CHAPTER 5.

Equilibrium Relaxation Theorems

“One has therefore rigorously proved that, whatever the distribution of kinetic energy at the initial time might have been, it will, after a very long time, always necessarily approach that found by Maxwell”. Boltzmann 1872.

5.1 INTRODUCTION

Understanding the relaxation of systems to equilibrium has been fraught with difficulties. The first reasonably general approach to this problem is the Boltzmann H-theorem. Beginning with the definition of the H-function, Boltzmann proved that the Boltzmann equation for the time evolution of the single particle probability density in a uniform ideal gas, implies a monotonic decrease in the H-function [2, 4, 6] – see the review by Lebowitz [7] for a modern discussion of Boltzmann’s ideas. However, there are at least two problems with Boltzmann’s treatment. Firstly the Boltzmann equation is only valid for an ideal gas – its extension to higher densities has proven to be impossible. Secondly and more problematically, unlike Newton’s equations the Boltzmann equation itself is not time reversal symmetric. It is therefore completely unsurprising that it can be used to derive time asymmetric results.

The early 1930’s saw significant progress in ergodic theory with the proof that if an autonomous Hamiltonian dynamical system is mixing, then in the long time limit, averages of smooth phase functions approach those of the uniform microcanonical distribution. In this Chapter we will give a proof of this result. However this ergodic theory proof reveals almost nothing about the relaxation process itself. The proof only tells us that the relaxation process takes place.

Later we will use the Dissipation Theorem and a corollary of the Evans-Searles Fluctuation Theorem (ESFT), namely the Second Law Inequality, to prove the relaxation to equilibrium of both autonomous Hamiltonian systems and also of such systems in contact with a heat bath. We use these proofs to follow the details of the relaxation process. Finally we prove that the negative logarithm of the microscopic canonical partition function is equal to the thermodynamic Helmholtz free energy divided by the thermodynamic temperature and Boltzmann's constant. Our results complement and extend the findings of modern ergodic theory and show the importance of dissipation in the process of relaxation towards equilibrium.

The results given in this chapter finally resolve the puzzle felt so keenly by R.C. Tolman (1938) concerning Boltzmann's postulate of equal *a priori* probabilities for the equilibrium state of autonomous Hamiltonian systems: "Although we shall endeveavour to show the reasonable character of this hypothesis, it must nevertheless be regarded as a postulate which can be ultimately justified only by the correspondence between conclusions which it permits and the regularities in the behaviour of actual systems which are empirically found." - R. C. Tolman, p 59, The Principles of Statistical Mechanics, Dover, 1979.

5.2 RELAXATION TOWARDS MICROCANONICAL EQUILIBRIUM: THE ERGODIC THEORY APPROACH

It is known from ergodic theory that for a finite, autonomous, Hamiltonian system that is mixing, an arbitrary initial state described by an initial phase space distribution $f(\mathbf{\Gamma};0)$ will eventually relax towards microcanonical equilibrium.

Definition

A system is said to be *mixing* if for integrable phase functions, time correlation functions computed with respect to a stationary distribution factorize into products of averages computed with respect to the same distribution:

$$\lim_{t\to\infty}\left\langle A(\mathbf{\Gamma})B(S^t\mathbf{\Gamma})\right\rangle_\infty - \left\langle A(\mathbf{\Gamma})\right\rangle_\infty \left\langle B(\mathbf{\Gamma})\right\rangle_\infty = 0\ . \qquad (5.2.1)$$

Here the brackets, $\langle ...\rangle_\infty$, denote an ensemble average with respect to an invariant (*i.e.* time-stationary) probability distribution μ_∞. In case μ_∞ has density $f(\mathbf{\Gamma};\infty)$, one may write:

$$\langle A\rangle_\infty = \int d\mu_\infty(\mathbf{\Gamma})A(\mathbf{\Gamma}) = \int d\mathbf{\Gamma}\ f(\mathbf{\Gamma};\infty)A(\mathbf{\Gamma}) \qquad (5.2.2)$$

where $d\mu_\infty = d\mathbf{\Gamma}\ f(\mathbf{\Gamma};\infty)$ is a (dimensionless and normalized) distribution.

If $f(\mathbf{\Gamma};\infty)$ is singular, one would write only the first equality $\langle A\rangle_\infty = \int d\mu_\infty(\mathbf{\Gamma})A(\mathbf{\Gamma})$, where $d\mu_\infty(\mathbf{\Gamma})$ is dimensionless and normalized.

Implicit in this definition is the fact that the invariant measure must be preserved by the dynamics. If it is not, $\lim_{t\to\infty}\left\langle B(S^t\mathbf{\Gamma})\right\rangle_\infty \neq \left\langle B(\mathbf{\Gamma})\right\rangle_\infty$ because by definition, $\mathbf{\Gamma}$ is sampled from

$d\mu_\infty = d\mathbf{\Gamma}\, f(\mathbf{\Gamma};\infty)$ but $S^t\mathbf{\Gamma}$ will be sampled from some other distribution entirely. So mixing systems must, as a prerequisite, have an invariant measure that is preserved by the dynamics *and* additionally they must satisfy (5.2.1) with respect to this invariant distribution or measure.

We note that if the system has nonzero angular momentum no stationary long-time measure is possible (unless we transform to a non-inertial, co-rotating coordinate frame where Hamiltonian dynamics breaks down). So if angular momentum is conserved in our system we set it to zero.

The mixing property is a property of the stationary state of interest, in which observables take the average values denoted by $\langle ... \rangle_\infty$. It represents the fact that, in the macroscopically stationary state, correlations among time evolving physical properties (measured by using averages of smooth phase functions) decay in time. Therefore, in general the mixing condition does not guarantee relaxation to an invariant state. Mixing already *assumes* stationarity of the macrostate, whether it is reached asymptotically in time, as implied by our notation, or it is initially prepared in that state by some means. Only in the special case of autonomous Hamiltonian systems, does mixing actually imply relaxation towards the (microcanonical) stationary state (*cf.* below). This is also a case, for which the density, $f(\mathbf{\Gamma};\infty)$, exists.

Definition

Our version of the standard ergodic theory proof of relaxation for autonomous Hamiltonian systems begins by noting that the *microcanonical distribution*, $f_{\mu c}(\mathbf{\Gamma})$:

$$f_{\mu c}(\mathbf{\Gamma}) \equiv \frac{1}{\int_{H(\mathbf{\Gamma})=E,\mathbf{\Gamma}\in D} d\mathbf{\Gamma}} \qquad (5.2.3)$$

has zero dissipation for autonomous Hamiltonian dynamics and is therefore a time-stationary equilibrium distribution, preserved by the autonomous Hamiltonian dynamics – see (4.2.1). We assume that if our ensemble is somehow distributed according to this naturally invariant distribution, the ensemble of finite systems is *mixing.* We will now prove that if our ensemble is initially *not* distributed according to this distribution, the ensemble will relax towards this distribution - at least for the purposes of computing time averages of smooth physical phase functions.

We compute the time dependent average of an integrable phase function $A(\mathbf{\Gamma})$:

$$\begin{aligned}\langle A \rangle_t &= \int d\mathbf{\Gamma}\, A(\mathbf{\Gamma}) f(\mathbf{\Gamma};t) \\ &= \int d\mathbf{\Gamma}\, A(\mathbf{\Gamma}) f(S^{-t}\mathbf{\Gamma};0) \qquad (5.2.4) \\ &= \int d\mathbf{\Gamma}\, A(S^t\mathbf{\Gamma}) f(\mathbf{\Gamma};0)\end{aligned}$$

where the second and third line follow from the fact that the dynamics is Hamiltonian, hence coordinate changes from $\mathbf{\Gamma}$ to $S^t\mathbf{\Gamma}$ or to $S^{-t}\mathbf{\Gamma}$ have unitary Jacobians. In (5.2.4) stationarity is not assumed. However since the dynamics is driven by an autonomous Hamiltonian, the energy is fixed.

Now we multiply and divide the last expression in (5.2.3) by the (necessarily finite!) volume of the phase space. This casts the first line in a form to which the mixing property can be applied,

$$\begin{aligned}\langle A\rangle_t &= \frac{1}{\int_D d\mathbf{\Gamma}}\bullet\int_D d\mathbf{\Gamma}\, A(S^t\mathbf{\Gamma})f(\mathbf{\Gamma};0)\bullet\int_D d\mathbf{\Gamma} \\ &\equiv \left\langle A(S^t\mathbf{\Gamma})f(\mathbf{\Gamma};0)\right\rangle_{\mu c}\bullet\int_D d\mathbf{\Gamma}\end{aligned}\quad . \qquad (5.2.5)$$

We emphasise that in order to derive (5.2.5) the ostensible phase space volume needs to be finite.

A few more words need to be said about $\left\langle A(S^t\mathbf{\Gamma})f(\mathbf{\Gamma};0)\right\rangle_{\mu c}$. This function is an equilibrium microcanonical, cross-time correlation function. It results from the fact that for Hamiltonian dynamics, any time dependent nonequilibrium ensemble average, say $\langle A\rangle_t$, equals a time dependent nonequilibrium average $\left\langle A(S^t\mathbf{\Gamma}\right\rangle_0$ computed with respect to the initial distribution $f(\mathbf{\Gamma};0)$. It also assumes that we can, to any desired level of accuracy represent the distribution function $f(\mathbf{\Gamma};0)$ as some finely divided histogram of a characteristic function. This allows us to treat the distribution function as though it was in fact a phase function. This is only done at $t=0$.

Using (5.2.1) we now take the long time limit:

$$\begin{aligned}\lim_{t\to\infty}\langle A\rangle_t &= \langle A(\mathbf{\Gamma})\rangle_{\mu c}\langle f(\mathbf{\Gamma};0)\rangle_{\mu c}\bullet\int_D d\mathbf{\Gamma} \\ &= \langle A(\mathbf{\Gamma})\rangle_{\mu c}\frac{1}{\int_D d\mathbf{\Gamma}}\int_D d\mathbf{\Gamma}\, f(\mathbf{\Gamma};0)\bullet\int_D d\mathbf{\Gamma} \quad . \qquad (5.2.6) \\ &= \langle A(\mathbf{\Gamma})\rangle_{\mu c}.1 = \langle A(\mathbf{\Gamma})\rangle_{\mu c}\end{aligned}$$

We have used the mixing assumption (5.2.1), to allow us to factorize the naturally invariant (microcanonical) time correlation function into a product of two invariant (microcanonical) averages. Lastly, we use the normalization of the initial distribution function. Note: we do not need to *assume* the existence of a stationary state, since the microcanonical distribution is indeed invariant for Hamiltonian dynamics. This is because (as noted above already) the dissipation function is identically zero for autonomous Hamiltonian dynamics with an initial ensemble being the uniform microcanonical distribution.

So $\langle A \rangle_t$ tends to a microcanonical average, whatever smooth phase function $A(\mathbf{\Gamma})$, or initial probability density $f(\mathbf{\Gamma};0)$ one considers - as long as it lies on an energy hypersurface. By definition, this amounts to a proof of relaxation *towards* the microcanonical equilibrium state denoted by $\langle . \rangle_{\mu C}$.

There is another crucial requirement in the proof above: the proof cannot be extended to thermostatted dissipative systems because the asymptotic steady state would be singular, and have no smooth density. If the system is subject to a time independent dissipative external field and no thermostat is applied, then the total Hamiltonian is constant and if the system is mixing, the derivation above still applies. Finite mixing Hamiltonian systems, regardless of whether external or only internal fields are applied, ultimately relax towards microcanonical equilibrium.

Unless one starts at *t*=0 with the microcanonical distribution, this proof shows only that *averages* of smooth thermodynamic quantities *approach* microcanonical averages in the long time limit. The actual distribution never *becomes* the microcanonical distribution. At any time no matter how large, we can always apply a time reversal map and return (eventually!) to the initial distribution. As time increases in the relaxation process, the long time distribution function becomes ever more tightly folded upon itself, never *becoming* the smooth microcanonical equilibrium distribution.

If it were in fact to eventually become *precisely* the microcanonical distribution one could never return to the initial distribution by applying a time reversal map. This gives a proof that the relaxation process cannot be complete in finite time.

5.3 RELAXATION OF AUTONOMOUS HAMILTONIAN SYSTEMS UNDER T-MIXING

Definition

From the definition of the dissipation function it is trivial to see that if the states are distributed as,

$$
\begin{aligned}
f_{\mu c}(\mathbf{\Gamma}) &= \frac{\delta(H_0(\mathbf{\Gamma})-E)\delta(\mathbf{P})\delta(\mathbf{L})}{\int d\mathbf{\Gamma}\ \delta(H_0(\mathbf{\Gamma})-E)\delta(\mathbf{P})\delta(\mathbf{L})} \\
&= \frac{1}{\int d\mathbf{\Gamma}}\ if\ \mathbf{\Gamma}\in D, \quad =0\ if\ \mathbf{\Gamma}\notin D
\end{aligned}
\qquad (5.3.1)
$$

the dissipation function is identically zero, everywhere in ostensible phase space, D , which has a fixed energy E and zero linear, $\mathbf{P}$, and angular momentum, $\mathbf{L}$. Basically the energy, linear and angular momenta are all constants of the motion and the phase space expansion factor is also zero.

The distribution function in equation (5.3.1) is therefore an equilibrium distribution function. It is referred to as the equilibrium micro-canonical distribution ($f_{\mu c}(\mathbf{\Gamma})$). Within this ostensible domain D , T-mixing systems have no non-trivial constants of the motion. Later we will prove this statement from the T-mixing definition[11]. Of course if the particular Hamiltonian, with which we are dealing, contains more symmetries than those discussed here, there will be additional nontrivial constants of the motion. These should be handled by inserting additional delta functions into the microcanonical distribution (5.3.1) so that the ostensible phase space is constrained to a fixed value for these additional constants of the motion.

Mixing is closely related to, but subtly different from, the T-mixing condition, which we introduced in slightly different versions to obtain correspondingly different results[15]. In particular, to prove steady state fluctuation relations, only the weak and hence very general form of T-mixing is required. This proves how general the steady state fluctuation relation for the dissipation function[11,15] actually is. On the other hand the Equilibrium Relaxation Theorems require sufficiently fast rates of correlation decay. Below this is made part of the definition of T-mixing itself[11].

Definition

The *T-mixing condition* assumes that for a real sufficiently smooth phase function $A(\mathbf{\Gamma})$:

$$\int_0^\infty ds \left\langle \Omega(\mathbf{\Gamma}) A(S^s \mathbf{\Gamma}) \right\rangle_0 = L_0 \in \Re \qquad (5.3.2)$$

i.e. that L_0 is real and finite, where $\Omega(\mathbf{\Gamma})$ is the instantaneous dissipation at the phase $\mathbf{\Gamma}$ and $A(S^s \mathbf{\Gamma})$ is the phase function $A(\mathbf{\Gamma})$ evaluated at the time evolved phase $S^s \mathbf{\Gamma}$. In contradistinction to the well known mixing condition of ergodic theory, the T-mixing condition considers time correlation functions referred to the *initial* state, here denoted by $\langle . \rangle_0$, where the distribution of phases is usually known.

Definition

The *weak T-mixing condition* that looks very similar to the mixing condition, is that,

$$\lim_{t\to\infty} [\left\langle A(\mathbf{\Gamma}) B(S^t \mathbf{\Gamma}) \right\rangle_0 - \left\langle A(\mathbf{\Gamma}) \right\rangle_0 \left\langle B(S^t \mathbf{\Gamma}) \right\rangle_0] = 0 \qquad (5.3.3)$$

where $A(\mathbf{\Gamma}),B(\mathbf{\Gamma})$ are any integrable phase functions. The main difference, between weak T-mixing (5.3.3), and standard mixing (5.2.1), lies in the fact that the second factor in the second term inside square brackets in (5.3.3), is not time *independent*. It takes the form $\left\langle B(S^t\mathbf{\Gamma})\right\rangle_0 = \left\langle B(\mathbf{\Gamma})\right\rangle_t$, hence it cannot be taken out of the limit. This time dependence is a reflection of the fact that the ensemble averages in (5.3.3) are taken with respect to the initial distribution rather than an invariant long-time distribution.

For weak T-mixing (5.3.3) and T-mixing (5.3.2), the relevant probability distribution in (5.3.2,3) is not the invariant one; it is the initial ensemble $d\mu_0(\mathbf{\Gamma}) = d\mathbf{\Gamma}\ f(\mathbf{\Gamma};0)$, whose averages are denoted by $\langle . \rangle_0$. Mixing (5.2.1) and weak T-mixing (5.3.3) do not say anything about the rate of convergence to a stationary state or even whether such convergence actually occurs.

Throughout the rest of this book if we use the term “T-mixing” we are referring to the condition given in equation (5.3.2). If we refer to “mixing”, we are discussing mixing given by equation (5.3.2). If we discuss “weak T-mixing” we are referring only to the condition given in equation (5.3.3).

We obviously exclude the constants of the motion inherent in the Hamiltonian symmetries from being possible phase functions in (5.3.2) (*i.e.* $A(\mathbf{\Gamma}),B(\mathbf{\Gamma}) \notin H_0(\mathbf{\Gamma}),P_\alpha(\mathbf{\Gamma}),L_\alpha(\mathbf{\Gamma}),\alpha = x,y,z$), since each of these variables is obviously a constant of the motion. So our ostensible phase space domain D is some specified physical volume on a zero linear and angular momentum energy hypersurface. The zero linear momentum condition could be relaxed but the total angular momentum must be fixed at zero.

If the space is orientationally isotropic, the total angular momentum is a constant of the motion and for reasons that are rather obvious the system cannot possibly be T-mixing

(5.3.2). When viewed from an inertial coordinate frame, the measure required for mixing (5.2.1) cannot be time invariant but rather will be periodic. Likewise the integrals required for the T-mixing property (5.3.2) will not in general converge but may also be periodic functions of the integration time. Rotating systems may however be weak T-mixing (5.3.3).

In a T-mixing system there can be no non-trivial constants of the motion other than those inherent in the Hamiltonian symmetries. If there were such constants we could form transient time correlation functions that violated equations (5.3.2) and (5.3.3). The fixed values of the various constants of the motion must be chosen to provide an inertial coordinate frame, within which we can construct a Hamiltonian dynamical system.

All T-mixing systems are physically ergodic over the ostensible phase space because if the phase space broke up into nonintersecting phase space subdomains characterised by different macroscopic averages for smooth phase functions, we could form constants of the motion depending on whether a system was on one subdomain or another. These subdomain occupation numbers could then be substituted as $A(\mathbf{\Gamma})$ in (5.3.2) thereby violating the T-mixing condition.

If the relevant time correlation functions (5.3.2) decay asymptotically as t^{-1} or more slowly, the system may be weak T-mixing (5.3.3) but cannot be T-mixing (5.3.2). In contradistinction to mixing (5.2.1), if a system is T-mixing (5.3.2), it *must* relax to a time-stationary state at long times, whether this state is characterized by a smooth probability density $f(\mathbf{\Gamma};\infty)$ or not. If a system is weak T-mixing, but not T-mixing, relaxation to an invariant state from a non-invariant initial state, will not occur.

In general it is exceedingly difficult to prove that a given system is mixing and perhaps even harder to prove whether it is T-mixing. However, because of the many properties of T-mixing systems it is easy to perform numerical/experimental tests of whether a system is T-mixing.

We now give a proof of relaxation to the stationary state based on the strong T-mixing condition (5.3.2). From the T-mixing assumption, there can be no constants of the motion other than the trivial ones, the internal energy, H_0 and the linear and angular momenta, $\mathbf{P},\mathbf{L}$ which are assumed to take on fixed values of $E,\mathbf{0},\mathbf{0}$, respectively.

If we consider *any* deviation from the microcanonical form (5.3.1) generated by a real-valued smooth deviation function, $g(\mathbf{\Gamma})$ that is even in the momenta and differentiable,

$$f_g(\mathbf{\Gamma}) = \frac{\exp[-g(\mathbf{\Gamma})]\delta(H_0(\mathbf{\Gamma})-E)\delta(\mathbf{P})\delta(\mathbf{L})}{\int d\mathbf{\Gamma}\,\exp[-g(\mathbf{\Gamma})]\delta(H_0(\mathbf{\Gamma})-E)\delta(\mathbf{P})\delta(\mathbf{L})}\ , \qquad (5.3.4)$$

the dissipation function will not vanish and we would have:

$$\Omega(\mathbf{\Gamma}) = \dot{g}(\mathbf{\Gamma}) \qquad (5.3.5)$$

where $\dot{g}(\mathbf{\Gamma}) \equiv \dot{\mathbf{\Gamma}} \bullet \partial g(\mathbf{\Gamma}) / \partial \mathbf{\Gamma}$ denotes the time derivative.

Since the system is T-mixing if $g(\mathbf{\Gamma}) \neq 0$, then $\Omega(\mathbf{\Gamma}) \neq 0$ because $g(\mathbf{\Gamma})$ cannot be a constant of the motion and the strict form of the Second Law Inequality applies.

Definition

In (5.3.4) the real phase function, even in the momenta, namely $g(\mathbf{\Gamma})$, is termed a *deviation function*.

The strict Second Law Inequality[3] states that the ensemble average of the time integral of the dissipation from 0 to some time t is positive for all values of t. It is only equal to zero if the system is at equilibrium and $g(\mathbf{\Gamma}) = 0, \forall \mathbf{\Gamma}$. Thus for finite values of the deviation function g we have,

$$\left\langle \Omega_t \right\rangle_0 = \left\langle g(S^t\mathbf{\Gamma}) - g(\mathbf{\Gamma}) \right\rangle_0 \equiv \left\langle \Delta g(t) \right\rangle_0 > 0, \quad g(\mathbf{\Gamma}) \neq 0, \forall t > 0 . \qquad (5.3.6)$$

Thus if there is *any* deviation from the equilibrium distribution (5.3.1), the dissipation function will not vanish (because there are no other constants of the motion) and further, the ensemble average of the time integrated dissipation function *must* be positive. In fact

$$\left\langle \Delta g(t) \right\rangle_0 = \int_0^\infty dA\, A(1 - e^{-A}) p(\Delta g(t) = A) > 0, \quad \forall t, g(\mathbf{\Gamma}) \neq 0 \qquad (5.3.7)$$

If there is any non-zero dissipation $\Delta g(t) \neq 0$ the ensemble averaged change in dissipation $\left\langle \Delta g(t) \right\rangle_0$ must be greater than zero. This means that for T-mixing systems, the equilibrium distribution function is unique and given by eq.(5.3.1).

One can prove that the system must relax towards equilibrium by using the T-mixing property (5.3.2) and the Dissipation Theorem for the deviation function itself.

$$\begin{gathered} \lim_{t\to\infty} \left\langle g(t) \right\rangle_0 = \left\langle g(0) \right\rangle_0 + \int_0^t ds \left\langle \dot{g}(0) s(s) \right\rangle_0 = const \\ \Rightarrow \lim_{t\to\infty} \left\langle \dot{g}(t) \right\rangle_0 = \lim_{t\to\infty} \left\langle \Omega(t) \right\rangle_0 = 0 \end{gathered} \qquad (5.3.8)$$

Thus for the T-mixing systems treated here in the long time limit the ensemble averaged instantaneous dissipation is zero. Using (3.7.4) we see that

$$\lim_{t\to\infty} \Omega(S^t\mathbf{\Gamma}) = 0, \forall \mathbf{\Gamma} \in D \qquad (5.3.9)$$

because if there is any dissipation any where in the system $\langle \Omega(t) \rangle > 0$. This means that from the point of view of calculating averages of smooth phase functions like the dissipation, the system must be relaxing *towards* the unique equilibrium state (5.3.1), which obviously has zero dissipation. Of course for any time no matter how large, the fine-grained phase space distribution is never given by the equilibrium state (5.3.1), because if this did happen the system could never return to the initial distribution after the application of a time reversal map.

This implies that for T-mixing autonomous Hamiltonian systems the relaxation to true thermal equilibrium *must* take an infinite amount of time. Relaxation to equilibrium *cannot* occur in a finite time!

5.4 THERMAL RELAXATION TO EQUILIBRIUM: THE CANONICAL ENSEMBLE

Consider a classical system of N interacting particles in a volume V. The microscopic state of the system is represented by a phase space vector of the coordinates and momenta of all the particles, $\{\mathbf{q}_1,\mathbf{q}_2,..\mathbf{q}_N,\mathbf{p}_1,..\mathbf{p}_N\}\equiv(\mathbf{q},\mathbf{p})\equiv\mathbf{\Gamma}$ where $\mathbf{q}_i,\mathbf{p}_i$ are the position and momentum of particle, i. Initially (at $t = 0$), the microstates of the system are distributed according to a normalized probability distribution function $f(\mathbf{\Gamma};0)$. To apply our results to realistic systems, we separate the N particle system into a system of interest and a wall region containing N_W particles. Within the wall a subset of N_{th}, particles is subject to a fictitious thermostat. The thermostat employs a switch, S_i, 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

$$K_{th} \equiv \sum_{i=1}^{N} S_i \frac{p_i^2}{2m_i}, \tag{5.4.1}$$

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

$$\begin{aligned}
\dot{\mathbf{q}}_i &= \mathbf{p}_i / m_i \\
\dot{\mathbf{p}}_i &= \mathbf{F}_i(\mathbf{q}) - S_i(\alpha \mathbf{p}_i + \boldsymbol{\gamma}_{th}) \qquad (5.4.2) \\
\dot{\alpha} &= \left[\frac{2K_{th}}{3(N_{th}-1)k_B T_{th}} - 1 \right] \frac{1}{\tau^2},
\end{aligned}$$

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, $-S_i\alpha\mathbf{p}_i$ 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 α, T_{th} is the *target* parameter that controls the time averaged kinetic energy of the thermostatted particles and τ is the time constant for the Nosé-Hoover thermostat. The 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.

Note 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 $\mathbf{\Gamma}^* \equiv (\mathbf{\Gamma},\alpha)$ and from here on represent this implicitly using $\mathbf{\Gamma}$. In the absence of the thermostatting terms the (Newtonian) equations of motion preserve the phase space volume so the system satisfies AI$\mathbf{\Gamma}$ [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 (5.4.1,2) are time reversible and heat can be either absorbed or given out by the thermostat.

For the Nosé-Hoover dynamics (5.4.1,2), consider the initial distribution

$$f(\mathbf{\Gamma};0) \equiv f_c(\mathbf{\Gamma}) = \frac{\delta(\mathbf{p}_{th})\exp[-\beta_{th}H_E(\mathbf{\Gamma})]}{\int_D d\mathbf{\Gamma}\ \delta(\mathbf{p}_{th})\exp[-\beta_{th}H_E(\mathbf{\Gamma})]}, \quad \forall\mathbf{\Gamma}\in D\,, \tag{5.4.3}$$

where $H_0(\mathbf{\Gamma})$ is the internal energy of the system, $H_E(\mathbf{\Gamma}) = H_0(\mathbf{\Gamma}) + \tfrac{3}{2}(N_{th}-1)k_B T_{th}\alpha^2\tau^2$ is the so called extended Nosé-Hoover Hamiltonian, $k_B T_{th} \equiv \beta_{th}^{-1}$, and $\delta(\mathbf{p}_{th}) \equiv \delta\left(\sum S_i p_{xi}\right)\delta\left(\sum S_i p_{yj}\right)\delta\left(\sum S_i p_{zk}\right)$ fixes the total momenta of the thermostatted particles in each Cartesian dimension, at zero.

Definition

We shall call the distribution in (5.3.3) the *canonical distribution* even though it includes extra degrees of freedom for the thermostat multiplier α.

It is easy to show that for this distribution (5.4.3) and the dynamics (5.4.1,2) the dissipation function, $\Omega_c(\mathbf{\Gamma})$, is identically zero at all points sampled by the canonical distribution

$$\Omega_c(\mathbf{\Gamma}) = 0, \quad \forall\, \mathbf{\Gamma} \in D \qquad (5.4.4)$$

Proof:

From (5.4.3) and the definition of the dissipation function we see that (see Appendix A2 for how to evaluate Λ exactly),

$$\Omega_{c,t}(\mathbf{\Gamma}(0)) = \beta_{th}[H_E(\mathbf{\Gamma}(t)) - H_E(\mathbf{\Gamma}(0))] + 3(N_{th}-1)\int_0^t ds\, \alpha(s)\,. \qquad (5.4.5)$$

Now from the definition of the extended Hamiltonian and the equations of motion we see that if we take the time derivative of (5.4.5) we obtain,

$$\Omega_c = \beta_{th}[-2K_{th}\alpha + 3(N_{th}-1)k_B T_{th}\alpha\dot{\alpha}\tau^2] - (3N_{th}-1)\alpha\,. \qquad (5.4.6)$$

Now using the equation of motion for the thermostat multiplier we see that,

,

$$\begin{aligned}
\Omega_c &= \beta_{th}[-2K_{th}\alpha + 3(N_{th}-1)k_B T_{th}\alpha \frac{2K_{th} - 3(N_{th}-1)k_B T_{th}}{3(N_{th}-1)k_B T_{th}}] + 3(N_{th}-1)\alpha \\
&= \beta_{th}[-2K_{th}\alpha + \alpha[2K_{th} - 3(N_{th}-1)k_B T_{th}]] + 3(N_{th}-1)\alpha \qquad , \qquad (5.4.7) \\
&= 0
\end{aligned}$$

where we have used the fact that $k_B T_{th} \equiv \beta_{th}^{-1}$. We note that in the proof we are using exact calculations. Often approximations that are only valid in the large *N* limit are used in statistical mechanics. This calculation is exact for arbitrary *N*.

We know from §4.2 that this initial (equilibrium) distribution is preserved by the dynamics (5.4.1,2)

$$f(\mathbf{\Gamma},t) = f_c(\mathbf{\Gamma}), \quad \forall\, \mathbf{\Gamma} \in D, \forall t\,. \qquad (5.4.8)$$

Since we know that (5.4.3) is an equilibrium distribution for the dynamics we consider and since we also know that T-mixing systems are physically ergodic we know from §4.3 that (5.4.3), is the unique equilibrium distribution for this system. However because of the importance of this point we will explore the matter in greater detail.

Consider an arbitrary deviation from the canonical distribution

$$f(\mathbf{\Gamma};0) \equiv \frac{\delta(\mathbf{p}_{th})\exp[-\beta_{th}H_E(\mathbf{\Gamma}) - \gamma g(\mathbf{\Gamma})]}{\int_D d\mathbf{\Gamma}\; \delta(\mathbf{p}_{th})\exp[-\beta_{th}H_E(\mathbf{\Gamma}) - \gamma g(\mathbf{\Gamma})]\int}\;, \qquad (5.4.9)$$

where $g(\mathbf{\Gamma})$ is an arbitrary integrable real deviation function and since $f(\mathbf{\Gamma};0)$ must be an even function of the momenta, $g(\mathbf{\Gamma})$ must also be even in the momenta. Without loss of generality we assume $0 \le \gamma$. The factor γ is a scale parameter that we can use to control the magnitude of the deviation from equilibrium.

For such a system (5.4.9) evolving under our dynamics (5.4.1,2), the instantaneous dissipation function is

$$\Omega(\mathbf{\Gamma}) = \gamma \, {}^{\partial g(\mathbf{\Gamma})}\!\big/_{\partial \mathbf{\Gamma}} \bullet \dot{\mathbf{\Gamma}}(\mathbf{\Gamma}) = \gamma \, {}^{dg(\mathbf{\Gamma})}\!\big/_{dt} \qquad (5.4.10)$$

Since $g(\mathbf{\Gamma})$ is even in the momenta we know that

$$\left\langle \Omega(0) \right\rangle_g = \left\langle \gamma \dot{g}(\mathbf{\Gamma}) \right\rangle_g = 0 \qquad (5.4.11)$$

where the subscript g on the ensemble average denotes the fact that the average is carried out over the initial distribution (5.4.9).

Now (5.4.10) implies

$$f(\mathbf{\Gamma};t) = \exp[-\gamma \Delta g(\mathbf{\Gamma},-t)] f(\mathbf{\Gamma};0) \, . \qquad (5.4.12)$$

where $\Delta g(\mathbf{\Gamma},t) \equiv g(S^t \mathbf{\Gamma}) - g(\mathbf{\Gamma})$. Because the system is T-mixing there can be no constants of the motion additional to those specified in (5.3.3). Thus, if $g(\mathbf{\Gamma}) \neq 0$ there must be dissipation and the distribution function cannot be a time independent equilibrium distribution. Thus the equilibrium distribution given by (5.4.3) is unique.

Summarising: since the system is T-mixing there is a *unique* time symmetric, equilibrium state characterized by being dissipationless everywhere in the phase space domain D. For the system considered here that distribution is the canonical distribution (5.4.3). Thus we have derived an expression for the unique equilibrium state corresponding to the thermostatted equations of motion and shown that it takes on the standard form for the canonical distribution, modulo the facts that: in the thermostatting region the momentum is a constant of the motion that is set to zero, and that there is an extended degree of freedom for the thermostat.

The dissipation function satisfies the strict Second Law Inequality

$$\begin{aligned} \gamma \left\langle \Delta g(\mathbf{\Gamma},t) \right\rangle_g &= \int_0^\infty dA\, A(1-e^{-A}) p[\gamma\Delta g(\mathbf{\Gamma},t)=A] \\ &> 0, \end{aligned} \qquad . \qquad (5.4.13)$$

If $p[\gamma\Delta g(\mathbf{\Gamma},t)=A]$ is non-zero for *any A > 0*, then $p[\gamma\Delta g(\mathbf{\Gamma},t)=A]A(1-e^{-A})>0$ and the integrand in (5.4.13), as well as the integral will be strictly positive. Thus in a T-mixing system if the initial distribution differs in any way from the canonical distribution there will be dissipation and the ensemble average of the time integral of the dissipation is *positive*. This remarkable result is true for arbitrary $\gamma,\ g(\mathbf{\Gamma})$.

If we start the system at time zero from a nonequilibrium distribution (5.4.9) we can ask the question how does the ensemble average of the deviation function change with time. Substituting (5.4.13) into (5.4.12) gives:

$$\left\langle \Delta g(\mathbf{\Gamma},t) \right\rangle_g = \gamma \int_0^t ds \left\langle \dot{g}(0) g(s) \right\rangle_g > 0, \quad \forall t > 0 \qquad (5.4.14)$$

where the ensemble averages are taken with respect to the initial nonequilibrium distribution function (5.4.9).

Because the initial distribution is an even function of the momenta we know from (5.4.11) that the transient time correlation function appearing in (5.4.14) can be regarded as involving the product of two zero mean phase variables (see (4.3.5)), and the T-mixing condition can be directly applied to the correlation function. Applying the T-mixing condition shows that the time integral on the right hand side of (5.5.1) converges as $t \to \infty$. This implies that:

$$\lim_{t\to\infty}\left|\langle \Delta g(\mathbf{\Gamma},t+\tau)\rangle_g - \langle \Delta g(\mathbf{\Gamma},t)\rangle_g\right| = 0 \qquad (5.4.15)$$

This means that as t becomes ever larger, the dissipation over a fixed time interval τ, becomes ever smaller and in the long time limit, the ensemble averaged instantaneous dissipation vanishes. This implies the system is relaxing towards its unique equilibrium state and

$$\lim_{t\to\infty}\gamma\langle \dot{g}(t)\rangle_{f(\mathbf{\Gamma},0)} = \lim_{t\to\infty}\langle \Omega(t)\rangle_{f(\mathbf{\Gamma},0)} = 0 \qquad (5.4.16)$$

Equation (5.5.3) follows by differentiating (5.5.2) with respect to τ and then letting t increase without bound.

We have therefore proved that subject to the conditions stated above, arbitrary initial nonequilibrium distributions eventually relax, perhaps not monotonically, towards equilibrium. As in the microcanonical case this relaxation process cannot be completed in finite time.

From (5.4.2) and (4.1.5) we see that

$$\left\langle B(t)\right\rangle_g = \left\langle B(0)\right\rangle_g + \gamma \int_0^t ds \left\langle \dot{g}(0)B(s)\right\rangle_g , \qquad (5.4.17)$$

and substituting $\dot{g}(\mathbf{\Gamma}) = B(\mathbf{\Gamma})$ we see that

$$\lim_{t\to 0^+} \left\langle \dot{g}(t)\right\rangle_g = \gamma \left\langle \dot{g}^2(0)\right\rangle_g > 0 \ . \qquad (5.4.18)$$

This proves that *initially*, on average, the system always moves towards, rather than away from, equilibrium. At later times the system may move, for a short time, away from equilibrium (*e.g.* as in the case of an under-damped oscillator) but such movement is never enough to make the time integrated ensemble average dissipation negative (or even zero). The time integrated average dissipation from the initial state to any intermediate state (including the final equilibrium state) is strictly positive. At any sufficiently later instant in the relaxation process, the *instantaneous* dissipation may be negative. This shows that, in general, the relaxation process may not be monotonic in time. Such non-monotonic relaxation is extremely common in Nature.

Definition

From (5.4.9) we see that if the perturbation relaxes *conformally* in phase space (*i.e.* the deviation function simply scales by a single time dependent parameter $\gamma(t)$), we see that in order to increase the value of $\left\langle g(t)\right\rangle_g$ we must decrease the magnitude of the scale parameter γ, implying that the system is moving closer to the equilibrium distribution. This also means that we can continuously redefine a new time origin. In this case equation (5.4.18) implies there is a monotonic relaxation to equilibrium. Conformal relaxation processes occur when for example when the deviation function corresponds to the decay in the slowest possible hydrodynamic mode available to a system. In this case the only possible relaxation process is conformal.

Since we now know that under the conditions specified here the system will at long times relax towards its unique equilibrium state, we therefore know the following:

$$\begin{aligned}\lim_{t\to\infty}\left\langle \dot{g}(t)\right\rangle_g &= \left\langle \dot{g}(0)\right\rangle_g + \gamma\int_0^\infty ds\,\left\langle \dot{g}(0)\dot{g}(s)\right\rangle_g \\ &= \gamma\int_0^\infty ds\,\left\langle \dot{g}(0)\dot{g}(s)\right\rangle_g = 0\end{aligned} \quad , \qquad (5.4.19)$$

where the first term on the right hand side of the top line is zero by (5.4.3) and the subscript zeros signify that the initial ensemble is given by (5.4.9). Equation (5.4.19) is true for any deviation function that is even in the momenta.

Definition

We call (5.4.19) the *heat death equation*. It shows that for systems arbitrarily far from equilibrium initially, the infinite time integral of the transient autocorrelation function of fluxes of nonconserved quantities vanishes.

If we take the weak deviation limit where $\gamma \to 0$ we see that equilibrium time auto correlation functions of fluxes of nonconserved quantities also vanish

$$\begin{aligned}\lim_{\gamma\to 0}\lim_{t\to\infty}\left\langle \dot{g}(t)\right\rangle_g &= \left\langle \dot{g}(0)\right\rangle_{eq} + \gamma\int_0^\infty ds\,\left\langle \dot{g}(0)\dot{g}(s)\right\rangle_{eq} \\ &= \gamma\int_0^\infty ds\,\left\langle \dot{g}(0)\dot{g}(s)\right\rangle_{eq} = 0\end{aligned} \qquad (5.4.20)$$

This equation was first written down in 1963 by Zwanzig and has been called the ZBBR equation.

Equation (5.4.18) implies another important point. Occasionally one sees in the literature the correct observation that dissipative systems have phase space trajectories that are more stable than their time reversed anti-dissipative conjugates – see for example William Thomson quote, Chapter 4 or p247 of “Time Reversibility, Computer Simulation and Chaos” by W G Hoover. This comment on the relative mechanical stability is easily seen to be correct because if we consider a nonequilibrium steady state the sum of all the Lyapunov exponents must be negative. This implies that for systems satisfying the Conjugate Pairing Rule, that the largest positive exponent for a steady state is smaller in magnitude than the largest positive exponent for an anti-steady state. This is obvious because the largest positive exponent for an antisteady state is -1 times the value of the most negative exponent of a steady state and the sum of the extremal exponents for a steady state must be negative. For systems that don’t satisfy the Conjugate Pairing Rule, the Kolmogorov-Sinai entropy for the antisteady state is greater than that for the steady state.

However this difference in the relative stability of steady states and antisteady state trajectories has nothing directly to do with the Second “Law” of Thermodynamics being satisfied. Equation (5.4.18) shows that on average systems respond *immediately* in a direction favoured by the Second “Law”. They do not rely on the slow build-up of instabilities before they begin to satisfy the Second “Law”.

In summary we have demonstrated that for any T-mixing Hamiltonian system of fixed volume and fixed number of particles, in contact with a heat reservoir whose initial (nonequilibrium) distribution is even under time reversal symmetry:

- there is a *unique* dissipationless state, and this state has the canonical distribution, [Although a Nosé-Hoover thermostat was used in this derivation, essentially the same result is obtained with other thermostatting mechanisms such as a Gaussian isokinetic thermostat];
- in T-mixing systems with decaying temporal correlations the system relaxes towards canonical equilibrium;
- this relaxation towards equilibrium is *not* necessarily monotonic [We note that the Boltzmann H-theorem applied to uniform dilute gases, implies a monotonic relaxation to equilibrium, thus the Relaxation Theorem allows for much more complex behaviour as seen experimentally];
- the time integrated ensemble averaged dissipation satisfies the strict inequality: $\left\langle \Delta g(\mathbf{\Gamma},t) \right\rangle_g > 0$,
- if the deviation function relaxes conformally, the system will relax to equilibrium monotonically;
- the *initial* ensemble average response is always towards, rather than away from, equilibrium;
- the relaxation process cannot take place in finite time.

We have also shown quite generally that for T-mixing dynamical systems obeying time reversible dynamics, states have properties that are time reversal symmetric (*i.e.* probabilities of observing any set of trajectories and its conjugate set of antitrajectories are equal) if and only if the dissipation function is zero everywhere in phase space. If there is dissipation anywhere in the phase space the distribution function is not time independent and the system cannot be in equilibrium.

5.5 RELAXATION TO QUASI-EQUILIBRIUM FOR NONERGODIC SYSTEMS

If the system is not T-mixing over the full domain *D,* the system may split into non-ergodic subdomains, D_i i=1,2,….. each characterized by different ensemble averages for physical properties. If these states are individually T-mixing then the two relaxation theorems given above (for Hamiltonian systems and for such systems in contact with a heat reservoir) still apply individually to each subdomain. The systems will still relax to either microcanonical or canonical equililibrium within each subdomain. Examples of such systems are relatively common for example solid glassy systems. Many solid systems are not really completely relaxed to true thermodynamic equilibrium; their macroscopic physical properties are history dependent – *e.g.* work hardened metals or metals that are rapidly cooled.

None of these systems are T-mixing over the ostensible phase space. However most are expected to be T-mixing over the history dependent phase space subdomains within which these solid samples are trapped. The topology of these subdomains can be incredibly complex! However whatever the topology, we expect that the vast majority of such solids to be T-mixing over whatever subdomain a particular solid sample is trapped. Depending on whether such a system is in contact with a thermal reservoir or not, at long times such systems will relax towards microcanonical or canonical equilibrium within their particular phase space subdomain.

Quasi-equilibrium is very common in solids because their physical properties (essentially infinite shear viscosities and very low diffusion coefficients) mean that the full exploration of phase space is kinetically restricted.

5.6 ASIDE: THE THERMODYNAMIC CONNECTION

This section is not necessary for the logical exposition of this book. It is included for those who already know classical thermodynamics. Statistical mechanics has been traditionally taught *assuming* the "laws" of thermodynamics. As we will see later in this book the Zeroth and Second Laws will be proved from mechanics, in §7.2 and §8.5 respectively.

We give a proof here that the microscopic expressions defined below (5.6.2 and 5.6.3,5.6.10) on average are indeed equal to the thermodynamic entropy and temperature respectively. We take as our starting point, known expressions for the Galilei invariant energy and pressure to equal, on average, their thermodynamic counterparts. Energy and pressure are, as the first "law" of thermodynamics makes clear, completely mechanical in nature.

To begin we note that from thermodynamics we have two equations for the entropy (*S*), in terms of the energy (*U*), the volume (*V*), and the pressure (*p*):

$$\left.\frac{\partial S}{\partial U}\right|_V = \frac{1}{T},$$
$$\left.\frac{\partial S}{\partial V}\right|_U = \frac{p}{T} \qquad (5.6.1)$$

Consider the function $\tilde{S}$ defined (up to an additive constant) as:

$$\tilde{S} = k_B \ln \int \delta(H_0(\mathbf{\Gamma}) - U)\delta(\mathbf{P})d\mathbf{\Gamma} \equiv k_B \ln V_{\mathbf{\Gamma}} \qquad (5.6.2)$$

We can identify the internal energy, U with the value of the Hamiltonian in a co-moving coordinate frame, $H_0(\mathbf{\Gamma})$, because internal energy is the Galilei invariant mechanical energy.

Consider a phase vector displacement in phase space $\mathbf{\Gamma}' = \mathbf{\Gamma} + d\mathbf{\Gamma}$ where $d\mathbf{\Gamma} = dU(\nabla_{\mathbf{p}} H)/(\nabla_{\mathbf{p}} H \bullet \nabla_p H)$ and $\nabla_{\mathrm{p}} ... \equiv ({}^{\partial}\!/\!_{\partial \mathrm{p}_{\mathrm{x}1}}, {}^{\partial}\!/\!_{\partial \mathrm{p}_{\mathrm{y}1}}, ..., {}^{\partial}\!/\!_{\partial \mathrm{p}_{\mathrm{zN}}}, 0,0,...,0)...$ that is normal to the kinetic energy hypersurface $\sum_i p_i^2 / 2m = K_0$, and that (to leading order in *N*) changes the energy of any phase point $\mathbf{\Gamma}$ by a constant infinitesimal amount dU. Since the

Jacobian of the transformation $J(\mathbf{\Gamma}) = \left| \partial \mathbf{\Gamma}' / \partial \mathbf{\Gamma} \right| = 1 + dU \, \nabla^2_{\mathbf{p}} H / (\nabla_{\mathbf{p}} H \bullet \nabla_{\mathbf{p}} H)$ it can be seen from (5.6.2) that [i]:

$$\left. \frac{\partial \tilde{S}}{\partial U} \right|_V \equiv \frac{1}{\tilde{T}} = \frac{3Nk_B}{2\langle K_0 \rangle_{\mu c}} \equiv \frac{1}{\langle T_K(\mathbf{\Gamma}) \rangle_{\mu c}} + O(1/N), \qquad (5.6.3)$$

where the ensemble average is microcanonical and taken with respect to (5.3.1) and $T_K(\mathbf{\Gamma})$ is the instantaneous kinetic temperature. (Note: there are obviously infinitely many other phase space projections that one could use to move between two infinitely close energy hypersurfaces. These lead to infinitely many different phase functions whose microcanonical and canonical averages equate, in the thermodynamic limit, to the equilibrium thermodynamic temperature – see for example (5.6.19).)

If we now use the SLLOD equations, (note the SLLOD equations of motion give an exact description of arbitrary homogeneous flows – see [ii]) to accomplish an infinitesimal volume change at constant energy using an ergostat to fix the energy we see that from the ergostatted equations of motion

$$dH_0 = dU = 0 = -pdV - 2K_0 \alpha dt \qquad (5.6.4)$$

where from the SLLOD equations, p is the microscopic expression for the pressure in a bulk system which is spatially uniform over the range of intermolecular forces,

$$3pV = \sum_{i \in V} p_i^2 / m - \frac{1}{2} \sum_{i \in V, \forall j} \mathbf{r}_{ij} \bullet \mathbf{F}_{ij} \qquad (5.6.5)$$

where $\mathbf{r}_{ij} \equiv \mathbf{r}_j - \mathbf{r}_i$, $\mathbf{F}_{ij}$ force on particle *i* due to particle *j,* and α alpha is the ergostat multiplier. As Irving and Kirkwood showed [iii] this microscopic expression for *p,* is easily identified with the microscopic mechanical force "across" a surface and is therefore on average, equal to the thermodynamic pressure.

We also know from the phase continuity equation $df / dt = 3N\alpha f$, that the change in phase space volume, $\mathrm{d}V_{\Gamma}$, caused by this constant energy volume change is,

$$dV_\Gamma = -3N\langle\alpha\rangle_{\mu c} V_\Gamma dt \tag{5.6.6}$$

From our proposed microscopic equation for the entropy we see that

$$\left.\frac{\partial \tilde{S}}{\partial V}\right|_U = \left\langle \frac{3Nk_B p}{2K_0} \right\rangle_{\mu c} = \left\langle \frac{p}{T_K} \right\rangle_{\mu c} = \frac{\langle p\rangle_{\mu c}}{\tilde{T}} + O(1/N) \tag{5.6.7}$$

Comparing equations (5.6.3,7) and (5.6.1) and noting that the classical entropy is only defined up to an arbitrary constant, we conclude that S and $\tilde{S}$ satisfy the same partial differential equation:

$$\left.\frac{\partial X}{\partial V}\right|_U \Big/ \left.\frac{\partial X}{\partial U}\right|_V = p\ . \tag{5.6.8}$$

This means that up to an arbitrary additive constant the entropy and $\tilde{S}(U,V)$ are the same function of U,V:

$$S(V,U) = \tilde{S}(V,U) + O(1/N) + const\ . \tag{5.6.9}$$

Note that $T,\tilde{T}$ which are yet unresolved, both individually cancel from the two versions of (5.6.8) (when $X = S,\tilde{S}$). Substituting the thermodynamic entropy into (5.6.7) and comparing with the second equation in (5.6.1) then shows that

$$T(V,U) = \tilde{T}(V,U) + O(1/N)\,. \qquad (5.6.10))$$

The $O(1/N)$ corrections disappear in the thermodynamic limit where classical thermodynamics is valid. Having identified microscopic expressions for the entropy and the temperature in equilibrium microcanonical systems, we can now apply the usual text book arguments to calculate expressions for the Helmholtz free energy in canonical systems.

We can also give a microscopic expression for the Helmholtz Free energy of equilibrium canonical systems directly.

We *postulate* that the Helmholtz Free energy $A(T,N,V)$ is the same function of the thermodynamic temperature as Q is of the Nose-Hoover target temperature T_{th} in (5.4.2),

$$\begin{aligned} A(T=T_{th},N,V) &= Q(T_{th},N,V)+(1/N) \\ &\equiv -k_B T_{th}\ln\left[\int d\boldsymbol{\Gamma}\,\delta(\mathbf{p}_{th})\exp[-\beta_{th}H_E(\boldsymbol{\Gamma})]\right]. \end{aligned} \qquad (5.6.11)$$

That is when $T_{th}=T$ the Helmholtz free energy $A(T)$, at the thermodynamic temperature *T*, is equal to the value of the statistical mechanical expression $Q(T_{th})$ that is defined in (5.6.11). From thermodynamics we note that the Helmholtz free energy satisfies the differential equation

$$U = A - T\frac{\partial A}{\partial T}, \qquad (5.6.12)$$

where *U* is the internal energy. Whereas if we differentiate *Q* with respect to T_{th} we see that

$$\langle H_0 \rangle = Q - T_{th}\frac{\partial Q}{\partial T_{th}}. \qquad (5.6.13)$$

Since $U=\langle H_0\rangle$ and noting that when $T=T_{th}=0$, that $A(0)=U(0)=Q(0)$, we observe, treating T,T_{th} as integration parameters *x*, that *A* and *Q* satisfy the same differential equation $U(x)=Y(x)-x\,{\partial Y(x)}/{\partial x}$ with the same initial $x=0$ condition and therefore $A(T)=Q(T_{th})$ and our hypothesis (5.6.11) is proved [29, 30]. The pressure can be verified using the SLLOD equations but since pressure is a mechanical property obtaining a microscopic expression for pressure presents no difficulties.

Lastly we derive a microscopic expression for the entropy of an equilibrium canonical system. Firstly we note from classical thermodynamics that,

$$S = \frac{U - A}{T} . \tag{5.6.14}$$

Substituting the microscopic expressions for both the Helmholtz Free energy and the internal energy into (5.6.14) gives

$$\begin{aligned} S &= \frac{\int d\mathbf{\Gamma}\; e^{-\beta H_0} H_0}{T\int d\mathbf{\Gamma}\; e^{-\beta H_0}} + k_B \ln[\int d\mathbf{\Gamma}\; e^{-\beta H_0}] \\ &= \frac{-k_B \int d\mathbf{\Gamma}\; e^{-\beta H_0}[-\beta H_0 - \ln\int d\mathbf{\Gamma}\; e^{-\beta H_0}]}{\int d\mathbf{\Gamma}\; e^{-\beta H_0}} \qquad (5.6.15) \\ &= -k_B \int d\mathbf{\Gamma}\; f_c(\mathbf{\Gamma}) \ln[f_c(\mathbf{\Gamma})] \end{aligned}$$

where $f_c(\mathbf{\Gamma})$ is given by the equilibrium canonical distribution (5.4.3).

Definition

The *Gibbs entropy* (see (2.6.1)) of a phase space distribution $f(\mathbf{\Gamma})$ is defined as:

$$S_G \equiv -k_B \int d\mathbf{\Gamma}\; f(\mathbf{\Gamma}) \ln[f(\mathbf{\Gamma})] \tag{5.6.16}$$

The Gibbs entropy of the equilibrium canonical distribution (5.4.3), is the thermodynamic entropy of the equilibrium system in contact with a heat bath at the specified thermodynamic temperature.

The Gibbs entropy of a microcanonical distribution of states is the thermodynamic entropy of the isolated autonomous Hamiltonian system with internal energy *U.* We can calculate the Gibbs entropy of a microcanonical distribution of states:

$$\begin{aligned} S_{G,\mu c} &= -k_B \int_D d\mathbf{\Gamma} \, \frac{1}{\int_D d\mathbf{\Gamma}} \{\ln[1] - \ln[\int_D d\mathbf{\Gamma}]\} \\ &= +k_B \ln[\int_D d\mathbf{\Gamma}] \end{aligned} \tag{5.6.17}$$

where *D* denotes the energy-momentum hypersurface in phase space given in (5.3.1). Equation (5.6.17) is of course consistent with (5.6.2).

We will now evaluate the Gibbs entropy directly at some time *t:*

$$\begin{aligned} S_G(t) &= -k_B \int d\mathbf{\Gamma} \; f(\mathbf{\Gamma};t) \ln(f(\mathbf{\Gamma};t) \\ &= -k_B \int d\mathbf{\Gamma} \; f(S^{-t}\mathbf{\Gamma};0) \ln(f(S^{-t}\mathbf{\Gamma};0) \\ &= -k_B \int d(S^{-t}\mathbf{\Gamma}) \left| \frac{\partial \mathbf{\Gamma}}{\partial S^{-t}\mathbf{\Gamma}} \right| f(S^{-t}\mathbf{\Gamma};0) \ln(f(S^{-t}\mathbf{\Gamma};0) \\ &= -k_B \int d\mathbf{\Gamma} \; f(\mathbf{\Gamma};0) \ln(f(\mathbf{\Gamma};0) = S_G(0) \end{aligned} \tag{5.6.18}$$

In deriving (5.6.17) we have used the facts that: the energy is a constant of the motion for autonomous Hamiltonian systems; the phase space expansion factor is identically zero for Hamiltonian systems implying (as Liouville discovered) that the what we call the Jacobian for time translation of phase space vectors is unity.

Lastly we should make a comment seldom made in textbooks. We have derived a number of "standard" microscopic expressions for thermodynamic quantities. However each such expression is not unique. For example there are infinitely many different expressions for the equilibrium pressure or temperature. At equilibrium you can even calculate the equilibrium thermodynamic temperature using an expression that is purely configurational. If $\nabla_{\mathbf{q}} \equiv (\partial/\partial\mathbf{q}_1,....,\partial/\partial\mathbf{q}_N)$ then

$$\frac{1}{k_B T} = \left\langle \frac{\nabla_{\mathbf{q}}^2 \Phi(\mathbf{q})}{\left|\nabla_{\mathbf{q}} \Phi(\mathbf{q})\right|} \right\rangle + O\left(\frac{1}{N}\right) \qquad (5.6.19)$$

where $\Phi(\mathbf{q})$ is the interparticle potential energy of the system and $\mathbf{q}$ denotes all the Cartesian coordinates of all the particles in the system. Typically what happens is that away from equilibrium these different expressions for equilibrium thermodynamic quantities each take on very different values, again pointing out how special the equilibrium state is.

REFERENCES

[1] D.A. McQuarrie, Statistical Mechanics, (Harper and Row, New York, 1976).

[2] R. C. Tolman, The Principles of Statistical Mechanics, (Dover Publications, New York, 1979).

[3] D. Chandler, Introduction to Modern Statistical Mechanics, (Oxford University Press, Oxford, 1987).

[4] K. Huang, Statistical Mechanics, (John Wiley & Sons, New York, 1963).

[5] T.L. Hill, Statistical Mechanics, Principles and Selected Applications, (Dover, New York, 1987).

[6] J.R. Dorfman, An Introduction to Chaos in Nonequilibrium Statistical Mechanics, (Cambridge University Press, Cambridge, 1999).
V.I. Arnold and A. Avez, Ergodic Problems of Classical Mechanics, (W.A. Benjamin, New York, 1968) see also
A.I. Khinchin, Mathematical Foundations of Statistical Mechanics, Dover Publications, New York, 1949.

[7] P. and T. Ehrenfest, The Conceptual Foundations of the Statistical Approach in Mechanics, (Dover, New York, 1990).

[8] D. J. Evans and D. J. Searles, Phys. Rev. E 50, 1645 (1994), ibid Phys. Rev. E 52, 5839 (1995); D. J. Evans and D. J. Searles, Adv. in Phys. 51, 1529 (2002); D. J. Searles and D. J. Evans, J. Chem. Phys. 113, 3503 (2000) and L. Rondoni and C. Mejia-Monasterio, Nonlinearity, 20, R1(2007).

[9] E. M. Sevick, R. Prabhakar, S. R. Williams and D. J. Searles, Ann. Rev. Phys. Chem., 59, 603 (2008).

[10] G. E. Crooks, Phys. Rev. E 61, 2361 (2000).

[11] C. Jarzynski, Phys. Rev. Lett., 78, 2690 (1997); C. Jarzynski, Phys. Rev. E, 56, 5018 (1997).

[12] C. Jarzynski, J. Stat. Phys., 98, 77 (2000).

[13] D. J. Evans, D. J. Searles and S.R. Williams, J. Chem. Phys., 128, 014504 (2008), ibid 128, 249901 (2008).

[14] D. J. Searles and D. J. Evans, Aust. J. Chem. 57, 1119 (2004).

[15] D. J. Evans and G. P. Morriss, Statistical Mechanics of Nonequilibrium Liquids, second edition, (Cambridge University Press, Cambridge, 2008).

[16] G. Ayton, D. J. Evans and D. J. Searles, J. Chem. Phys., 115, 2035 (2001); S. R. Williams, D. J. Searles and D. J. Evans, Phys. Rev. E, 70, 066113 (2004) and G. Gallavotti, arXiv:cond-mat/0601049v2 (2006).

[17] D. J. Evans, S. R. Williams and D. J. Searles, Thermodynamics of Small Systems, in Nonlinear Dynamics of Nanosystems, Eds. G. Radons, B. Rumpf and H. G. Schuster (Wiley-VCH, Weinheim, in press).

[18] S. R. Williams and D. J. Evans, Phys. Rev. E, 78, 021119 (2008).

[19] S. R. Williams and D. J. Evans, J. Chem. Phys., 127, 184101 (2007).

REFERENCES

[1] D. A. McQuarrie, *Statistical Mechanics*, (Harper and Row, New York, 1976).

[2] R. C. Tolman, *The Principles of Statistical Mechanics*, (Dover Publications, New York, 1979).

[3] D. Chandler, *Introduction to Modern Statistical Mechanics*, (Oxford University Press, Oxford, 1987).

[4] K. Huang, *Statistical Mechanics*, (John Wiley & Sons, New York, 1963).

[5] T. L. Hill, *Statistical Mechanics, Principles and Selected Applications*, (Dover, New York, 1987).

[6] P. and T. Ehrenfest, *The Conceptual Foundations of the Statistical Approach in Mechanics*, (Dover, New York, 1990).

[7] J. L. Lebowitz, *Phys. Today*, **46**, 32 (1993).

[8] U. M. B. Marconi, A Puglisi, L. Rondoni and A. Vulpiani, Phys. Rep., **461**, 111 (2008).

[9] J. R. Dorfman, *An Introduction to Chaos in Nonequilibrium Statistical Mechanics*, (Cambridge University Press, Cambridge, 1999).

[10] V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, (W.A. Benjamin, New York, 1968) see also

A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics*, Dover Publications, New York, 1949 and,

D. Szasz, Physica A, **194**, 86 (1993).

[11] D. J. Evans and D. J. Searles, Phys. Rev. E **50**, 1645 (1994); D. J. Evans and D. J. Searles, Adv. in Phys. **51**, 1529 (2002); D. J. Searles and D. J. Evans, J. Chem. Phys.

113, 3503 (2000), L. Rondoni and C. Mejia-Monasterio, Nonlinearity, **20**, R1(2007) and U. M. B. Marconi *et. al.,* Phys, Repts. **461**, 111(2008).

[12] D. J. Evans and D. J. Searles, Phys. Rev. E, **52**, 5839 (1995).

[13] E. M. Sevick, R. Prabhakar, S. R. Williams and D. J. Searles, Ann. Rev. Phys. Chem., **59**, 603 (2008).

[14] G. E. Crooks, Phys. Rev. E **61**, 2361 (2000).

[15] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett., 74, 2694 (1995); G. Gallavotti and E. G. D. Cohen, J. Stat. Phys., **80**, 931 (1995); C. Jarzynski, Phys. Rev. Lett., **78**, 2690 (1997); C. Jarzynski, Phys. Rev. E, **56**, 5018 (1997).

[16] G. E. Crooks, J. Stat. Phys., **90**, 1481 (1998).

[17] D. J. Evans, D. J. Searles and S.R. Williams, J. Chem. Phys., **128**, 014504 (2008), ibid **128,** 249901 (2008).

[18] D. J. Searles and D. J. Evans, Aust. J. Chem. **57**, 1119 (2004).

[19] D. J. Evans and G. P. Morriss, Statistical Mechanics of Nonequilibrium Liquids, second edition, (Cambridge University Press, Cambridge, 2008).

[20] G. Ayton, D. J. Evans and D. J. Searles, J. Chem. Phys., **115**, 2035 (2001); S. R. Williams, D. J. Searles and D. J. Evans, Phys. Rev. E, **70**, 066113 (2004) and G. Gallavotti, arXiv:cond-mat/0601049v2 (2006).

[21] D. J. Evans, S. R. Williams and D. J. Searles, *Thermodynamics of Small Systems*, in Nonlinear Dynamics of Nanosystems, Eds. G. Radons, B. Rumpf and H. G. Schuster (Wiley-VCH, Weinheim, in press).

[22] T. Yamada and K. Kawasaki, Prog. Theo. Phys., **38**, 1031 (1967), *ibid* **53**, 111 (1975), *ibid* **53**, 437(1975), see also,

K. Kawasaki and J.D. Gunton, *Phys. Rev. A*, **8**, 2048 (1973).

[23] D. J. Evans and G. P. Morriss, *Phy.Rev.*, **32**, 2923 (1985).

[24] Our definition of ergodic is that for all initial phase points (except a set of zero measure) a phase space trajectory emanating from that point explores (almost) all of phase space.

[25] S. R. Williams and D. J. Evans, Phys. Rev. E, **78**, 021119 (2008).

[26] If there are other constants of the motion the system cannot be ergodic. The proof is simple: at $t = 0$ only consider phases having a common fixed value for this constant of the motion; as time evolves the ensemble cannot explore the rest of phase space.

[27] This is to be contrasted with a nonequilibrium steady state where on average heat flows from the system to a heat reservoir. This is clearly a time asymmetric state.

[28] S. R. Williams and D. J. Evans, J. Chem. Phys., **127**, 184101 (2007).

[29] This proof is really proof by induction except that the iterations are infinitesimal.

[30] Although Eq. **Error! Reference source not found.** refers to the Helmholtz free of the extended Hamiltonian, it is trivial to reduce this to the system Hamiltonian.

---

[i] B. D. Butler, G. Ayton, O. G. Jepps and D. J. Evans, J. Chem. Phys. **109**, 6519 *1998*.

[ii] P. J. Daivis and B. D. Todd, J. Chem. Phys. **124**, 194103 *2006*.

[iii] J. H. Irving and J. G. Kirkwook, J. Chem. Phys. **18**, 817 *1950.*