{th} / 2, \qquad (2.2.10)$$

a constant of the motion; in which case we speak of isokinetic dynamics. In (2.2.10) $N_{th} = \sum S_i$ .

**Definition**

The quantity $T_{th}$ defined by (2.2.10) is called the *kinetic temperature* of the thermostat.

Both of these latter thermostatting methods involve differential feedback and the equations of motion can be derived using Gauss' Principle of Least Constraint to fix either the internal energy or the thermostat peculiar kinetic energy. In both cases the first two equations of (2.2.5) still apply but the third equation in (2.2.5) is replaced by an explicit expression for the multiplier. These Gaussian thermostats were in fact the first time-reversible deterministic thermostats. Hoover developed the isokinetic thermostat while Evans (simultaneously) developed the ergostat. It is a trivial matter to check that these thermostats are time reversal symmetric. We note from (2.2.9) that as in the Nosé-Hoover thermostat, $\alpha$ is odd under time reversal.

**Definitions**

The *dissipative flux* **J** due to the driving *dissipative field*, $\mathbf{F}_e$ is defined as,

$$\dot{H}_0^{ad} \equiv -\mathbf{J}(\mathbf{\Gamma})V \bullet \mathbf{F}_e \equiv -\sum [\mathbf{p}_i/m \bullet \mathbf{D}_i - \mathbf{F}_i \bullet \mathbf{C}_i] \bullet \mathbf{F}_e, \qquad (2.2.11)$$

$\dot{\mathrm{H}}_0^{\mathrm{ad}}$ is the adiabatic time derivative of the internal energy (*i.e.* it is computed without the contributions from the thermostat) and V is the volume of the system. It is always assumed that the equations of motion for the driven system satisfy the adiabatic incompressibility of phase space condition. In Chapter 6, we will show in detail why the dissipative flux is so-named.

**Definitions**

The *adiabatic time derivative* of $H_0$ is in fact the *work* performed on the system by the dissipative field because it the total change of energy minus the heat removed by the thermostat/ergostat.

Equation (2.2.6) is a statement of the First Law of Thermodynamics for an ergostatted nonequilibrium system. The energy removed from (or added to) the system by the ergostat must be balanced instantaneously by the work done on (or removed from) the system by the external dissipative field, $F_e$. For ergostatted dynamics we solve (1.2.3) for the ergostat multiplier and substitute this phase function into the equations of motion. For isokinetic dynamics we solve an equation which is analogous to (1.2.3) but which ensures that the kinetic temperature of the walls or system, is fixed [16]. The equations of motion (2.2.5) are reversible where the thermostat multiplier is defined in this way (2.2.7).

A simple example system is the case of electrical conductivity. There we could model the charged ions of a molten salt ($\mathbf{C}_i = \mathbf{0}, \mathbf{D}_i = c_i\mathbf{I}$ where $c_i$ is the electric charge of particle *i*) subject to an electric field, $\mathbf{F}_e = \mathbf{E}$. We could surround these ions with neutral atoms ($c_i = 0$) of a solid wall that contains the electrically conducting molten salt. Further, outside these realistically modelled wall particles we could then have a layer of thermostatted or ergostatting electrically neutral, wall particles. These thermostatting particles can be located arbitrarily far from the system of interest.

One might object that our analysis is compromised by our use of these artificial (time reversible) thermostats. Since, the thermostat can be made arbitrarily remote from the system of physical interest [17], the system cannot ‘know’ the precise details of how the heat is ultimately removed. This means that the results obtained for the system using our simple mathematical thermostat must be the same as the those we would infer for the same system surrounded (at a distance) by a real physical thermostat (say with a huge heat capacity). These mathematical thermostats may be unrealistic, however in the final analysis they are very convenient but ultimately irrelevant devices. Importantly, they allow us to do the mathematical bookkeeping that is necessary in the study of systems that exchange heat with their surroundings. Ultimately the work that is on average performed on the system of interest is ultimately on average, transformed into heat that is absorbed by an infinitely large system that can be regarded as being arbitrarily close to thermodynamic equilibrium, arbitrarily far from the system of interest. That reservoir has a known kinetic temperature which as we will see in chapter 5 is the equilibrium thermodynamic temperature of that reservoir.

## 2.3 EXAMPLE: HOMOGENEOUSLY THERMOSTATTED SLLOD EQUATIONS FOR PLANAR COUETTE FLOW

A very important dynamical system is the standard model for thermostatted planar Couette flow – the so-called SLLOD equations for shear flow. Consider $N$ particles under shear. In this system the external field is the shear rate, $\partial u_x / dy = \dot{\gamma}(t)$ (the y-gradient of the x-streaming velocity, $u_x$). The equations of motion for the particles are given by the so-called homogeneously thermostatted SLLOD equations,

$$\dot{\mathbf{q}}_i = \mathbf{p}_i / m + \mathbf{i}\dot{\gamma} y_i, \quad \dot{\mathbf{p}}_i = \mathbf{F}_i - \mathbf{i}\dot{\gamma} p_{yi} - \alpha \mathbf{p}_i \,. \tag{2.3.1}$$

Here, **i** is a unit vector in the positive x-direction. At low Reynolds number where a planar velocity profile is expected to be stable, the SLLOD momenta are in fact peculiar momenta (*i.e.* they are measured relative to the average streaming velocity of the individual particles, $\mathbf{u}(\mathbf{q}_i, t) = \mathbf{i}\dot{\gamma} y_i$).

As first pointed out be Evans and Morriss in 1984, the adiabatic SLLOD equations of motion give an exact description of planar Couette flow arbitrarily far from equilibrium. This is because the adiabatic SLLOD equations for a step function strain rate $\partial u_x(t) / \partial y = \dot{\gamma}(t) = \dot{\gamma}\Theta(t)$, are equivalent to Newton's equations after the impulsive imposition of a linear velocity gradient at t = 0 (*i.e.* $d\mathbf{q}_i(0^+) / dt = d\mathbf{q}_i(0^-) / dt + \mathbf{i}\dot{\gamma} y_i(0)$ [16]. There is thus a remarkable subtlety in the SLLOD equations of motion. If one starts at t = $0^-$, with a canonical ensemble of systems then at t = $0^+$, the SLLOD equations of motion transform this initial ensemble into the local equilibrium ensemble for planar Couette flow at a shear rate $\dot{\gamma}$.

Because the effects of thermostatting are asymptotically quadratic in the strain rate, the homogeneously thermostatted SLLOD equations of motion give an exact description of the *linear* response of a system to planar Couette flow – even a time dependent planar Couette flow. For a mathematical proof see Evans and Morriss.

At low Reynolds number, the SLLOD momenta, $\mathbf{p}_\mathrm{i}$, are peculiar momenta and $\alpha$ is determined using Gauss's Principle of Least Constraint to keep the internal energy, $H_0 = \sum p_i^2 / 2m + \Phi(\mathbf{q})$, fixed [16] (Note: the internal energy is the sum of the peculiar kinetic energy and the potential energy. It is *not* the sum of the laboratory kinetic energy and the potential energy.). Thus, for our system (2.3.1)

$$\dot{H}_0 = \sum \dot{\mathbf{p}}_i \bullet \frac{\partial K}{\partial \mathbf{p}_i} + \dot{\mathbf{q}}_i \bullet \frac{\partial \Phi}{\partial \mathbf{q}_i} - 2K(\mathbf{p})\alpha$$

$$= \sum -\mathbf{i}\gamma p_{yi} \bullet \frac{\mathbf{p}_i}{m} - \mathbf{i}\dot{\gamma} y_i \bullet \mathbf{F}_i - 2K(\mathbf{p})\alpha \qquad (2.3.2)$$

$$= -\dot{\gamma} \sum \frac{p_{yi} p_{xi}}{m} + y_i F_{xi} - 2K(\mathbf{p})\alpha \equiv -\dot{\gamma} P_{xy} V - 2K(\mathbf{p})\alpha$$

where $P_{xy}$ is the well known expression for the *xy*-element of the pressure tensor in a homogeneous system. We can fix the internal energy of the system by choosing $\alpha$ as:

$$\alpha = -P_{xy} \dot{\gamma} V / 2K(\mathbf{p}), \qquad (2.3.3)$$

The *xy*-element of the pressure tensor, is the dissipative flux, J [16]. From the equations of motion we can see that the rate of change of internal energy for for isoenergetic SLLOD dynamic is:

$$\dot{H}_0 = -P_{xy} \dot{\gamma} V - 2K\alpha = -P_{xy} \dot{\gamma} V - \dot{Q} = 0$$

Thus the ergostat increases the internal energy of the system at a rate $-\dot{Q}$, that is precisely and instantaneously equal to the rate at which work is expended on the system by the shearing deformation namely $-P_{xy} \dot{\gamma} V$

The corresponding isokinetic form for the thermostat multiplier is,

$$\alpha = \frac{\sum_{i}^{N} \mathbf{F}_i \bullet \mathbf{p}_i - \dot{\gamma} \sum_{i=1}^{N} p_{xi} p_{yi} / m}{\sum_{i=1}^{N} \mathbf{p}_i^2 / m}. \qquad (2.3.3)$$

We note there are obviously Nosé-Hoover thermostatted forms for the SLLOD equations but we do not give these here.

If the thermostatted SLLOD system comes to a nonequilibrium steady state the time average rate of shearing work that is performed by the shear on the system is equal to the time average rate at which heat is removed by the thermostat from the system. If the system is ergostated this balance is achieved instantaneously.

The ergostatted and thermostatted SLLOD equations of motion ((2.3.1),(2.3.2), (2.3.3)) are time reversible [16]. In the weak flow limit these equations yield the correct Green-Kubo relation for the linear shear viscosity of a fluid [16]. We have also proved that in this limit, the linear response obtained from the equations of motion or equivalently from the Green-Kubo relation are identical to leading order in N the number of particles. In the far from equilibrium regime, Brown and Clarke [20] have shown that the results for homogeneously thermostatted SLLOD dynamics are indistinguishable from those for boundary thermostatted shear flow, up to the limiting shear rate above which a steady state for boundary thermostatted systems is not stable[2] .

In computer simulations if one wants to carry out nonequilibrium molecular dynamics simulations one has to supplement the SLLOD equations of motion with appropriate boundary conditions – you cannot simulate infinite systems. If you start with a cubic periodic system at $t$=0, then the shear motion causes the unit cells above and below the primitive cell will slide to the right and left above and below the primitive cell at constant speeds $V^{1/3}\dot{\gamma}$ . These cells move to positions $\pm V^{1/3}\dot{\gamma}t$ . The positions of these cells affects the forces $\mathbf{F}_i$ on the $N$ particles in the primitive cell. This means that for finite periodic systems the SLLOD equations of motion as implemented are in fact non-autonomous!

In practice for the short range forces like Lennard-Jones or WCA forces these non-autonomous effects can be hard or near impossible to observe in 3 dimensional systems where $N \geq\sim 100$ . In working with small systems the non-autonomous effects can be easily observed particularly close to the freezing density.

---

[2] Dissipation is extensive $=$ O($N$) while entropy absorption by the thermostat $=$ O( $N^{2/3}$ ). So for any given system there is a limiting shear rate beyond which boundary thermostatting is not possible.

## 2.4 PHASE CONTINUITY EQUATION

We have introduced the phase space distribution function, $f(\mathbf{\Gamma};t)$. It gives the probability per unit phase space volume, of finding ensemble members near the phase vector $\mathbf{\Gamma}$, at time *t.*

There is a simple exact equation of motion for this density. That equation is called the Phase Continuity Equation (Gibbs 1901). [In most text books this equation is called the Liouville Equation. In fact Liouville's 1838 paper doesn't refer to the phase space density at all. This is because the concept of a statistical mechanical ensemble hadn't been developed at this early date.]

The proof of the Phase Continuity equation

$$\frac{df(\mathbf{\Gamma},t)}{dt} = -f(\mathbf{\Gamma},t)\frac{\partial}{\partial \mathbf{\Gamma}} \bullet \dot{\mathbf{\Gamma}}(\mathbf{\Gamma}) \qquad (2.4.1)$$

is identical to the proof of the mass continuity equation for a compressible fluid. Both equations express the fact that the total mass of a compressible fluid or the total number of ensemble members in phase space, is conserved. The total number $N_\Gamma$ of ensemble members inside an enclosing phase space volume, $V_\Gamma$ must be related to the total integrated flux taken over the enclosing surface, $S_\Gamma$ .

$$\begin{aligned} \frac{dN_\Gamma}{dt} &= \int_{V_\Gamma} d\mathbf{\Gamma} \; \frac{\partial f(\mathbf{\Gamma},t)}{\partial t} = -\int_{S_\Gamma} d\mathbf{S}_\Gamma \bullet \dot{\mathbf{\Gamma}} f(\mathbf{\Gamma},t) \\ &= -\int_{V_\Gamma} d\mathbf{\Gamma} \; \frac{\partial}{\partial \mathbf{\Gamma}} \bullet [\dot{\mathbf{\Gamma}} f(\mathbf{\Gamma},t)] \end{aligned} \qquad (2.4.2)$$

Since this equation is true for arbitrary phase space volumes $V_\Gamma$ we see that

$$\frac{\partial f(\mathbf{\Gamma};t)}{\partial t} = -\frac{\partial}{\partial \mathbf{\Gamma}} \bullet [\dot{\mathbf{\Gamma}} f(\mathbf{\Gamma};t)] \qquad (2.4.3)$$

Use of the chain rule then yields the streaming equivalent form of the Phase Continuity equation given above. For isolated systems that do not exchange heat with their surroundings, the streaming derivative of the *N*-particle distribution function is identically zero. For thermostatted systems the phase space expansion factor $\Lambda$ is for isokinetic or isoenergetic systems – see A2

$$\Lambda \equiv \frac{\partial}{\partial \boldsymbol{\Gamma}} \bullet \dot{\boldsymbol{\Gamma}} = -(D_C N_W - 1)\alpha \,. \qquad (2.4.4)$$

**Definition**

In order to carry out symbolic calculations of the time dependent *N*-particle distribution function it is convenient to define the *f-Liouvillian* $L$ where

$$\frac{\partial}{\partial \boldsymbol{\Gamma}} \bullet [\dot{\boldsymbol{\Gamma}} ...] \equiv iL... \qquad (2.4.5)$$

[We note in passing that while Liouville never discussed phase space density as required in the phase continuity equation he did discuss, in a different notation, what we call the *p*-Liouvillean.] Using this operator we see that the phase continuity equation can be written as

$$\frac{\partial f(\boldsymbol{\Gamma};t)}{\partial t} = -iLf(\boldsymbol{\Gamma};t) \qquad (2.4.6)$$

This equation has a formal solution

$$f(\boldsymbol{\Gamma};t) = \exp[-iLt] f(\boldsymbol{\Gamma};0) \qquad (2.4.7)$$

The correctness of this solution can be checked by differentiation.

If we return to (2.4.3) we see that

$$\frac{df(\mathbf{\Gamma};t)}{dt}=\frac{\partial f(\mathbf{\Gamma};t)}{\partial t}+\dot{\mathbf{\Gamma}}(\mathbf{\Gamma})\bullet\frac{\partial f(\mathbf{\Gamma};t)}{\partial\mathbf{\Gamma}}$$

$$=-\frac{\partial}{\partial\mathbf{\Gamma}}\bullet[\dot{\mathbf{\Gamma}}(\mathbf{\Gamma})f(\mathbf{\Gamma};t)]+\dot{\mathbf{\Gamma}}(\mathbf{\Gamma})\bullet\frac{\partial f(\mathbf{\Gamma};t)}{\partial\mathbf{\Gamma}} \qquad (2.4.8)$$

$$=-f(\mathbf{\Gamma};t)\frac{\partial}{\partial\mathbf{\Gamma}}\bullet\dot{\mathbf{\Gamma}}(\mathbf{\Gamma})=-f(\mathbf{\Gamma};t)\Lambda(\mathbf{\Gamma}),\ \forall\mathbf{\Gamma}\in D$$

If we set $\mathbf{\Gamma}\to S^t\mathbf{\Gamma}$ we obtain

$$\frac{df(S^t\mathbf{\Gamma};t)}{dt}=-\Lambda(S^t\mathbf{\Gamma})f(S^t\mathbf{\Gamma};t)\,. \qquad (2.4.9)$$

**Definition**

This equation (2.4.9) is termed the *streaming or Lagrangian form of the phase continuity equation*. For a given initial phase $\mathbf{\Gamma}$ (2.4.9) is a simple first order ordinary differential equation for the density along the phase space trajectory. Its solution can be written as

$$f(S^t\mathbf{\Gamma};t)=\exp[-\int_0^t ds\ \Lambda(S^s\mathbf{\Gamma})]f(\mathbf{\Gamma};0) \qquad (2.4.10)$$

The correctness of which can easily be checked by differentiation: $df(S^t\mathbf{\Gamma};t)/dt=-\Lambda(S^t\mathbf{\Gamma})f(S^t\mathbf{\Gamma};t)$. So the distribution function at time *t* at the streamed position of the phase vector originating at $\mathbf{\Gamma}$ is related to the path integral of the phase space expansion factor along the phase space trajectory.

We could also consider the time dependence of the measure of an infinitesimal phase space volume $dV_\Gamma(S^s\mathbf{\Gamma})$ centred on the streamed position $S^s\mathbf{\Gamma}:0\le s\le t$ along the phase space trajectory. This phase space volume contains a fixed number of ensemble members and obeys the following equation of motion:

$$dV_\Gamma(S^t\mathbf{\Gamma})=\exp[\int_0^t ds\ \Lambda(S^s\mathbf{\Gamma})]dV_\Gamma(\mathbf{\Gamma}) \qquad (2.4.11)$$

In nonequilibrium steady states, experience shows that the time average value of the thermostat multiplier is positive and therefore the time average phase space expansion factor is negative. This implies that for almost every initial phase space vectoror the streamed density $f(S^t\mathbf{\Gamma};t) \sim \exp[\bar{\alpha}t]f(\mathbf{\Gamma};0) \to +\infty$ while the corresponding streamed phase space volume goes to zero exponentially in time.

**Definition**

We are now in a position to compute *ensemble averages*, $\langle B(t)\rangle$ of an arbitrary integrable phase function $B(\mathbf{\Gamma})$

$$\langle B(t)\rangle \equiv \int d\mathbf{\Gamma}\; B(\mathbf{\Gamma})f(\mathbf{\Gamma};t) = \int d\mathbf{\Gamma}\; B(S^t\mathbf{\Gamma})f(\mathbf{\Gamma};0) \qquad (2.4.12)$$

Physically this equation is rather obvious. We can formally prove the correctness of this equation by noting that by integrating by parts:

$$\int d\mathbf{\Gamma}\; f(\mathbf{\Gamma};t)iLB(\mathbf{\Gamma}) = -\int d\mathbf{\Gamma}\; B(\mathbf{\Gamma})i\mathsf{L}f(\mathbf{\Gamma};t)\,, \qquad (2.4.13)$$

from which we deduce that

$$\begin{aligned}\langle B(t)\rangle &= \int d\mathbf{\Gamma}\; B(\mathbf{\Gamma})\exp[-i\mathsf{L}t]f(\mathbf{\Gamma};0)\\ &= \int d\mathbf{\Gamma}\; f(\mathbf{\Gamma};0)\exp[iLt]B(\mathbf{\Gamma}) \qquad (2.4.14)\\ &\equiv \int d\mathbf{\Gamma}\; f(\mathbf{\Gamma};0)B(S^t\mathbf{\Gamma})\end{aligned}$$

We note that if the system satisfies the AI$\mathbf{\Gamma}$ condition and if there are no thermostats applied then $\Lambda = 0$, $L = \mathsf{L}$ and the single Liouville operator is Hermitian. For thermostatted systems the *f*- and *p*-Liouvilleans are not in fact equal.

Finally we note that all phase space integrals given above should be carried out over some specified phase space domain. We have omitted this for simplicity but if one wants to verify these equations *e.g.* (2.4.9) by integrating by parts then the specification of this domain is essential.

## 2.5 LYAPUNOV INSTABILITY AND STATISTICAL MECHANICS

In this section we give the briefest of introductions to an entire new field of research. Many statements are made without proof. To fill out the details would fill another book. This interested reader should consult the references cited in this section.

We include this material because without some knowledge of the dimensional reduction processes in time reversible deterministic steady states, the reader will be puzzled by many apparent contradictions. How can it be that in a nonequilibrium *steady state* the entropy is not time independent but instead decreases at a constant average *rate* towards negative infinity?

The Lyapunov exponents are used in dynamical systems theory to characterise the stability of phase space trajectories. If one imagines two systems that evolve in time from phase vectors $\mathbf{\Gamma}_1, \mathbf{\Gamma}_2$ that initially are very close together $|\mathbf{\Gamma}_1 - \mathbf{\Gamma}_2| \equiv |\delta\mathbf{\Gamma}| \to 0$, then one can ask how the separation between these two systems evolves in time. Oseledec's Theorem says for nonintegrable systems under very general conditions, that the separation vector asymptotically grows or shrinks *exponentially* in time. Of course this does not happen for integrable systems, but then again very few systems are integrable. A system is said to be chaotic if the separation vector asymptotically *grows* exponentially with time. Most systems in Nature are chaotic: the world weather and high Reynolds Number flows are chaotic. In fact all systems that obey thermodynamics are chaotic.

In 1990 the first of a remarkable set of relationships between phase space stability measures (*i.e.* Lyapunov exponents) and thermophysical properties were discovered by Evans, Cohen and Morriss [21] and separately by Gaspard and Nicolis [22]. More recently Lyapunov exponents have been used to assign dynamical probabilities to the observation of phase space trajectory segments [14, 15, 23]. This is something quite new to statistical mechanics where hitherto probabilities had been given (only for equilibrium systems!), on the basis of the value of the Hamiltonian (*i.e.* the weights are static).

Suppose the autonomous equations of motion (1.2.1), are written

$$\dot{\mathbf{\Gamma}} = \dot{\mathbf{\Gamma}}(\mathbf{\Gamma}) . \qquad (2.5.1)$$

**Definition**

It is trivial to see that the equation of motion for an infinitesimal phase space separation vector, $\mathrm{d}\mathbf{\Gamma}$, can be written as:

$$d\dot{\mathbf{\Gamma}} = \mathbf{T}(\mathbf{\Gamma}) \bullet d\mathbf{\Gamma} \tag{2.5.2}$$

where $\mathbf{T} \equiv \partial\dot{\mathbf{\Gamma}}(\mathbf{\Gamma},t)/\partial\mathbf{\Gamma}$ is the *stability matrix* for the flow.

**Definition**

The propagation of the tangent vectors is therefore given by,

$$dS^t\mathbf{\Gamma} = \Xi(\mathbf{\Gamma};t) \bullet d\mathbf{\Gamma} \tag{2.5.3}$$

where the *tangent vector propagator* is:

$$\Xi(\mathbf{\Gamma};t) = \exp_L\left(\int_0^t ds\, \mathbf{T}(S^s\mathbf{\Gamma})\right) \tag{2.5.4}$$

and $\exp_L$ is a left time-ordered exponential. The correctness of (2.5.4) can be checked by differentiation. The time evolution of these tangent vectors are used to determine the Lyapunov spectrum for the system. The Lyapunov exponents thus represent the rates of divergence of nearby points in phase space.

If $d\mathbf{\Gamma}_i$ is an eigenvector of $\Xi^T(\mathbf{\Gamma};t) \bullet \Xi(\mathbf{\Gamma};t)$ and if the Lyapunov exponents are defined as [24]:

$$\left\{\lambda_i; i = 1,..2D_C N\right\} \equiv \lim_{t\to\infty} \frac{1}{2t} \ln\left(eigenvalues\left(\Xi^T(\mathbf{\Gamma};t) \bullet \Xi(\mathbf{\Gamma};t)\right)\right), \tag{2.5.5}$$

then the Lyapunov exponents describe the growth rates of the set of orthogonal tangent vectors (eigenvectors of $\left(\Xi^T(t) \bullet \Xi(t)\right)$), $\{d\mathbf{\Gamma}_i(t); i = 1,2d_C N\}$,

$$\lim_{t\to\infty}\frac{1}{2t}\ln\frac{\left|dS^t\mathbf{\Gamma}_i \bullet dS^t\mathbf{\Gamma}_i\right|}{\left|d\mathbf{\Gamma}_i \bullet d\mathbf{\Gamma}_i\right|} = \lim_{t\to\infty}\frac{1}{2t}\ln\frac{\left|d\mathbf{\Gamma}_i^T \bullet \Xi^T(t) \ \bullet \Xi(t) \bullet d\mathbf{\Gamma}_i\right|}{\left|d\mathbf{\Gamma}_i \bullet d\mathbf{\Gamma}_i\right|}$$

$$= \frac{1}{2t}\ln\frac{\left|d\mathbf{\Gamma}_i^T \bullet \exp[2\lambda_i t]\mathbf{1} \bullet d\mathbf{\Gamma}_i\right|}{\left|d\mathbf{\Gamma}_i \bullet d\mathbf{\Gamma}_i\right|} \qquad (2.5.6)$$

$$= \lambda_i, \quad i = 1,...,2D_C N$$

(Note: the transpose matrix not only transposes the rows and column it also transposes time ordering, $\Xi^T(\mathbf{\Gamma};t) = \exp_R\left(\int_0^t ds\, \mathbf{T}^T(S^s\mathbf{\Gamma})\right)$.)

By convention the exponents are ordered such that $\lambda_1 \geq \lambda_2 \geq . . . \geq \lambda_{2D_C N}$. It can be shown that the Lyapunov exponents are independent of the metric used to measure phase space lengths. They are also independent for T-mixing steady states (see Chapter 6) of the initial position $\mathbf{\Gamma}$ of the "mother" phase space trajectory.

In general there will be a number of Lyapunov exponents that are zero. For example there will $2D_C$ zero exponents for each Cartesian momentum component that is conserved since momentum conservation also means that the associated position of the centre of mass of each Cartesian coordinate is constant. In autonomous systems there will be another zero exponent associated with time translation invariance. In isokinetic or isoenergetic systems there will each be another zero exponent associated with this additional constant of the motion. To keep the notation flexible we will say that there are $f$ zero Lyapunov exponents.

In order to calculate the Lyapunov spectrum, one does not normally use (2.5.5). Benettin et al. developed a technique whereby the finite but small displacement vectors are periodically rescaled and orthogonalised during the course of a solution of the equations of motion [25, 26]. Hoover and Posch [27] pointed out that this rescaling and orthogonalisation can be carried out continuously by introducing constraints to the equations of motion of the tangent vectors [28]. With this modification, orthogonality and tangent vector length are maintained at all times during the calculation in much the same way as our thermostats and ergostats maintain fixed values for the kinetic temperature or the internal energy.

In theory, the $2D_C N$ eigenvalues of the real symmetric matrix $\Xi^T(t) \bullet \Xi(t)$ can also be used to calculate the Lyapunov spectrum in the limit $t \to \infty$. Since

$\Xi(S^t\boldsymbol{\Gamma})$ is dependent only on the mother trajectory, $S^t\boldsymbol{\Gamma}$, calculation of the Lyapunov exponents from the eigenvalues of $\Xi^T(t)\bullet\Xi(t)$ does not require the solution of 2dN tangent trajectories as in the methods mentioned in the previous paragraph. However, after a short time, numerical difficulties are encountered using this method due to the enormous difference in the magnitude of the eigenvalues of the $\Xi^T(t)\bullet\Xi(t)$ matrix[3]. The use of QR decompositions (where where $\Xi(t)=\mathbf{Q}\bullet\mathbf{R}$ and $\mathbf{R}$ is a real upper triangular matrix with positive diagonal elements and $\mathbf{Q}$ is a real orthogonal matrix) reduces this problem [24, 29]. Use of the QR-decomposition is equivalent to the reorthogonalisation/rescaling of the displacement vectors in the scheme discussed above [30].

We note that the Lyapunov exponents are only defined in the long time limit and if the simulated *nonequilibrium* fluid does not reach a stationary state, the exponents will not converge to constant values. It is useful for the purposes of this work to define time-dependent exponents as:

$$\left\{\lambda_i(t;\boldsymbol{\Gamma});i=1,...2D_CN\right\}=\frac{1}{2t}\ln\left(eigenvalues\left(\Xi^T(t;\boldsymbol{\Gamma})\bullet\Xi(t;\boldsymbol{\Gamma})\right)\right). \qquad (2.5.9)$$

Unlike the Lyapunov exponents, these finite time exponents will depend on the initial phase space vector, $\boldsymbol{\Gamma}$ and the length of time over which the tangent vectors are integrated, and we therefore will refer to them as finite-time, local Lyapunov exponents.

The systems considered here are chaotic: they have at least one positive Lyapunov exponent. This means that (except for a set of zero measure) points that are initially close will diverge after some time, and therefore information on the initial phase space position of the trajectory will be lost. Points that are initially close will eventually span the accessible phase space of the system.

The sum of the first two Lyapunov exponents shows how the fastest growing area grows. The sum of the first three Lyapunov exponents gives the rate of growth of the fastest growing 3-volume grows,.. etc. The Lyapunov exponents of an equilibrium (Hamiltonian) system sum to zero, reflecting the phase space conservation of these systems.

---

[3] It rapidly becomes an illconditioned matrix.

The sum of all the Lyapunov exponents is in fact equal to the time averaged phase space expansion factor so that for isokinetic systems – see A2:

$$\sum_{i=1}^{2D_C N-f} \lambda_i = \lim_{t\to\infty} \overline{\frac{\partial}{\partial\mathbf{\Gamma}}\bullet\dot{\mathbf{\Gamma}}}_t = \lim_{t\to\infty}\overline{\Lambda}_t = -\lim_{t\to\infty}[D_C(N-1)-1]\overline{\alpha}_t \,. \qquad (2.5.10)$$

This is because the phase space expansion is the average rate of increase of the ostensibly dimensioned phase space volume element. Comparing this equation with (2.4.11) shows that the sum of all the Lyapunov exponents gives the exponential rate at which the streamed phase space volume vanishes

$$\lim_{t\to\infty}\delta V_{\mathbf{\Gamma}}(S^t\mathbf{\Gamma}) = \exp[\overline{\Lambda}t]\delta V_{\mathbf{\Gamma}}(\mathbf{\Gamma}) = \exp[\sum_{i=1}\lambda_i t]\delta V_{\mathbf{\Gamma}}(\mathbf{\Gamma}) \,. \qquad (2.5.11)$$

As we will see later, in equilibrium systems all properties including Lyapunov exponents, must be invariant under time reversal. This implies that time reversal of Lyapunov spectra for equilibrium systems must transform the spectrum into itself which in turn means that for all equilibrium systems, the exponents must arrange themselves into conjugate pairs that each sum to zero.

$$\lambda_{\max}^{eq} + \lambda_{\min}^{eq} = \lambda_{\max+1}^{eq} + \lambda_{\min-1}^{eq} = ... = 0 \qquad (2.5.12)$$

If the ostensible phase space dimension is odd, the unpaired exponent must be zero. In fact there could be multiple exponents that are zero since this would not violate the time reversal property.

The symplectic eigenvalue theorem shows that for all autonomous symplectic dynamical systems with time independent Lyapunov exponents the conjugate exponents must pair about zero as in (2.5.11). If the system is stationary in time, this pairing about zero can only happen if the system eventually relaxes towards equilibrium.

Equation (2.5.10) also shows how Lyapunov exponents are related to time averaged dissipative fluxes and thereby to transport coefficients.  We define a nonlinear transport coefficient in terms of the long time average of the dissipative flux divided by the dissipative field as

$$L(F_e) \equiv -\lim_{t\to\infty} \frac{\overline{J}_t}{F_e} \tag{2.5.13}$$

**Definition**

Then using equations we see derive what is known as the *Lyapunov Sum Rule* which for an *N*-particle isokinetic systems reads:

$$L_N(F_e) = -\frac{\sum_{i=1}^{2D_C N-f} \lambda_i}{(D_C(N-1)-1)F_e^2 V k_B T}\,. \tag{2.5.14}$$

For homogeneously thermostatted symplectic systems, Evans, Cohen and Morriss showed in 1990 that the Lyapunov spectrum has a conjugate pairing symmetry about the time averaged value of the thermostat multiplier so that the nonlinear transport coefficient can be calculated by summing any conjugate pair of Lyapunov exponents. Since the largest and the smallest exponents are the easiest to compute, for such systems we can write:

$$L_N(F_e) = -\frac{(\lambda_{\max} + \lambda_{\min})}{F_e^2 V k_B T} \tag{2.5.15}$$

**Definition**

This relation (2.5.14), is called the *Conjugate Pairing Rule* for homogeneously thermostatted adiabatically symplectic systems.

Equations (2.5.14,15) show how apparently abstract mathematical quantities such as Lyapunov exponents which characterize the stability or otherwise of phase space trajectories, are related to measurable physical properties such as transport coefficients.

For thermostatted steady states, the Lyapunov sum is negative. This indicates that the phase space collapses onto a lower dimensional attractor in the original phase space. The set of Lyapunov exponents, can be used calculate the dimension of the object whose volume is preserved by the dynamics.

**Definition**

This dimension is called the *Kaplan-Yorke dimension* defined as

$$D_{KY,N} \equiv N_{KY} + \sum_{i=1}^{N_{KY}} \lambda_i / \left|\lambda_{N_{KY}+1}\right| , \qquad (2.5.16)$$

where $N_{N_{KY}}$ is the largest integer or which $\sum_{i=1}^{N_{KY}} \lambda_i > 0$. As you sum the Lyapunov exponents from the largest to the smallest you start by summing at least one positive number – because the system is chaotic. If the system satisfies the Second Law of Thermodynamics the time averaged thermostat multiplier is positive indicating that on average work is converted to heat that must be removed by the thermostat in order to maintain steady state conditions. From the Lyapunov Sum Rule, (2.5.13) we see that summing all the Lyaponov exponents gives a negative number. Somewhere during the summation process the running sum changed from being positive to negative. Using linear interpolation between the integer exponents, the Kaplan-Yorke dimension is the dimension of that object whose volume is preserved by the dynamics.

In the linear response regime close to equilibrium there is an exact relation between the Kaplan-Yorke dimension of the steady state and the zero field transport coefficient for an *N*-particle system:

$$L_N(F_e = 0) = \lim_{F_e \to 0} \frac{(2D_C N - f - D_{KY,N}(F_e))\lambda_{\max,N}(F_e) k_B T}{V F_e^2} \qquad (2.5.17)$$

In this limit the dimensional reduction is less than 1: $2D_C N - f - D_{KY,N}(F_e) < 1$.

These remarkable equations show not only how to calculate the dimension of the invariant steady state attractor but how this dimensional reduction is related to a physical property, namely a transport coefficient.

## 2.6 ENTROPY IN DETERMINISTIC NONEQUILIBRIUM STATES.

The fine-grained Gibbs entropy $S_G$, is defined as (see §5.6 for a discussion about why this quantity is useful at equilibrium):

$$S_G(t) \equiv -k_B \int_D d\mathbf{\Gamma}\ f(\mathbf{\Gamma};t) \ln f(\mathbf{\Gamma};t)] \tag{2.6.1}$$

We do not discuss why entropy is defined in this way. We will however, explain why this quantity is so problematic in deterministic nonequilibrium steady states. From the discussion of Lyapunov exponents and the Kaplan-Yorke dimension we know that the nonequilibrium density is distributed in a space of lower dimension than the ostensible phase space dimension, $2D_C(N-1)-1$. This is analogous to condensing a density from a two dimensional area onto a one dimensional line. When the entropy is defined as above the integral is to be taken over the ostensible phase space, *D*. This is highly problematic because almost everywhere in phase space the density measured with respect to the ostensible dimension is zero!

The fact is that you can only calculate the entropy if you know the dimension and the topology of the invariant measure – that object that is preserved by the dynamics. However the dimension is only known approximately and the topology is not known at all.

If we use the phase continuity equation we can attempt to calculated how this fine-grained entropy changes in time for an isokinetic system:

$$
\begin{aligned}
\dot{S}_G(t) &= -k_B \int d\mathbf{\Gamma}\, [1+\ln(f)] \frac{\partial f}{\partial t} \\
&= k_B \int d\mathbf{\Gamma}\, [1+\ln(f)] \frac{\partial}{\partial \mathbf{\Gamma}} \bullet [\dot{\mathbf{\Gamma}} f] \\
&= -k_B \int d\mathbf{\Gamma}\, f \dot{\mathbf{\Gamma}} \bullet \frac{\partial}{\partial \mathbf{\Gamma}} [1+\ln(f)] \qquad (2.6.2) \\
&= -k_B \int d\mathbf{\Gamma}\, \dot{\mathbf{\Gamma}} \bullet \frac{\partial f}{\partial \mathbf{\Gamma}} \\
&= k_B \int d\mathbf{\Gamma}\, f(\mathbf{\Gamma};t) \frac{\partial}{\partial \mathbf{\Gamma}} \bullet \dot{\mathbf{\Gamma}}(\mathbf{\Gamma},t) = -k_B (D_C(N-1)-1) \langle \alpha(t) \rangle
\end{aligned}
$$

In a nonequilibrium steady state the entropy apparently diverges at a constant rate towards negative infinity! If there are no thermostats as in an autonomous Hamiltonian system, the fine grain Gibbs entropy is simply constant.

Now we consider an autonomous Hamiltonian system that has some arbitrary initial distribution. We know from the streaming form of the phase continuity equation (2.4.1) that $f(S^t\mathbf{\Gamma};t) = f(\mathbf{\Gamma};0)$. This is because $f(S^s\mathbf{\Gamma};t) = \exp[-\int_0^t ds\ \Lambda(S^s\mathbf{\Gamma})] f(\mathbf{\Gamma};0)$ and for Hamiltonian dynamics $\Lambda(\mathbf{\Gamma}) \equiv 0$. We use (2.6.1) to calculate the entropy at some time *t*.

$$
\begin{aligned}
S_G(t) &= -k_B \int dS^t\mathbf{\Gamma}\ f(S^t\mathbf{\Gamma};t) \ln[f(S^t\mathbf{\Gamma};t)] \\
&= -k_B \int dS^t\mathbf{\Gamma}\ f(\mathbf{\Gamma};0) \ln[f(\mathbf{\Gamma};0)] \qquad (2.6.3) \\
&= -k_B \int d\mathbf{\Gamma} \left| \frac{\partial S^t \mathbf{\Gamma}}{\partial \mathbf{\Gamma}} \right| f(\mathbf{\Gamma};0) \ln[f(\mathbf{\Gamma};0)] \\
&= -k_B \int d\mathbf{\Gamma}\ f(\mathbf{\Gamma};0) \ln[f(\mathbf{\Gamma};0)] = S_G(0)
\end{aligned}
$$

The first line uses $S^t\mathbf{\Gamma}$ as a dummy integration variable in (2.6.1). The second line uses the streaming phase continuity relation for Hamiltonian systems. The fourth line uses the fact that the Jacobian in line 3 is, for Hamiltonian systems, unity.

Because the entropy is so problematic in nonequilibrium systems, it will play no role in our discussions of nonequilibrium phenomena. We will meet it again when we discuss equilibrium systems in chapter.

## 2.7 AN HISTORICAL ASIDE: COINCIDENCES?

It took 45 years for the Boltzmann equation to be solved. When it was solved, it was solved simultaneously and independently by Chapman and by Enskog in 1917! The solution was obtained by slightly different methods but what is now called the Chapman-Enskog solution of the Boltzmann equation is in fact David Enskog's adaptation of Sydney Chapman's solution.

The first time reversible deterministic algorithm for the linear response of adiabatic planar Couette flow namely the Dolls tensor algorithm, was proposed by Hoover et al. in 1980. The Sllod algorithm for adiabatic planar Couette flow far from equilibrium was proposed by Evans and Morriss in 1984.

Two very similar algorithms that are, for heat flow the analogs of the Dolls tensor for shear, were proposed simultaneously but independently by Gillan and Dixon, and separately by Evans in 1982. The more complex to implement, Gillan-Dixon algorithm, has since been rarely used but the two algorithms are closely related. The Gillan-Dixon algorithm which is exact for the linear response violates AIΓ.

Amazingly, in the same year, 1982, time reversible deterministic thermostats were developed simultaneously but independently by Hoover in California and Evans in Australia. These discoveries (the Gaussian thermostat by Hoover and the Gaussian ergostat by Evans) were made within one week of each other!

In 1990 the first practical connections between chaotic measures and thermophysical properties were made by Evans, Cohen and Morriss with the Conjugate Pairing Rule and independently by Gaspard and Nicolis with the Escape Rate Formalism.

It seems that an idea has its time and when that time comes so does the idea.

## A2 Appendix

Here we consider the slightly tricky issue of computing the exact phase space expansion factor for Gaussian isokinetic dynamics. We treat the isokinetic case because it is a little more tricky than the Nosé-Hoover case.

For simplicity consider a 3 dimensional *N*-particle system obeying the following dynamics:

$$\dot{\mathbf{q}}_i = \mathbf{p}_i \big/ m_i$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i(\mathbf{q}) - \alpha \mathbf{p}_i \qquad \text{(A2.1)}$$

$$\alpha = \frac{\sum_{i=1}^{N} \mathbf{F}_i \bullet \mathbf{p}_i}{\sum_{i=1}^{N} p_i^2},$$

As always the momenta are peculiar so

$$\sum_{i=1}^{N} \mathbf{p}_i = \mathbf{0} \qquad \text{(A2.2)}$$

With this choice for the thermostat multiplier the peculiar kinetic energy is also constant

$$\sum_{i=1}^{N} p(t)_i^2 / 2m = K, \forall t \qquad \text{(A2.3)}$$

The 4 constraints (A2.2,3) imply that the 3*N* Cartesian momentum components are not all independent so one cannot compute the usual phase space expansion factor

$$\Lambda = {\partial}/{\partial \mathbf{\Gamma}} \bullet \dot{\mathbf{\Gamma}}(\mathbf{\Gamma}) = \sum_{i=1}^{N} {\partial}/{\partial \mathbf{p}_i} \bullet -\alpha \mathbf{p}_i \tag{A2.4}$$

where we have assumed AI$\mathbf{\Gamma}$. The difficulty is that, in general, you cannot vary one Cartesian momentum component keeping all other $3N-1$ components fixed and still satisfy the constraint (A2.2).

We resolve this situation by effectively eliminating the degrees of freedom associated with the $N^{\text{th}}$ particle and we compute the phase space expansion factor as,

$$\begin{aligned}
\Lambda &= \sum_{i=1}^{N-1} {\partial}/{\partial \mathbf{p}_i} \bullet -\alpha \mathbf{p}_i \\
&= -(3N-3)\alpha - \sum_{i=1}^{N-1} \mathbf{p}_i \bullet {\partial}/{\partial \mathbf{p}_i}\, \alpha \\
&= -(3N-3)\alpha - \sum_{i=1}^{N-1} \mathbf{p}_i \bullet {\partial}/{\partial \mathbf{p}_i} \frac{\sum_{j=1}^{N-1} \mathbf{F}_j \bullet \mathbf{p}_j + (\sum_{j=1}^{N-1} \mathbf{F}_j) \bullet (\sum_{j=1}^{N-1} \mathbf{p}_j)}{\sum_{j=1}^{N-1} p_j^2 + (\sum_{j=1}^{N-1} \mathbf{p}_j)^2} \\
&= -(3N-3)\alpha - \frac{\sum_{i=1}^{N-1} \mathbf{F}_i \bullet \mathbf{p}_i + (\sum_{i=1}^{N-1} \mathbf{F}_j) \bullet (\sum_{i=1}^{N-1} \mathbf{p}_i)}{2mK} \\
&\quad + 2 \frac{\sum_{i=1}^{N-1} \mathbf{F}_i \bullet \mathbf{p}_i + (\sum_{j=1}^{N-1} \mathbf{F}_j) \bullet (\sum_{j=1}^{N-1} \mathbf{p}_i)}{(2mK)^2} [\sum_{i=1}^{N-1} p_i^2 + (\sum_{j=1}^{N-1} \mathbf{p}_i)^2] \\
&= -(3N-4)\alpha
\end{aligned} \tag{A2.5}$$

In calculating this derivative we still have one constraint. However the two terms involving the partial derivatives: $\sum_{i=1}^{N-1} \partial / \partial \mathbf{p}_i \bullet \mathbf{p}_i$ and $\sum_{i=1}^{N-1} \mathbf{p}_i \bullet \partial / \partial \mathbf{p}_i$ are independent of the value of the peculiar kinetic energy. So although the virtual displacement taken in the derivative violates the kinetic energy constraint, the answer that is computed is independent of the value of the kinetic energy. In fact one could transform to a normalized momentum $\mathbf{p}'_i$ for which the scaled kinetic energy could not vary. The results so obtained are still given by (A2.5) because $\sum_{i=1}^{N-1} \partial / \partial \mathbf{p}_i \bullet \mathbf{p}_i = \sum_{i=1}^{N-1} \partial / \partial \mathbf{p}'_i \bullet \mathbf{p}'_i$ etc.

The same calculation for Nosé-Hoover thermostats in the phase space extended to include the thermostat multiplier $\alpha$ shows that in that case the phase space expansion factor is $-(3N-3)\alpha$ because the second term in the second line of (A2.5) is absent.