

Lecture 8. Disequilibrium I: Stochastic Processes

1. INTRODUCTION

The problem of determining the appropriate probability distributions of statistical dynamical systems not in equilibrium lacks a general solution unlike the equilibrium case in which a Gibbs measure (which is a maximum entropy distribution) is appropriate. A number of different approaches to this rather difficult issue have been attempted over the past century but it is fair to say that none is entirely satisfactory. In the next several lectures we will look at various approaches that have had some success.

A very common approach is to divide dynamical variables into fast and slow. This can be assessed in various ways. For example, the ensemble mean of the variables could be calculated and the timescale of variations of this assessed. Alternatively temporal decorrelation of the random variables could be calculated. One then assumes that the fast variables can be replaced with “noise” which is a random variable with very rapid temporal decorrelation. White noise has instantaneous decorrelation and is the most commonly considered case. The objective then is provide a model for the probability distribution for the slow variables of the system. A common assumption in this modelling process is that the probability distributions satisfy a so-called Markov assumption in that the distribution at any one time determines future distributions. The associated random processes are commonly referred to as stochastic processes and have been extensively studied mathematically. Several excellent graduate courses here at the Courant cover this theory in detail and are thoroughly recommended. Our focus here will be on studying the time evolution of entropic functionals on such distributions in an effort to see whether there exists a second law of thermodynamics for such systems. We shall restrict our attention to the case that the slow random variables have outcomes sets which are continuous. We shall also focus to begin with on the case where any realization of such variables is continuous in time i.e. does not exhibit “jumps” in time.

2. CONTINUOUS MARKOV STOCHASTIC PROCESSES

2.1. Fokker Planck Equation. Stochastic processes in which the outcomes belong to the reals are immensely important in physics and other application fields. The most commonly studied such processes satisfy the Markov condition discussed above. A particularly important subclass of such processes are associated with the so-called Fokker Planck Equation (FPE). They are often referred to as solutions of stochastic differential equations (see [1], for a good introduction). Crudely speaking, realisations of these processes are continuous (but not differentiable) in time. There exist other Markov processes in which a realisation may jump discontinuously at a particular time. Examples of those are given by the so-called Master equations which we consider briefly below and in the next Lecture. The FPE is an evolution equation for probability density functions (pdfs) defined on an N dimensional space of the form

$$(2.1) \quad \partial_t p = - \sum_{i=1}^N \partial_i [A_i(\mathbf{x}, t)p] + \frac{1}{2} \sum_{i,j=1}^N \partial_i \partial_j \{C_{ij}p\}$$

where it is usually assumed that the matrix C_{ij} is non-negative definite and symmetric. The reason this equation is important physically is because it has the intuitive interpretation that realisations represent possible evolutions of a dynamical system which is stochastically forced by Gaussian noise white in time (i.e. completely uncorrelated in the time dimension). The dynamical system has the form

$$(2.2) \quad \frac{\partial x_i}{\partial t} = A_i(\mathbf{x}, t)$$

and the Gaussian forcing has a covariance matrix given by C_{ij} (hence the restriction on it mentioned previously). More rigorously the above equation is associated with the Ito stochastic differential equation

$$dx_i = A_i(\mathbf{x}, t)dt + B_{ij}(\mathbf{x}, t)dW_j$$

where W_i are so-called random walk Wiener processes; $C_{ij} = B_{ki}B_{kj}$ and where we are assuming that repeated indices are summed.

Three interesting results are available regarding the evolution of the ordinary and relative entropy within such a system. The first two apply to systems without stochastic forcing

Theorem 1. *Suppose we have a realization of a stochastic process obeying equation (2.1) with $\mathbf{B} = 0$*

then the ordinary (differential) entropy satisfies the evolution equation

$$(2.3) \quad h_t = \int p \nabla \cdot \mathbf{A} dx = \langle \nabla \cdot \mathbf{A} \rangle_p$$

Proof. Let the realization of the process have pdf f then it follows that

$$\begin{aligned} -(f \ln(f))_t &= -f_t (\ln f + 1) \\ &= \nabla \cdot (\mathbf{A} f) (\ln f + 1) \\ &= f \nabla \cdot \mathbf{A} (\ln f + 1) + \mathbf{A} \cdot (\nabla f) (\ln f + 1) \\ &= f \nabla \cdot \mathbf{A} (\ln f + 1) + \nabla \cdot (\mathbf{A} f \ln f) - \nabla \cdot \mathbf{A} f \ln f \\ &= f \nabla \cdot \mathbf{A} + \nabla \cdot (\mathbf{A} f \ln f) \end{aligned}$$

Consider $r = -f \ln f$ as an entropy density whose integral over all space is the (Shannon) entropy. Corresponding to this density is a current $\mathbf{j} \equiv \mathbf{A} r$ whose divergence may be seen on the RHS of the above equation. For suitable restrictions on \mathbf{A} and f required for the definition of Shannon entropy and for the FPE to be well behaved, the surface integral of \mathbf{j} is zero for an infinite surface. The divergence theorem then ensures that the integral of the second term above vanishes and we are then left with equation (2.3). \square

Notice the importance of $\nabla \cdot \mathbf{A}$ to the entropy evolution. Since \mathbf{A} is the “velocity vector” of the dynamical system space (see equation(2.2)) then $\nabla \cdot \mathbf{A}$ measures the time rate at which an infinitesimal volume element expands or contracts in the dynamical system. When it vanishes the system is said to satisfy a Liouville condition. As we saw in an earlier lecture systems of molecules obey Hamiltonian dynamics in which this divergence vanishes so entropy is conserved for such systems. This issue is important for the study of dilute gases as we shall see in the next lecture.

The relative entropy on the other hand is conserved in all systems with $\mathbf{B} = 0$:

Theorem 2. *Suppose we have two realizations of a stochastic process obeying equation (2.1) which have the additional condition that $\mathbf{B}(\mathbf{x}, t) = 0$ then the relative entropy of the two realizations (if defined) is time invariant.*

Proof. Let the two realizations of the process have pdfs f and g then it follows that

$$\begin{aligned} (f \ln(f/g))_t &= f_t \ln(f/g) + f_t - g_t(f/g) = f_t(\ln(f/g) + 1) - g_t(f/g) \\ &= -\nabla \cdot (\mathbf{A}f)(\ln(f/g) + 1) + \nabla \cdot (\mathbf{A}g)(f/g) \\ &= [-f\nabla \cdot \mathbf{A} - \mathbf{A} \cdot (\nabla f)] [\ln(f/g) + 1] + [g\nabla \cdot \mathbf{A} + \mathbf{A} \cdot (\nabla g)] (f/g) \\ &= -\nabla \cdot \mathbf{A}f \ln(f/g) - \mathbf{A} \cdot (\nabla f [\ln(f/g) + 1] - \nabla g(f/g)) \\ &= -\nabla \cdot \mathbf{A}f \ln(f/g) - \mathbf{A} \cdot \nabla (f \ln(f/g)) \\ &= -\nabla \cdot (\mathbf{A}f \ln(f/g)) \end{aligned}$$

In this case the entire right hand side of the evolution equation is in the form of a divergence of a “relative entropy” density current and as argued in the previous theorem this implies that the global integral of the left hand side vanishes by the divergence theorem. \square

In many classical systems with $\mathbf{B} = 0$ if one calculates the relative entropy with respect to a particular finite partitioning of state space rather in the limit of infinitesimal partitioning then the conservation property no longer holds and in nearly all interesting cases it declines with time instead and the system equilibrates. This reflects the fact that as time increases the difference in the distributions tends to occur on the unresolved scales which are not measured by the second relative entropy calculation. This *coarse graining* effect is often related to the next result.

In the stochastically forced case we have:

Theorem 3. *Suppose we have two distinct¹ stochastic processes obeying (2.1) with $C = \mathbf{B}(\mathbf{x}, t)\mathbf{B}^t(\mathbf{x}, t)$ positive definite almost everywhere then the relative entropy strictly declines.*

Proof. As in the previous theorem we consider the relative entropy “density” function $r = f \ln f/g$. Clearly the proof of this shows we need only consider the time rate of change in this function due to C since that due to \mathbf{A} leads to no change in time of the global integral of r . The change in r due to C is easily calculated using equation (2.1):

$$(2.4) \quad (r_c)_t = (\ln(f/g) + 1) \partial_i \partial_j (C_{ij} f) - \frac{f}{g} \partial_i \partial_j (C_{ij} g)$$

where we are using the summation convention for repeated latin indices. Now it is easy to see that

$$(2.5) \quad \partial_i \partial_j (C_{ij} u w) = w \partial_i \partial_j (C_{ij} u) + 2(\partial_j (C_{ij} u))(\partial_i w) + C_{ij} u \partial_i \partial_j (w)$$

where we are using the symmetry of C . Writing $g = f(g/f)$ and applying the last relation we derive that the second term of equation (2.4) is

$$(2) = -\frac{f}{g} \left[\frac{g}{f} \partial_i \partial_j (C_{ij} f) + 2\partial_i (C_{ij} f) \partial_i \left(\frac{g}{f} \right) + C_{ij} f \partial_i \partial_j \left(\frac{g}{f} \right) \right]$$

combining this with the first term we get a cancellation of the first term of (2) with part of the first term of equation (2.4) and so

$$(r_c)_t = \ln(f/g) \partial_i \partial_j (C_{ij} f) - 2 \left(\frac{f}{g} \right) \partial_j (C_{ij} f) \partial_i \left(\frac{g}{f} \right) - \left(\frac{f}{g} \right) C_{ij} f \partial_i \partial_j \left(\frac{g}{f} \right)$$

¹In other words differing on a set of measure greater than zero.

Now to this equation we add and subtract the terms

$$2\partial_i(\ln \frac{f}{g})\partial_j(C_{ij}f) + C_{ij}f\partial_i\partial_j\left(\ln \frac{f}{g}\right)$$

and use equation (2.5) to deduce that

$$(r_c)_t = \partial_i\partial_j(C_{ij}r) - \left(C_{ij}f\left[\frac{f}{g}\partial_i\partial_j\left(\frac{g}{f}\right) + \partial_i\partial_j\left(\ln \frac{f}{g}\right)\right]\right)$$

where we are using the definition of r as well as cancelling two terms involving $\partial_j(C_{ij}f)$. It is straightforward (albeit tedious) to simplify the expression in the square brackets and obtain finally

$$(2.6) \quad (r_c)_t = \partial_i\partial_j(C_{ij}r) - fC_{ij}\partial_i(\ln \frac{f}{g})\partial_j(\ln \frac{f}{g})$$

The first term on the right is of the form of a divergence and so using similar arguments as previously does not contribute to the evolution of the global integral of r_C . Actually the positive definite nature of C shows that it is purely diffusive of the “relative entropy” density r . The second term is negative almost everywhere due to the fact that C is positive definite almost everywhere and that f and g differ on a set of measure greater than zero. Thus in that situation if we take the global integral of r_C we conclude that the relative entropy declines strictly with time. \square

This third theorem shows the central role of stochastic forcing in causing relative entropy to decline monotonically. This behaviour is sometimes called a generalized second law of thermodynamics and $-D(f, f_{eq})$ with f_{eq} the equilibrium density for the system has been proposed as a non-equilibrium entropy function (see the book by Mackey mentioned in the first lecture) where it referred to as the conditional entropy. Note also that this result holds for any two densities satisfying the FPE not just f and f_{eq} . This result was first proven in 1957 by Joel Lebowitz from Rutgers [2]. It is a crucial result in the theory of the Fokker Planck equation in that it allows a deduction of the convergence of a density to an equilibrium density as $t \rightarrow \infty$. For more on this see the book [3].

It is possible to extend the Fokker Planck equation to include discontinuous jump processes and then this equation becomes the more general Chapman-Kolmogorov equation. The additional terms are often referred to (on their own) as the Master equation. It is then possible by similar arguments to those given above to conclude that the jump processes result in an additional strict monotonic decline in relative entropy. The interested reader is referred to Chapter 3 of [1] for a sketch proof and more information and references.

There is also a very interesting connection between these results and the so-called kinetic theory of Boltzmann which applies to dilute gases. We take this up in the next Lecture.

REFERENCES

- [1] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, volume 13 of *Springer Series in Synergetics*. Springer, Berlin Heidelberg, 2004.
- [2] J.L. Lebowitz and P.G. Bergmann. Irreversible gibbsian ensembles. *Annals of Physics*, 1(1):1–23, 1957.
- [3] H. Risken. *The Fokker-Planck Equation*. Springer Verlag, Berlin, 2nd edition, 1989.