Applications of the Maximum Entropy Principle to Time Dependent Processes

Johann-Heinrich Christiaan Schönfeldt

Faculty of Natural & Agricultural Science

University of Pretoria

Pretoria

Submitted in partial fulfilment of the requirements for the degree

Magister Scientae

March 2007

I declare that the dissertation, which I hereby submit for the degree Magister Scientae at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Acknowledgements

I would like to thank the following people:

- my supervisor Prof. Plastino for being understanding and being a great source of advice and inspiration.
- my parents for the immense contribution that they have made during my studies, my education and my upbringing.
- the love of my life, Elna, for being that and my best friend, giving me support all along the way.
- my fellow students, past and present for making the study environment fun, and for being there for advice.
- Dr. Walter Meyer for all the conference trips and helping with numerous small things and some critical disasters, from computer woes to general advice.
- The coffee club team for providing an (almost) uninterrupted source of caffeine.

Applications of the Maximum Entropy Principle to Time Dependent Processes

Johann-Heinrich Christiaan Schönfeldt
Supervised by Prof. A.R. Plastino
Department of Physics
Submitted for the degree: M.Sc (Physics)

Summary

The maximum entropy principle, pioneered by Jaynes, provides a method for finding the least biased probability distribution for the description of a system or process, given as prior information the expectation values of a set (in general, a small number) of relevant quantities associated with the system. The maximum entropy method was originally advanced by Jaynes as the basis of an information theory inspired foundation for equilibrium statistical mechanics. It was soon realised that the method is very useful to tackle several problems in physics and other fields. In particular it constitutes a powerful tool for obtaining approximate and sometimes exact solutions to several important partial differential equations of theoretical physics. In Chapter 1 a brief review of Shannon's information measure and Jaynes' maximum entropy formalism is provided. As an illustration of the maximum entropy principle a brief explanation of how it can be used to derive the standard grand canonical formalism in statistical mechanics is given.

The work leading up to this thesis has resulted in the following publications in peer-review research journals:

- J.-H. Schönfeldt and A.R. Plastino, Maximum entropy approach to the collisional Vlasov equation: Exact solutions, Physica A, 369 (2006) 408-416,

- J.-H. Schönfeldt, N. Jimenez, A.R. Plastino, A. Plastino and M. Casas, Maximum entropy principle and classical evolution equations with source terms, Physica A, 374 (2007) 573-584,
- J.-H. Schönfeldt, G.B. Roston, A.R. Plastino and A. Plastino, Maximum entropy principle, evolution equations, and physics education, Rev. Mex. Fis. E, 52 (2)(2006) 151-159.

Chapter 2 is based on Schönfeldt and Plastino (2006). Two different ways for obtaining exact maximum entropy solutions for a reduced collisional Vlasov equation endowed with a Fokker-Planck like collision term are investigated.

Chapter 3 is based on Schönfeldt *et al.* (2007). Most applications of the maximum entropy principle to time dependent scenarios involved evolution equations exhibiting the form of a continuity equations and, consequently, preserving normalization in time. In Chapter 3 the maximum entropy principle is applied to evolution equations with source terms and, consequently, not preserving normalization. We explore in detail the structure and main properties of the dynamical equations connecting the time dependent relevant mean values , the associated Lagrange multipliers, the partition function, and the entropy of the maximum entropy scheme. In particular, we compare the H-theorems verified by the maximum entropy approximate solutions with the H-theorems verified by the exact solutions.

Chapter 4 is based on Schönfeldt *et al.* (2006). In chapter 4 it is discussed how the maximum entropy principle can be incorporated into the teaching of aspects of theoretical physics related to, but not restricted to, statistical mechanics. We focus our attention on the study of maximum entropy solutions to evolution equations that exhibit the form of continuity equations (eg. Liouville equation, the diffusion equation the Fokker-Planck equation, etc.).

Contents

Chapter 1

Introduction

The maximum entropy principle (Beck and Schlögl (1993); Jaynes (1963); Katz (1967)) was pioneered in Jaynes (1957) where Jaynes used the maximum entropy principle to show how the computational rules for statistical mechanics can be inferred by considering statistical mechanics as a form of statistical inference. He showed how statistical mechanics is connected to information theory by adopting a viewpoint where thermodynamic entropy and information entropy appear as the same concept. The principle describes a method for finding the least biased probability distribution for the description of a system or process, given as prior information the expectation values of a set (in general a small number) of relevant quantities associated with the system. It constitutes a powerful tool for obtaining approximate (sometimes exact) solutions to several important partial differential equations of physics Borland *et al.* (1999); da Silva *et al.* (2004); Frank (2005); Plastino *et al.* (1997a,b,c); Tsallis and Bukman (1996) and other scientific disciplines Borland (2002); Gell-Mann and Tsallis (2004).

Consider the situation where a physical quantity x can assume the discrete values x_i $(i = 1, 2, \ldots, n)$. The probabilities f_i corresponding to the x_i are not known and form a probability distribution f. The only information known about the system is the expectation values for the $j = 1, 2, \ldots, m$ dynamical quantities $g_j(x)$ (for example through experiment),

$$\langle g_j(x)\rangle = \sum_i f_i g_j(x_i), \tag{1.1}$$

and the normalization condition for the probabilities

$$\langle 1 \rangle = \sum_i f_i = 1. \tag{1.2}$$

A number of different probability distributions f may yield expectation values that are the same as those obtained via experiment, and thus these different probability distributions all describe the observed physical phenomena. The question then must be which f is the best. The maximum entropy argument is that the probability distribution that contains the least information must be the better one. Any extra information that a probability distribution contains has no bearing on the expectation values obtained from the experiment, and thus has no physical justification. The extra information represents knowledge about certain expectation values that were never observed, and thus should not be described by the probability distribution. A probability distribution with no extra information will describe the physical results obtained without making any assumptions about measurements that were never made. Thus the f that has the maximum missing information or uncertainty (whilst still describing the observed expectation values) will be the least biased distribution. The problem is then how to decide which f has the most missing information.

In Chapter (3) it is discussed how the maximum entropy principle can also be applied to positive densities, in other words a distribution that is not normalised to 1. In fact the normalisation of the distribution could even vary with time.

1.1 Missing Information (Uncertainty) and Entropy

First a suitable measure for the missing information (I) of a probability distribution is required. Once we have such a measure we can try to maximise it in order to obtain the probability distribution with the least extra information. A measure of missing information is the same thing as a measure of the uncertainty. Such a measure must conform to certain requirements (Kapur (1989); Khinchin (1957); Nielsen and Chuang (2000)):

i. It should be a function of the probabilities, ie.

$$I(f) = I(f_1, f_2, f_3, \ldots, f_n). \tag{1.3}$$

ii. It should be a continuous function so that small changes in the probabilities only effect a small change in I.

iii. It should remain the same if the probabilities are shuffled. Thus I should be a symmetric function of the probabilities.

iv. I must not change if an impossible event is added to the probability distribution,

$$I(f_1, f_2, f_3, \ldots, f_n, 0) = I(f_1, f_2, f_3, \ldots, f_n). \tag{1.4}$$

v. As one of the f_i approaches one and the rest zero, uncertainty is clearly decreasing and thus the measure of missing information must approach zero.

vi. The measure must have its maximum, for a given n, when the uncertainty is maximum, ie., when the probabilities f_i are all equal:

$$f_i = 1/n \;\; (i = 1, 2, \ldots, n). \tag{1.5}$$

vii. The maximum for $I(f_1, f_2, f_3, \ldots, f_n)$ should increase as n increases.

viii. The measure for missing information should be additive, so that the sum of the missing information for two independent probability distributions f and p, corresponding to discrete physical quantities x_i $(i = 1, 2, \ldots, n)$ and y_h $(h = 1, 2, \ldots, l)$ respectively, is equal to the missing information of the joint scheme $f \cup p$. In other words,

$$I(f, p) = I(f) + I(p). \tag{1.6}$$

Thus the amount of information gained by performing an experiment on the joint scheme, and finding say $x_3 y_3$, should be the same as the information gained by doing two separate experiments with the results x_3 and y_3.

ix. A generalisation of (viii) is that the missing information for a joint probability scheme $f \cup p$, where f and p are not necessarily independent from each other, is equal to the missing information of one of the schemes and the mathematical expectation of the missing information of the other scheme, conditional on the realisation of the first scheme:

$$I(f,p) = I(f) + \sum_i f_i I_i(p), \tag{1.7}$$

where $I_i(p)$ is the conditional missing information of the probability distribution, p, corresponding to the value x_i in the probability distribution f. The term on the right of equation (1.7) is then the mathematical expectation of $I(p)$ with respect to the probability distribution f.

Shannon (1948), pioneered information theory and the measure of information. The measure for missing information I (as proposed by Shannon) for the discrete probability distribution f is given by

$$I = -k \sum_i^n f_i \ln f_i, \tag{1.8}$$

and for a continuous probability distribution, $f(x)$, by

$$I = -k \int f(x) \ln f(x) dx, \tag{1.9}$$

where the constant k is a positive number. These measures satisfy all the requirements (i-ix) as set out above. The constant k, will determine the unit of information used in a given problem, for convenience usually set to 1. Khinchin (1957), showed that any measure that satisfies requirements (ii,iv,vi and ix) must be of equation (1.8)'s form. Equations (1.8) and (1.9) are also referred to as the information entropy equations since, with $k = k_B$, the Boltzmann constant, they are identical to the thermodynamic entropy in statistical mechanics (Reif (1965)). There are however other formulations for the information measure that do not conform to all the requirements. Renyi's entropic measure for example satisfies requirements (i-viii) and a modified version of (ix), see (Kapur (1989); Rényi (1961)). Havrda and Charvat (1967) obtained non-additive measures of entropy by only satisfying requirements (i-vii),see also (Kapur (1989)). These other information measures sometimes prove useful to describe certain situations, and sometimes an alternative to Shannon information is the only measure that provides solutions for a given situation. In general, however, Shannon's measure remains the most useful. Tsallis entropy is a generalization of the Boltzmann-Gibbs entropy and is used in nonextensive statistical mechanics.

1.2 Formulation of the Maximum Entropy Principle

In order to maximize equation (1.8) we take the constraints, equations (1.2) and (1.1), into account by introducing the Lagrange multipliers $k(\lambda_0 - 1)$ and $k\lambda_j$ and then varying I':

$$\begin{aligned} I' &= I - k(\lambda_0 - 1)\langle 1 \rangle - k\sum_j^m \lambda_j \langle g_j(x) \rangle \\ &= -k\sum_i^n f_i \ln f_i - k(\lambda_0 - 1)\sum_i^n f_i - k\sum_j^m \lambda_j \sum_i^n f_i g_j(x_i) \\ &= -k\sum_i^n f_i \left(\ln f_i + \lambda_0 - 1 + \sum_j^m \lambda_j g_j(x_i) \right), \end{aligned} \tag{1.10}$$

the variation is without restriction and with respect to f_i

$$\begin{aligned} \delta I' &= -k\sum_i^n \delta f_i \left(\ln f_i + \lambda_0 - 1 + \sum_j^m \lambda_j g_j(x_i) \right) \\ &\quad -k\sum_i^n f_i\, \delta\!\left(\ln f_i + \lambda_0 - 1 + \sum_j^m \lambda_j g_j(x_i) \right) \\ &= -k\sum_i^n \delta f_i \left(\ln f_i + \lambda_0 + \sum_j^m \lambda_j g_j(x_i) \right) + k\sum_i^n \delta f_i \\ &\quad -k\sum_i^n f_i \;\; (\delta f_i / f_i) \\ &= -k\sum_i^n \delta f_i \left(\ln f_i + \lambda_0 + \sum_j^m \lambda_j g_j(x_i) \right). \end{aligned} \tag{1.11}$$

Since $\delta I'$ vanishes for arbitrary δf_i we have that

$$\begin{aligned}
0 &= \ln f_i + \lambda_0 + \sum_j^m \lambda_j g_j(x_i) \\
\ln f_i &= -\lambda_0 - \sum_j^m \lambda_j g_j(x_i) \\
f_i &= \exp\left(-\lambda_0 - \sum_j^m \lambda_j g_j(x_i)\right). \qquad (1.12)
\end{aligned}$$

By substituting into equation (1.2) λ_0 can be determined

$$\begin{aligned}
1 &= \sum_i^n \exp\left(-\lambda_0 - \sum_j^m \lambda_j g_j(x_i)\right) \\
\ln 1 &= \ln\left(e^{-\lambda_0} \sum_i^n \exp\left(-\sum_j^m \lambda_j g_j(x_i)\right)\right) \\
\lambda_0 &= \ln \sum_i^n \exp\left(-\sum_j^m \lambda_j g_j(x_i)\right) \\
\lambda_0 &= \ln Z(\lambda_1 \ldots \lambda_j), \qquad (1.13)
\end{aligned}$$

where

$$Z(\lambda_1 \ldots \lambda_j) = \sum_i^n \exp\left(-\sum_j^m \lambda_j g_j(x_i)\right), \qquad (1.14)$$

is the partition function.

Substitution of equation (1.12) into equation (1.1) gives a set of equations from which the λ_j can be determined

$$\begin{aligned}
\langle g_j(x) \rangle &= \sum_i^n \exp\left(-\lambda_0 - \sum_h^m \lambda_h g_h(x_i)\right) g_j(x_i) \\
&= e^{-\lambda_0} \sum_i^n \exp\left(-\sum_h^m \lambda_h g_h(x_i)\right) g_j(x_i) \\
&= -\frac{1}{Z} \sum_i^n \frac{\partial}{\partial \lambda_j} \exp\left(-\sum_h^m \lambda_h g_h(x_i)\right)
\end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{Z}\frac{\partial}{\partial\lambda_j}Z \\
&= -Z\frac{1}{Z^2}\frac{\partial}{\partial\lambda_j}Z \\
&= -Z\frac{\partial}{\partial\lambda_j}1/Z \\
&= -\frac{\partial}{\partial\lambda_j}\ln Z,
\end{aligned} \tag{1.15}$$

these equations are often intractable and have to be solved numerically.

The entropy of distribution (1.12) is the maximum entropy that can be obtained in an unbiased fashion with the given prior knowledge. This entropy, with $k = 1$ reduces to

$$\begin{aligned}
S_{\max} &= -\sum_i^n f_i \ln f_i \\
&= -\sum_i^n f_i \ln\left[\exp\left(-\lambda_0 - \sum_j^m \lambda_j\, g_j(x_i)\right)\right] \\
&= -\sum_i^n f_i \left(-\lambda_0 - \sum_j^m \lambda_j\, g_j(x_i)\right) \\
&= \lambda_0 + \sum_j^m \lambda_j \sum_i^n f_i\, g_j(x_i) \\
&= \lambda_0 + \sum_j^m \lambda_j \left\langle g_j(x)\right\rangle \\
&= \ln Z + \sum_j^m \lambda_j \left\langle g_j(x)\right\rangle .
\end{aligned} \tag{1.16}$$

The Lagrange multipliers are also related to the expectation values through the following differential equation

$$\begin{aligned}
\frac{\partial S}{\partial\left\langle g_j\left(x\right)\right\rangle} &= \frac{\partial}{\partial\left\langle g_j\left(x\right)\right\rangle}\left(\lambda_0 + \sum_h^m \lambda_h \left\langle g_h(x)\right\rangle\right) \\
&= \frac{\partial\lambda_0}{\partial\left\langle g_j\left(x\right)\right\rangle} + \sum_h^m\left(\frac{\partial\lambda_h}{\partial\left\langle g_j\left(x\right)\right\rangle}\left\langle g_h(x)\right\rangle + \lambda_h\frac{\partial\left\langle g_h\left(x\right)\right\rangle}{\partial\left\langle g_j\left(x\right)\right\rangle}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_k^m \frac{\partial \ln Z}{\partial \lambda_k} \frac{\partial \lambda_k}{\partial \langle g_j(x) \rangle} + \sum_h^m \frac{\partial \lambda_h}{\partial \langle g_j(x) \rangle} \langle g_h(x) \rangle + \sum_h^m \lambda_h \delta(h-j) \\
&= \sum_k^m \left[- \langle g_k(x) \rangle \frac{\partial \lambda_k}{\partial \langle g_j(x) \rangle} \right] + \sum_h^m \langle g_h(x) \rangle \frac{\partial \lambda_h}{\partial \langle g_j(x) \rangle} + \lambda_j \\
&= \lambda_j. \qquad (1.17)
\end{aligned}
$$

.

From (1.17) another very useful relation is also found,

$$
\begin{aligned}
\frac{\partial \lambda_h}{\partial \langle g_j(x) \rangle} &= \frac{\partial}{\partial \langle g_j(x) \rangle} \frac{\partial S}{\partial \langle g_h(x) \rangle} \\
&= \frac{\partial^2 S}{\partial \langle g_h(x) \rangle \partial \langle g_j(x) \rangle} \\
&= \frac{\partial \lambda_j}{\partial \langle g_h(x) \rangle}. \qquad (1.18)
\end{aligned}
$$

Equations (1.15-1.18) are also known as the Jaynes' relations. Due to their great importance these equations have been summarised in table (1.1). These relations are of crucial importance in that they form the basis of the connection between statistical mechanics and thermodynamics in Jaynes' information-theoretic approach to statistical mechanics (Jaynes (1983)). All the basic equations of equilibrium thermodynamics are particular instances of, or can be derived from, equations (1.15-1.18).

On occasion in the following chapters some of the equations in this section will be repeated, most notably those summarised in table (1.1). This is done so that each chapter will be self-contained in as far as is reasonable.

1.3 The Grand Canonical Ensemble

As an example we will briefly look at the grand canonical ensemble formalism in statistical mechanics. The system has a probability p_i to be in the i'th micro state at any given time. Each micro state i is characterised by a total energy $E(i)$ and a total number of particles $N(i)$. For instance, in the case of a quantum ideal gas the label i represents a string of occupation numbers: $i \rightarrow \left(n_0^{(i)}, n_1^{(i)}, n_2^{(i)}, \ldots\right)$

$$\langle g_j(x)\rangle = -\frac{\partial}{\partial \lambda_j}\ln Z,$$

$$S_{\max} = \ln Z + \sum_j^m \lambda_j \langle g_j(x)\rangle\,,$$

$$\frac{\partial S}{\partial \langle g_j(x)\rangle} = \lambda_j,$$

$$\frac{\partial \lambda_h}{\partial \langle g_j(x)\rangle} = \frac{\partial^2 S}{\partial \langle g_j(x)\rangle\,\partial \langle g_h(x)\rangle} = \frac{\partial \lambda_j}{\partial \langle g_h(x)\rangle}.$$

Table 1.1: Jaynes' relations

where $n_k^{(i)}$ stands for the occupation number of the k'th single particle state and ϵ_k is the energy of the single particle state. In this case we have:

$$N(i) = \sum_k n_k^{(i)}, \tag{1.19}$$

and

$$E(i) = \sum_k n_k^{(i)}\epsilon_k. \tag{1.20}$$

Having as prior knowledge the mean values, $\langle E\rangle = \sum_i p_i E(i)$ and $\langle N\rangle = \sum_i p_i N(i)$ we can write down a maximum entropy ansatz for the probability distribution by considering equations (1.12), (1.13) and (1.14):

$$p_i = \frac{\exp(-\beta E(i) - \alpha N(i))}{\sum_i \exp(-\beta E(i) - \alpha N(i))}, \tag{1.21}$$

where β and α are the Lagrange multipliers corresponding to the mean energy and the mean number of particles of the system respectively. The partition function is then given by

$$Z = \sum_i \exp\left(-\beta E(i) - \alpha N(i)\right). \tag{1.22}$$

As we shall see, the basic equations of thermodynamics are recovered from the Jaynes' relations if we make the following identifications:

$$\beta \rightarrow \frac{1}{kT}, \tag{1.23}$$

and

$$\alpha \rightarrow -\frac{\mu}{kT}, \tag{1.24}$$

where T is the absolute temperature, k is the Boltzmann constant and μ is the chemical potential of the system. Now, according to equation (1.15) we have that the mean energy is equal to

$$\begin{aligned} \langle E \rangle &= -\frac{\partial}{\partial \beta} \ln Z \\ &= -\frac{\partial}{\partial \beta} \ln \left[\sum_i \exp\left(-\beta E(i) - \alpha N(i)\right) \right] \\ &= -\frac{\sum_i \frac{\partial}{\partial \beta} \exp\left(-\beta E(i) - \alpha N(i)\right)}{\sum_i \exp\left(-\beta E(i) - \alpha N(i)\right)} \\ &= -\frac{\sum_i E(i) \exp\left(-\beta E(i) - \alpha N(i)\right)}{\sum_i \exp\left(-\beta E(i) - \alpha N(i)\right)}, \end{aligned} \tag{1.25}$$

and similarly for the mean number of particles we have

$$\begin{aligned} \langle N \rangle &= -\frac{\partial}{\partial \alpha} \ln \left[\sum_i \exp\left(-\beta E(i) - \alpha N(i)\right) \right] \\ &= -\frac{\sum_i N(i) \exp\left(-\beta E(i) - \alpha N(i)\right)}{\sum_i \exp\left(-\beta E(i) - \alpha N(i)\right)}. \end{aligned} \tag{1.26}$$

From these relations, that are well known in statistical physics, the values for β and α can be determined. Using the results of (1.16) and remembering that our entropic measure should now include the Boltzmann constant we find that

$$\begin{aligned} S &= k\sum_i p_i \ln p_i \\ S/k &= \ln Z + \beta \langle E \rangle + \alpha \langle N \rangle \\ \ln Z &= S/k - \beta \langle E \rangle - \alpha \langle N \rangle \\ &= \frac{TS - \langle E \rangle - \mu \langle N \rangle}{kT} \end{aligned} \tag{1.27}$$

where from thermodynamics $\mu \langle N \rangle$ is the Gibbs free energy

$$G = \langle E \rangle - TS + PV, \tag{1.28}$$

where P is the pressure and V is the volume of the system. Equation (1.27) can then be written as

$$\ln Z = \frac{PV}{kT}, \tag{1.29}$$

which is a well known equation in statistical physics. Lastly by considering equation (1.17) we get

$$\frac{\partial S}{\partial \langle E \rangle} = k\beta = \frac{1}{T}, \tag{1.30}$$

and

$$\frac{\partial S}{\partial \langle N \rangle} = k\alpha = -\frac{\mu}{T} \tag{1.31}$$

Summarising, we see that the thermodynamic formalism can be recovered from Jaynes' maximum entropy formalism. In general, the extensive thermodynamic quantities (energy, number of particles, etc.) are to be identified with the relevant mean values used as constraints in the maximum entropy approach. On the other hand, the intensive thermodynamic quantities (temperature, chemical potential, etc.) are related to the Lagrange multipliers appearing in the maximum entropy distribution.

Chapter 2

Maximum Entropy Approach to the Collisional Vlasov Equation: Exact Solutions

2.1 Introduction

The collisional Vlasov equation is also known as the Boltzmann equation (Cercignani (1988)), it describes the statistical distribution of particles in a low density fluid (gas), and is used to study the transport of physical quantities (such as heat, charge and density) through a fluid. The Boltzmann equation is only valid when the density of the gas is not too high so that collisions between three or more particles doesn't occur. It is used to study diverse systems such as collisional plasmas and a Brownian particle moving through a medium (El-Hanbaly and El-garayhi (1998)), and is used in cosmology for example to model galaxy formation (Steinmetz (1999)) by considering each star as a particle and assuming that the long-range "collisions" only take place between pairs of stars. When collisions are ignored (or the particles are only weakly interacting) the normal Vlasov (collisionless Boltzmann) equation is used to describe the short time behaviour of a system (Binney and Tremaine (1987)), (the collision term on the right of equation (2.1) is then equal to zero).

The Boltzmann equation (Uhlenbeck (1957)) is a time evolution equation of the particle density and in it's usual form it reads,

2. MAXIMUM ENTROPY APPROACH TO THE COLLISIONAL VLASOV EQUATION: EXACT SOLUTIONS

$$\frac{\partial f}{\partial t} + v_a \frac{\partial f}{\partial x_a} + X_a \frac{\partial f}{\partial v_a} = \int dv_1 \iint g\, I\,(g, \theta)\,(f' f_1' - f f_1)\, d\Omega, \tag{2.1}$$

where

$$X_a = -\frac{\partial}{\partial x_a} U, \tag{2.2}$$

and the density function $f(x, v, t)$ is a time dependent function in velocity and position space. U is the potential of an external force. The term on the right hand side of the equation is the collision term and $g = |v - v_1| = |v' - v_1'|$ is the relative velocity that turns over the angle θ during the collision. The prime and index of the $f's$ refer only to the velocity variable. $I\,(g, \theta)\, d\Omega$ is the differential collision cross-section for a collision in the solid angle $d\Omega$.

Recently, El-Wakil, Elhanbaly and Abdou (from now on EEA) in (El-Wakil *et al.* (2003)) proposed an interesting maximum entropy scheme for solving a particular instance of the collisional Vlasov equation. In (El-Wakil *et al.* (2003)), however, EEA obtained only *approximate* maximum entropy solutions of the alluded evolution equation. The aim is to show that the maximum entropy method can be used also to generate *exact* solutions of the collisional Vlasov equation studied by EEA.

In this chapter we consider two different ways for obtaining exact maximum entropy solutions of the aforementioned equation. On the one hand, we identify an appropriate set of five relevant mean values (moments) that evolve according to a closed set of coupled, ordinary, linear differential equations. We show that there are exact maximum entropy solutions associated with that set of moments. These solutions can be studied focusing either on the equations of motion of the moments themselves, or on the equations of motion of the corresponding Lagrange multipliers. On the other hand, we prove that it is possible to obtain exact solutions of the reduced equation considered by EEA, if the zeroth-order moment of the solutions is explicitly taken into account.

Exact maximum entropy solutions of the collisional Vlasov equation itself are discussed in Section II. Exact maximum entropy solutions of EEA's reduced equation are investigated in Section III. Finally, some conclusions are drawn in Section IV.

2.2 Direct Maximum Entropy Approach to the Collisional Vlasov Equation

The collisional Vlasov equation that EEA studies has a Fokker-Planck collision term (the term on the right of the equation), and is given by

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} - \frac{\partial \phi}{\partial x}\frac{\partial f}{\partial v} = \gamma\frac{\partial}{\partial v}\left[vf(x,v,t) + \alpha\frac{\partial f(x,v,t)}{\partial v}\right]. \tag{2.3}$$

The Fokker-Planck term (Cercignani (1988)) is derived when long range interactions that take place over relatively long periods of time, eg. plasmas, galactic systems etc., need to be considered.

After taking the derivative on the right and rearranging the terms in equation (2.3) it becomes,

$$\frac{\partial f}{\partial t} = -v\frac{\partial f}{\partial x} + \left[\frac{\partial \phi}{\partial x} + \gamma v\right]\frac{\partial f}{\partial v} + \gamma\alpha\frac{\partial^2 f}{\partial v^2} + \gamma f, \tag{2.4}$$

where γ and α are positive constants, and the potential ϕ is of a quadratic form,

$$\phi(x) = \phi_0 + \phi_1 x + \frac{1}{2}\phi_2 x^2. \tag{2.5}$$

Here we are also going to assume that $\phi_2 > 0$.

Let us now consider a maximum entropy ansatz (1.12) of the form

$$f(x,v,t) = \exp[-\lambda_0 - \lambda_1 x - \lambda_2 v - \lambda_3 x^2 - \lambda_4 xv - \lambda_5 v^2], \tag{2.6}$$

where the $\lambda_i, \quad i = 0,\ldots,5$ are appropriate Lagrange multipliers. The distribution (2.6) maximizes the Boltzmann-Gibbs entropic functional,

$$S[f] = -\int f(x,v,t)\,\ln f(x,v,t)\,dxdv, \tag{2.7}$$

under the constraints imposed by normalization and the instantaneous mean values of the quantities $B_1 = x$, $B_2 = v$, $B_3 = x^2$, $B_4 = xv$, and $B_5 = v^2$. All the time dependence of the ansatz (2.6) is through the Lagrange multipliers λ_i, which are time dependent. Inserting the ansatz (2.6) into the partial differential equation (2.4), one obtains

$$\begin{aligned}
0 \;=\; & -f\left(-\frac{d\lambda_0}{dt} - x\frac{d\lambda_1}{dt} - v\frac{d\lambda_2}{dt} - x^2\frac{d\lambda_3}{dt} - vx\frac{d\lambda_4}{dt} - v^2\frac{d\lambda_5}{dt}\right) \\
& - vf\left(-\lambda_1 - 2x\lambda_3 - v\lambda_4\right) \\
& + f\left(\phi_1 + \gamma v + \phi_2 x\right)\left(-\lambda_2 - x\lambda_4 - 2v\lambda_5\right) \\
& + \gamma\alpha f\left(-2\lambda_5 + \left(-\lambda_2 - x\lambda_4 - 2v\lambda_5\right)^2\right) \;+\; f\gamma \\
=\; & -\phi_1\lambda_2 \;+\; \gamma\alpha\lambda_2^2 \;-\; 2\gamma\alpha\lambda_5 \;+\; \gamma + \frac{d\lambda_0}{dt} \\
& + x\left(-\phi_1\lambda_4 - \phi_2\lambda_2 + 2\gamma\alpha\lambda_4\lambda_2 + \frac{d\lambda_1}{dt}\right) \\
& + v\left(\lambda_1 - \gamma\lambda_2 - 2\phi_1\lambda_5 + 4\gamma\alpha\lambda_2\lambda_5 + \frac{d\lambda_2}{dt}\right) \\
& + x^2\left(-\phi_2\lambda_4 + \gamma\alpha\lambda_4^2 + \frac{d\lambda_3}{dt}\right) \\
& + vx\left(2\lambda_3 - 2\phi_2\lambda_5 - \gamma\lambda_4 + 4\gamma\alpha\lambda_4\lambda_5 + \frac{d\lambda_4}{dt}\right) \\
& + v^2\left(\lambda_4 + 4\gamma\alpha\lambda_5^2 - 2\gamma\lambda_5 + \frac{d\lambda_5}{dt}\right),
\end{aligned} \tag{2.8}$$

and then equating to zero separately terms proportional to $x^i v^j$ with different exponents i, j, it is clear that the ansatz (2.6) constitutes an exact solution to (2.4), provided that the Lagrange multipliers comply with the set of coupled ordinary differential equations,

$$\frac{d\lambda_0}{dt} = \phi_1\lambda_2 - \gamma\alpha\lambda_2^2 + 2\gamma\alpha\lambda_5 - \gamma, \tag{2.9}$$

$$\frac{d\lambda_1}{dt} = \phi_1\lambda_4 + \phi_2\lambda_2 - 2\gamma\alpha\lambda_4\lambda_2, \tag{2.10}$$

$$\frac{d\lambda_2}{dt} = -\lambda_1 + \gamma\lambda_2 + 2\phi_1\lambda_5 - 4\gamma\alpha\lambda_2\lambda_5, \tag{2.11}$$

$$\frac{d\lambda_3}{dt} = \phi_2\lambda_4 - \gamma\alpha\lambda_4^2, \tag{2.12}$$

$$\frac{d\lambda_4}{dt} = -2\lambda_3 + 2\phi_2\lambda_5 + \gamma\lambda_4 - 4\gamma\alpha\lambda_4\lambda_5, \tag{2.13}$$

and

$$\frac{d\lambda_5}{dt} = -\lambda_4 - 4\gamma\alpha\lambda_5^2 + 2\gamma\lambda_5. \tag{2.14}$$

Alternatively, we can focus our attention on the set of ordinary differential equations governing the evolution of the selected set of relevant mean values, obtaining

$$\begin{aligned}
\frac{d}{dt}\langle x\rangle &= \int\!\!\int x\frac{df}{dt}\,dx\,dv \\
&= \int\!\!\int\left[-vx\frac{\partial f}{\partial x} + x\left(\phi_1+\phi_2 x+\gamma v\right)\frac{\partial f}{\partial v} + \gamma\alpha x\frac{\partial^2 f}{\partial v^2} + \gamma x f\right]dx\,dv \\
&= \int -v\int x\frac{\partial f}{\partial x}\,dx\,dv + \int x\left(\phi_1+\phi_2 x\right)\int\frac{\partial f}{\partial v}\,dv\,dx \\
&\quad + \int x\int\gamma v\frac{\partial f}{\partial v}dv\,dx + \gamma\alpha\int x\int\frac{\partial^2 f}{\partial v^2}\,dv\,dx + \gamma\int\!\!\int xf\,dx\,dv \\
&= \int vf\,dv + 0 - \gamma\int xf\,dx + 0 + \gamma\int xf\,dx \\
&= \langle v\rangle,
\end{aligned} \tag{2.15}$$

and in a similar fashion,

$$\frac{d}{dt}\langle v\rangle = -\phi_1 - \phi_2\langle x\rangle - \gamma\langle v\rangle, \tag{2.16}$$

$$\frac{d}{dt}\langle x^2\rangle = 2\langle xv\rangle, \tag{2.17}$$

$$\frac{d}{dt}\langle xv\rangle = -\phi_1\langle x\rangle - \phi_2\langle x^2\rangle - \gamma\langle xv\rangle + \langle v^2\rangle, \tag{2.18}$$

and

$$\frac{d}{dt}\langle v^2\rangle = -2\phi_1\langle v\rangle - 2\phi_2\langle xv\rangle - 2\gamma\langle v^2\rangle + 2\alpha\gamma. \tag{2.19}$$

Changing appropriately the origin of the x-coordinate, it is possible to set the linear term in the potential equal to zero. Consequently, and without loss of generality, we can set the coefficient $\phi_1 = 0$. In that case, the differential equations

(2.15-2.16) governing the evolution of the mean values $\langle x \rangle$ and $\langle v \rangle$ are decoupled from the three equations (2.17-2.19) governing the evolution of $\langle x^2 \rangle$, $\langle xv \rangle$, and $\langle v^2 \rangle$. The differential equations (2.15-2.16) admit the particular (linearly independent) solutions

$$\begin{aligned} \langle x \rangle_{1,2} &= -\left(\frac{\gamma + \sigma_{1,2}}{\phi_2}\right) \exp(\sigma_{1,2} t) \\ \langle v \rangle_{1,2} &= \exp(\sigma_{1,2} t), \end{aligned} \tag{2.20}$$

where

$$\sigma_{1,2} = \frac{1}{2} \left\{ -\gamma \pm \sqrt{\gamma^2 - 4\phi_2} \right\}. \tag{2.21}$$

The general solution for the equations (2.15-2.16) is then,

$$\begin{aligned} \langle x \rangle &= D_1 \langle x \rangle_1 + D_2 \langle x \rangle_2 \\ \langle v \rangle &= D_1 \langle v \rangle_1 + D_2 \langle v \rangle_2, \end{aligned} \tag{2.22}$$

where $D_{1,2}$ are appropriate coefficients.

The equations (2.17-2.19) constitute a closed set of inhomogeneous linear differential equations admitting the particular solution

$$\mathbf{W}_0 = \begin{pmatrix} \langle x^2 \rangle_0 \\ \langle xv \rangle_0 \\ \langle v^2 \rangle_0 \end{pmatrix} = \begin{pmatrix} \alpha/\phi_2 \\ 0 \\ \alpha \end{pmatrix}. \tag{2.23}$$

This stationary solution, along with the stationary solution $\langle x \rangle_0 = 0$, $\langle v \rangle_0 = 0$ of equations (2.15-2.16), corresponds to the stationary solution of the collisional Vlasov equation (2.4), exhibiting a Maxwellian velocity distribution.

The homogeneous set of differential equations associated with (2.17-2.19) can be cast under the guise

$$\frac{d}{dt} \begin{pmatrix} \langle x^2 \rangle \\ \langle xv \rangle \\ \langle v^2 \rangle \end{pmatrix} = \mathbf{A} \cdot \begin{pmatrix} \langle x^2 \rangle \\ \langle xv \rangle \\ \langle v^2 \rangle \end{pmatrix}, \tag{2.24}$$

where

$$\mathbf{A} = \begin{pmatrix} 0 & 2 & 0 \\ -\phi_2 & -\gamma & 1 \\ 0 & -2\phi_2 & -2\gamma \end{pmatrix}. \tag{2.25}$$

The general solution of the homogeneous set of differential equations is

$$\mathbf{W}_{\text{homog.}} = \begin{pmatrix} \langle x^2 \rangle_{homog.} \\ \langle xv \rangle_{homog.} \\ \langle v^2 \rangle_{homog.} \end{pmatrix} = \sum_{i=1}^{3} c_i \, e^{l_i \, t} \, \mathbf{W}_i, \tag{2.26}$$

where the (constant) coefficients c_i are determined by the initial conditions, and $(\mathbf{W}_i, l_i, \quad i = 1, 2, 3)$ are the eigenvectors and eigenvalues of the matrix $\mathbf{A}$. The general solution to the equations is then,

$$\mathbf{W} = \begin{pmatrix} \langle x^2 \rangle \\ \langle xv \rangle \\ \langle v^2 \rangle \end{pmatrix} = \mathbf{W}_0 + \mathbf{W}_{\text{homog.}}. \tag{2.27}$$

Now, the eigenvalues l of the matrix (2.25) are the roots of the equation

$$l^3 + 3\gamma \, l^2 + (2\gamma^2 + 4\phi_2) l + 4\phi_2 \gamma = 0, \tag{2.28}$$

which has the three roots,

$$l_1 = -\gamma + \sqrt{\gamma^2 - 4\phi_2}, \tag{2.29}$$

$$l_2 = -\gamma - \sqrt{\gamma^2 - 4\phi_2}, \tag{2.30}$$

and

$$l_3 = -\gamma, \tag{2.31}$$

with the concomitant eigenvectors

$$\mathbf{W}_1 = \begin{pmatrix} \frac{1}{4\phi_2^2} \left[\gamma + \sqrt{\gamma^2 - 4\phi_2} \right]^2 \\ -\frac{1}{2\phi_2} \left[\gamma + \sqrt{\gamma^2 - 4\phi_2} \right] \\ 1 \end{pmatrix}, \tag{2.32}$$

$$\mathbf{W}_2 = \begin{pmatrix} \frac{1}{4\phi_2^2}\left[\gamma - \sqrt{\gamma^2 - 4\phi_2}\right]^2 \\ -\frac{1}{2\phi_2}\left[\gamma - \sqrt{\gamma^2 - 4\phi_2}\right] \\ 1 \end{pmatrix}, \tag{2.33}$$

and

$$\mathbf{W}_3 = \begin{pmatrix} \frac{1}{\phi_2} \\ -\frac{\gamma}{2\phi_2} \\ 1 \end{pmatrix}. \tag{2.34}$$

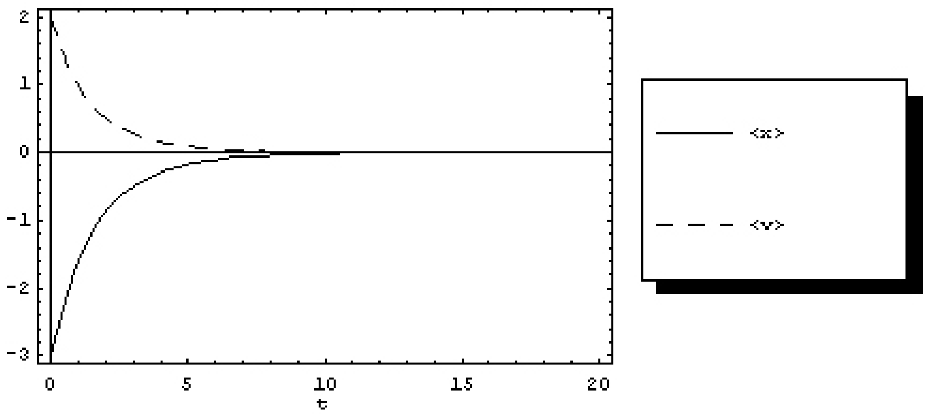

Figure 2.1: The time evolution of the expectation values of x and v

We can see that the three eigenvalues of the matrix (2.25) have negative real parts. Consequently, we have that, in the limit $t \to \infty$, the general solution of equations (2.17-2.19) tends to the stationary solution (with a Maxwellian distribution). It is important to realize that the mean values associated with any solution of the collisional Vlasov equation (being it of a maximum entropy form or not) evolve towards the aforementioned stationary values. This illustrates the fact that any exact solution of the collisional Vlasov equation relaxes towards the stationary distribution exhibiting a Maxwellian velocity distribution. This can also be proved by recourse to an appropriate H-theorem.

In figure (2.1) and (2.2) the time evolution of the mean values, $\langle x \rangle$, $\langle v \rangle$, $\langle x^2 \rangle$, $\langle xv \rangle$, and $\langle v^2 \rangle$ are plotted. The arbitrarily chosen values for the constants are:

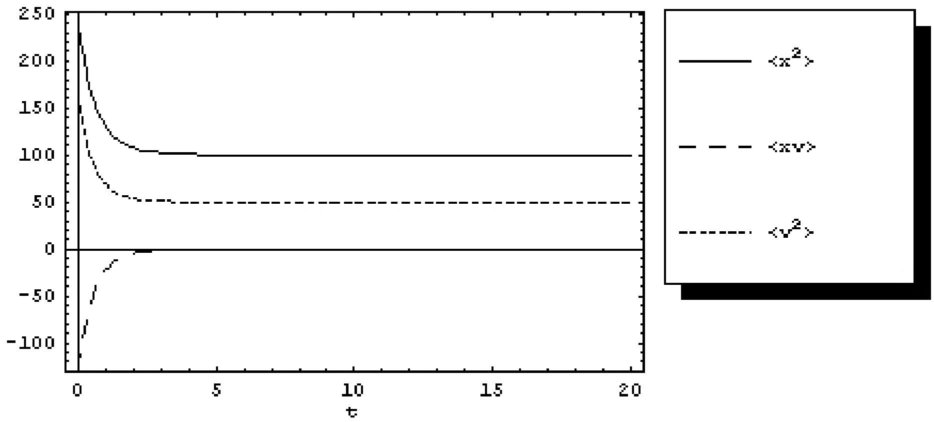

Figure 2.2: The time evolution of the expectation values of x^2, xv and v^2

$\gamma = 1.5$, $\phi_2 = 0.5$, $D_1 = D_2 = 1$, $\alpha = 50$, $c_1 = 10$, $c_2 = 100$, $c_3 = 1$. As expected the mean values evolve to a stationary solution.

The maximum entropy solutions based upon the set of relevant mean values can be analysed focusing either on the equations of motion for the moments themselves, or on the equations of motion of the corresponding Lagrange multipliers. Each of these two viewpoints is, of course, equivalent to the other one. The "translation" between the mean values language and the Lagrange multipliers language can be effected by recourse to the partition function,

$$\begin{aligned} Z &= \exp[\lambda_0] \\ &= \int \exp[-\lambda_1 x - \lambda_2 v - \lambda_3 x^2 - \lambda_4 xv - \lambda_5 v^2] dx dv. \end{aligned} \tag{2.35}$$

Evaluating explicitly the Gaussian integrals associated with Z (assuming $\lambda_3, \lambda_5 > 0$ and $4\lambda_3\lambda_5 - \lambda_4^2 > 0$) we obtain,

$$Z = \frac{2\pi}{\sqrt{4\lambda_3\lambda_5 - \lambda_4^2}} \exp\left\{ \frac{\lambda_1^2\lambda_5 + \lambda_2^2\lambda_3 - \lambda_1\lambda_2\lambda_4}{4\lambda_3\lambda_5 - \lambda_4^2} \right\}. \tag{2.36}$$

The relevant mean values $\langle B_i \rangle$ and the associated Lagrange multipliers λ_i are related by the Jaynes' relations (see equations (1.15) and (1.17))

$$\lambda_i = \frac{\partial}{\partial\langle B_i\rangle} S, \tag{2.37}$$

and

$$\langle B_i\rangle = -\frac{\partial}{\partial\lambda_i} (\ln Z). \tag{2.38}$$

If the time dependent Lagrange multipliers are determined by solving the equations of motion, then the (again, time dependent) mean values can be directly evaluated from equation (2.38).Thus, by substituting equation (2.36) into equation (2.38), we can write the mean values in terms of the Lagrange multipliers as:

$$\langle x\rangle = \frac{2\lambda_1\lambda_5 - \lambda_2\lambda_4}{\lambda_4^2 - 4\lambda_3\lambda_5}, \tag{2.39}$$

$$\langle v\rangle = \frac{2\lambda_2\lambda_3 - \lambda_1\lambda_4}{\lambda_4^2 - 4\lambda_3\lambda_5}, \tag{2.40}$$

$$\langle x^2\rangle = \frac{\lambda_2^2\lambda_4^2 - 2\lambda_4\left(2\lambda_1\lambda_2 + \lambda_4\right)\lambda_5 + 4\left(\lambda_1^2 + 2\lambda_3\right)\lambda_5^2}{\left(\lambda_4^2 - 4\lambda_3\lambda_5\right)^2}, \tag{2.41}$$

$$\langle xv\rangle = \frac{\lambda_4^3 - 2\lambda_2^2\lambda_3\lambda_4 - 2\left(\lambda_1^2 + 2\lambda_3\right)\lambda_4\lambda_5 + \lambda_1\lambda_2\left(\lambda_4^2 + 4\lambda_3\lambda_5\right)}{\left(\lambda_4^2 - 4\lambda_3\lambda_5\right)^2}, \tag{2.42}$$

and

$$\langle v^2\rangle = \frac{4\lambda_2^2\lambda_3^2 - 4\lambda_1\lambda_2\lambda_3\lambda_4 + \left(\lambda_1^2 - 2\lambda_3\right)\lambda_4^2 + 8\lambda_3^2\lambda_5}{\left(\lambda_4^2 - 4\lambda_3\lambda_5\right)^2}. \tag{2.43}$$

After having solved equations (2.9-2.14) for the Lagrange multipliers we can now in principle plot graphs for the expectation values, or determine the expectation values at a specific time.

On the other hand, if the equations of motion of the relevant mean values are solved, equation (2.38) can again be used, now to solve for the time dependent Lagrange multipliers. Here we take equations (2.39-2.43) and solve for the Lagrange multipliers in terms of the relevant mean values:

$$\lambda_1 = \frac{\langle v^2\rangle\langle x\rangle - \langle v\rangle\langle xv\rangle}{\langle v\rangle^2 + \langle v^2\rangle^2\langle x\rangle - 2\langle v\rangle\langle v^2\rangle\langle xv\rangle - \langle x\rangle\langle x^2\rangle + \langle xv\rangle^2\langle x^2\rangle}, \tag{2.44}$$

$$\lambda_2 = \frac{\langle xv\rangle\langle x^2\rangle - \langle v\rangle\langle v^2\rangle}{\langle v\rangle^2 + \langle v^2\rangle^2\langle x\rangle - 2\langle v\rangle\langle v^2\rangle\langle xv\rangle - \langle x\rangle\langle x^2\rangle + \langle xv\rangle^2\langle x^2\rangle}, \tag{2.45}$$

$$\lambda_3 = \frac{\langle xv\rangle^2 - \langle x\rangle}{2\left(\langle v\rangle^2 + \langle v^2\rangle^2\langle x\rangle - 2\langle v\rangle\langle v^2\rangle\langle xv\rangle - \langle x\rangle\langle x^2\rangle + \langle xv\rangle^2\langle x^2\rangle\right)}, \tag{2.46}$$

$$\lambda_4 = \frac{\langle v\rangle - \langle v^2\rangle\langle xv\rangle}{\langle v\rangle^2 + \langle v^2\rangle^2\langle x\rangle - 2\langle v\rangle\langle v^2\rangle\langle xv\rangle - \langle x\rangle\langle x^2\rangle + \langle xv\rangle^2\langle x^2\rangle}, \tag{2.47}$$

and

$$\lambda_5 = \frac{\langle v^2\rangle^2 - \langle x^2\rangle}{2\left(\langle v\rangle^2 + \langle v^2\rangle^2\langle x\rangle - 2\langle v\rangle\langle v^2\rangle\langle xv\rangle - \langle x\rangle\langle x^2\rangle + \langle xv\rangle^2\langle x^2\rangle\right)}. \tag{2.48}$$

It is then possible to write the maximum entropy solution explicitly in term of the relevant mean values (which are known through the prior knowledge). Thus we can draw a graph of the time evolution of the maximum entropy solution (that is, we can plot $f(x, v, t)$ as a function of (x, v) at any given time).

2.3 EEA Reduced Evolution Equation

After an appropriate change of variables, EEA transformed their original Vlasov equation,(2.4), into an evolution equation of the form (see, for instance, equations (11), (23), (26), and (36) of El-Wakil *et al.* (2003))

$$\frac{\partial F}{\partial s} = m\frac{\partial^2 F}{\partial z^2} + (nz + p)\frac{\partial F}{\partial z} + (q_1 + q_2 z)F, \tag{2.49}$$

s and z playing respectively the roles of the temporal and spatial variables, and m, n, p, q_1, and q_2 being constants (notice that, when writing down the evolution equation (2.49) our notation differs slightly from that of EEA. For instance, in connection with equation (11) of EEA, we have, $h_1 = m$, $h_3 = (q_1 + q_2 z)$, $h_4 = n$, $h_5 = p$). By recourse to an ingenious procedure based upon group-theoretical ideas, they obtained the explicit time dependence of the first two moments of $F(z, s)$, namely

$$\langle z \rangle(s) = \int F(z, s)\, z\, dz, \tag{2.50}$$

and

$$\langle z^2 \rangle(s) = \int F(z, s)\, z^2\, dz. \tag{2.51}$$

Finally, taking the instantaneous values of $\langle z \rangle(s)$ and $\langle z^2 \rangle(s)$, at each time s, as constraints, they found the concomitant maximum entropy distribution F_{ME}, which constitutes an approximate solution to the evolution equation (2.49). *Before going on, it is important to realize that, according to the developments presented in El-Wakil* et al. *(2003), each exact solution of EEA's reduced equation corresponds to an exact solution of the the original Vlasov equation (2.4). Obviously, such an exact solution is going to share all the general properties of the solutions of (2.4). In particular, such a solution is going to relax to the stationary solution associated with a Maxwellian velocity distribution.*

The point we want to make here is that, if the zeroth order moment,

$$N(s) = \int F(z, s)\, dz, \tag{2.52}$$

is explicitly incorporated as a constraint in the entropy maximization scheme, the maximum entropy procedure leads to exact solutions of the evolution equation (2.49). It is important to realize that the evolution equation (2.49) does not preserve the normalization of the solution F. In point of fact, taking the time derivative of the zeroth order moment of (2.49) we get,

$$\frac{dN}{ds} = (q_1 - n)N + q_2 \langle z \rangle, \tag{2.53}$$

which is, in general, different from zero. Now, if taking as the only constraints the instantaneous values of $\langle z \rangle$ and $\langle z^2 \rangle$ (as EEA does in El-Wakil *et al.* (2003)) we build up a normalized (that is, $N = 1$) maximum entropy solution for (2.49), we are disregarding an important piece of information concerning the behaviour of $F(z, s)$. Such a maximum entropy approximation has, by construction, a constant normalization, although we know that the exact solution has a time dependent one. Consequently, such a procedure is bound to yield only an approximate solution to (2.49). However, as we shall presently show, *exact* solutions can be obtained if the time dependence of N is explicitly taken into account in the maximum entropy procedure.

In connection with the new constraint N, a few remarks are in order. It is usually assumed that entropic concepts and, in particular, the maximum entropy principle, can be applied only to *probability distributions*. In order to be interpreted as a probability distribution, a function ρ must be non-negative and normalized to one. However, entropic concepts can be profitably applied also to the description of (positive) *densities*, which are non-negative quantities not necessarily normalized to 1. A (positive) density can be normalized to any positive number N. The application of the maximum entropy principle to the study of densities allows for the discussion of a wide family of interesting scenarios. For instance, densities may evolve according to non-linear evolution equations (Borland *et al.* (1999); da Silva *et al.* (2004); Plastino (2001)) (as opposed to ensemble probabilities which, strictly speaking, must evolve linearly, see van Kampen (1992)). In this regard, it is important to remember that Boltzmann himself introduced his celebrated entropic functional in order to study the evolution of the density of particles in the $(\mathbf{x}, \mathbf{v})$ space which, by the way, evolves according to a non-linear transport equation. When applying the maximum entropy principle to the evolution of a density the normalization N may even change with time. This is precisely the case with the (linear) problem studied by EEA in El-Wakil *et al.* (2003).

As already mentioned, our present purpose is to show that, when the zeroth moment (2.52) is explicitly incorporated to the maximum entropy solution based on the first and second moments (2.50-2.51), an exact solution of (2.49) is obtained. In order to do that we consider the concomitant maximum entropy ansatz based on the three constraints (2.50-2.52). The maximum entropy ansatz has the form,

$$F(z,s) \;=\; \exp\left[-\lambda_0(s) - \lambda_1(s)z - \lambda_2(s)z^2\right], \tag{2.54}$$

where the $\lambda_i(s)$'s are time dependent Lagrange multipliers. Alternatively, the ansatz can be written explicitly in terms of the relevant three moments (which account for its time dependence),

$$F(z,s) \;=\; \frac{N}{[2\pi(\frac{\langle z^2\rangle}{N} - \frac{\langle z\rangle^2}{N^2})]^{1/2}} \exp\left[\frac{-\left(z - \frac{\langle z\rangle}{N}\right)^2}{2[\frac{\langle z^2\rangle}{N} - \frac{\langle z\rangle^2}{N^2}]}\right]. \tag{2.55}$$

Which after setting

$$a = \frac{\langle z^2\rangle}{N} \quad \text{and} \quad b = \frac{\langle z\rangle}{N}, \tag{2.56}$$

becomes

$$F(z,s) \;=\; \frac{N}{[2\pi(a-b^2)]^{1/2}} \exp\left[\frac{-\left(z-b\right)^2}{2\left(a-b^2\right)}\right], \tag{2.57}$$

Inserting the maximum entropy ansatz (2.57) into the partial differential equation (2.49) we get (the $\prime$'s designate total derivatives with respect to s),

$$\begin{aligned}
0 \;&=\; m\frac{\partial^2 F}{\partial z^2} + (nz+p)\frac{\partial F}{\partial z} + (q_1 + q_2 z)F - \frac{\partial F}{\partial s} \\
&=\; z^2\left(\frac{m}{(a-b^2)^2} - \frac{n}{a-b^2} - \frac{a'-2bb'}{2(a-b^2)^2}\right) \\
&\quad + z\left(q_2 - \frac{2mb}{(a-b^2)^2} - \frac{p}{a-b^2} + \frac{nb}{a-b^2} - \frac{b'}{a-b^2} + \frac{b(a'-2bb')}{(a-b^2)^2}\right) \\
&\quad - \frac{N'}{N} + q_1 + \frac{mb^2}{(a-b^2)^2} - \frac{m}{a-b^2} + \frac{pb}{a-b^2} + \frac{bb'}{a-b^2} \\
&\quad - \frac{b^2(a'-2bb')}{2(a-b^2)^2} + \frac{a'-2bb'}{2(a-b^2)}.
\end{aligned} \tag{2.58}$$

Equating separately to zero the coefficients corresponding to different powers of z we obtain,

$$0 \;=\; \frac{m}{(a-b^2)^2} - \frac{n}{a-b^2} - \frac{a'-2bb'}{2(a-b^2)^2}, \tag{2.59}$$

$$0 = q_2 + \frac{nb - p - b'}{a - b^2} + \frac{b(a' - 2bb') - 2mb}{(a - b^2)^2}, \tag{2.60}$$

and

$$0 = q_1 - \frac{N'}{N} + \frac{2mb^2 - b^2(a' - 2bb')}{2(a - b^2)^2} + \frac{a' - 2bb' + 2p\,b - 2m + 2bb'}{2(a - b^2)}. \tag{2.61}$$

Equations (2.59) and (2.60) (considering the combination (2.60) + $2b$(2.59)) lead to

$$b' = -p + q_2 a - nb - q_2 b^2. \tag{2.62}$$

Combining now equations (2.59) and (2.62) we get

$$a' = 2(m - na - pb + q_2 ab - q_2 b^3). \tag{2.63}$$

Finally, from equations (2.61), (2.62), and (2.63), we derive the evolution equation for the parameter N,

$$N' = N\,(q_1 - n + q_2 b). \tag{2.64}$$

The three equations (2.62-2.64) constitute a closed set of (coupled) ordinary differential equations for the three parameters a, b, and N. It is clear that the maximum entropy ansatz (2.55) constitutes an *exact solution* to the evolution equation (2.49), provided that the three parameters a, b, and N evolve according to the equations (2.62-2.64).

The time derivatives of the three moments N, $\langle z \rangle$, and $\langle z^2 \rangle$, can also be obtained taking the corresponding moments of equation (2.49). That is, multiplying equation (2.49) respectively by 1, z, and z^2, and integrating over z. One obtains,

$$\frac{dN}{ds} = (q_1 - n)N + q_2 \langle z \rangle, \tag{2.65}$$

$$\frac{d\langle z \rangle}{ds} = -pN + (q_1 - 2n)\,\langle z \rangle + q_2 \langle z^2 \rangle, \tag{2.66}$$

and

$$\frac{d\langle z^2\rangle}{ds} = 2mN - 3n\langle z^2\rangle - 2p\langle z\rangle + q_1\langle z^2\rangle + q_2\langle z^3\rangle. \tag{2.67}$$

In the case $q_2 = 0$, it can be directly verified that the above three equations are equivalent to the equations (2.62-2.64) (taking into account (2.56)).

Now, the set of coupled ordinary differential equations (2.62-2.64) for the parameters b, a, and N, admits a closed analytical solution given by,

$$\begin{aligned} b \;=\; & \frac{1}{n^2}e^{-ns}\Big[nq_2\big(a_0 - b_0^2\big)\left(1-e^{-ns}\right) + mq_2 + \left(e^{ns}+e^{-ns}-2\right) \\ & \qquad - pn\left(e^{ns}-1\right) + b_0 n^2\Big], \end{aligned} \tag{2.68}$$

$$\begin{aligned} a = \frac{1}{n^4}\Big[& a_0 n^4 e^{-2ns} + 2b_0 p n^3 e^{-ns}\left(e^{-ns}-1\right) - mn^3\left(e^{-2ns}-1\right) \\ + \;& 2q_2 n^3 e^{-2ns}\left(e^{-ns}-1\right)\left(b_0^3 - a_0 b_0\right) + p^2 n^2\left(e^{-ns}-1\right)^2 \\ + \;& 2b_0 m q_2 n^2 e^{-ns}\left(e^{-ns}-1\right)^2 + 2pq_2 n^2 e^{-ns}\left(e^{-ns}-1\right)^2\left(b_0^2 - a_0\right) \\ + \;& q_2^2 n^2 e^{-2ns}\left(e^{-ns}-1\right)^2\left(b_0^2-a_0\right)^2 + 2mpq_2 n\left(e^{-ns}-1\right)^3 \\ + \;& 2mq_2^2 n e^{-ns}\left(e^{-ns}-1\right)^3\left(b_0^2-a_0\right) + m^2 q_2^2\left(e^{-ns}-1\right)^4\Big], \end{aligned} \tag{2.69}$$

and

$$\begin{aligned} N \;=\; & N_0 \exp\Bigg\{\frac{1}{2n^3}\Big[-2b_0 n^2 q_2\left(e^{-ns}-1\right) \\ & \qquad - n\,q_2^2\left(e^{-ns}-1\right)^2\left(b_0^2-a_0\right) - 2np\,q_2\left(e^{-ns}-1\right)\Big] \\ & \times e^{\frac{1}{2n^3}\left(-m\,q_2^2\left(3-4e^{-ns}+e^{-2ns}\right)-2s\left(n^4-n^3q_1+n^2p\,q_2-m\,n\,q_2^2\right)\right)}\Bigg\}, \end{aligned} \tag{2.70}$$

where b_0, a_0, and N_0, are the initial values of these parameters at $s = 0$.

2.4 Conclusions

Summing up, we have re-visited the maximum entropy approach to the collisional Vlasov equation studied by EEA in El-Wakil *et al.* (2003). *We have shown that there exist exact maximum entropy solutions to the collisional Vlasov equation (as opposed to the solutions advanced by EEA in El-Wakil* et al. *(2003), that are only approximate).* We considered two different approaches to the exact maximum entropy solutions of the collisional Vlasov equation. On the one hand, we identified an appropriate set of five relevant mean values (moments) that evolve according to a closed set of coupled, ordinary, linear differential equations. We then constructed exact maximum entropy solutions associated with that set of moments. We investigated these solutions using both the equations of motion of the moments themselves, as well as the equations of motion of the corresponding Lagrange multipliers.

On the other hand, we proved that it is possible to obtain exact solutions of the reduced equation considered by EEA, if the zeroth-order moment of the solutions is explicitly taken into account. In order to do this, we proved that some of the partial differential equations considered by EEA in El-Wakil *et al.* (2003) (that is, equations of the form (2.49)) admit exact maximum entropy solutions. These authors considered *approximate* maximum entropy solutions based upon the optimization of the Boltzmann-Gibbs entropy under the constraints imposed by the first two moments, $\langle z \rangle$ and $\langle z^2 \rangle$ (they implicitly assumed a constant normalization to 1 of their maximum entropy ansatz). However, the *exact* solutions to an evolution equation of the form (2.49) have a time-dependent normalization. Consequently, the maximum entropy solutions advanced by EEA are bound to be only approximate, since they do not take into account this important piece of information concerning the exact solutions. In particular, and contrary to what is asserted in El-Wakil *et al.* (2003) (see paragraph after equation (32) in El-Wakil *et al.* (2003)), some of the approximate, maximum entropy solutions obtained by EEA are going to differ drastically from the exact solutions at large times ($s \to \infty$). The reason for this is that EEA's maximum entropy solutions (for instance, solution (31) to equation (26) in El-Wakil *et al.* (2003)) have a constant normalization equal to 1, while the exact solution has a time dependent normalization. For example, the exact solutions to equation (26) of El-Wakil *et al.* (2003) have, for $h_3 > 0$ ($h_3 < 0$) a monotonously increasing (decreasing) normalization. In the cases of equations (26) and (36) of El-Wakil *et al.* (2003), EEA provide (formal) exact solutions. But they do not recognize those solutions as maximum entropy solutions easily obtainable if the zeroth order moment is incorporated.

What we have shown here is that, if the time-dependent zeroth order moment is explicitly taken into account as a further constraint in the maximum entropy procedure, it is possible to obtain *exact* maximum entropy solutions to some of the evolution equations considered by EEA in El-Wakil *et al.* (2003).

Chapter 3

Maximum Entropy Principle and Classical Evolution Equations with Source Terms

3.1 Introduction

The application of information-entropic variational principles to the study of diverse systems and processes in physics, astronomy, biology, and other fields, has been the focus of considerable research activity in recent years. A (by no means exhaustive) list of important examples is given in references: Beck and Schlögl (1993); Boghosian (1996); Borland (1996); Frank (2005); Frieden (1998, 2004); Frieden and Soffer (1995); Gell-Mann and Tsallis (2004); Jaynes (2003); Plastino and Curado (2005); Sieniutycz and Farkas (2005); Yamano (2000). The roots of this approach can be traced back (at least) to Gibbs (1902) who pointed out that the canonical probability distribution is the one maximizing the entropy under the constraints imposed by normalization and the mean energy value. However, it was Jaynes who elevated the principle of maximum entropy to the status of a foundational starting point for the development of statistical mechanics, and the first to recognize its relevance as a general statistical inference principle (Grandy and Milonni (1993); Jaynes (1983); Katz (1967)) see also Chapter (1).

A large amount of research has been devoted to the study of time dependent maximum entropy solutions (either exact or approximate) of diverse evolution equations, such as the Liouville equation, the Vlasov equation, diffusion equations, and Fokker-Planck equations (Baker-Jarvis *et al.* (1989); Borland *et al.* (1999); da Silva *et al.* (2004); El-Wakil *et al.* (2003); Hick and Stevens (1987); Malaza *et al.* (1999); Plastino and Plastino (1997); Plastino *et al.* (1997a,b,c); Tsallis and Bukman (1996)). Most of these applications of the maximum entropy method to time dependent scenarios involved evolution equations (linear or nonlinear) exhibiting the form of a *continuity equation* and, consequently, preserving normalization in time. Our purpose here is to explore some aspects of the application of the maximum entropy approach to a special type of evolution equations: those endowed with source terms and, consequently, *not preserving normalization.*

It is a common assumption that entropic concepts, including the maximum entropy principle, can be applied only to *probability distributions.* A given function ρ, if it is to be interpreted as a probability distribution, has to be non-negative and normalized to unity. However, entropic concepts can be profitably applied also to the study of (positive) *densities*, which are non-negative quantities not necessarily normalized to 1. Indeed, a (positive) density can be normalized to any positive number $\mathcal{N}$. The application of the maximum entropy principle to the study of densities allows for the analysis of a variegated family of interesting problems. For example, densities may evolve according to non-linear evolution equations (Borland *et al.* (1999); da Silva *et al.* (2004); Tsallis and Bukman (1996)) (as contrasted to ensemble probabilities which, strictly speaking, must evolve linearly, see van Kampen (1992)). In this regard, it is worthwhile to remember that Boltzmann himself introduced his celebrated entropic functional in order to study the evolution of the density of particles in the $(\mathbf{x}, \mathbf{v})$ space which, by the way, obeys a non-linear transport equation. When applying the maximum entropy principle to the evolution of a density the normalization $\mathcal{N}$ may even change with time (i.e., $\mathcal{N} = \mathcal{N}(t)$). This is precisely the case with the (linear) evolution equations with source terms that we are going to consider in the present work.

There are several possible scenarios where these equations with sources may arise. For instance, when considering the diffusion of a certain type of particles we may need to include explicitly, in the description of the diffusion process, the sources of those particles. This situation may arise in several problems in physics, astronomy, or biology. For example, when dealing with the transport equation of cosmic rays (Hick and Stevens (1987)), if we want to include the sources of cosmic rays into our model, we have to incorporate the corresponding source-terms into the evolution equation. In spite of its possible practical applications, our principal interest in the present contribution will be to explore the structure of the dynamical equations connecting the (time dependent) main characters of our maximum entropy scheme: the relevant mean values (constituting, at an initial time t_0, the available prior information), the associated Lagrange multipliers, the partition function, and the entropy. In particular, we are going to investigate the relationships between H-theorems verified by the exact solutions and the H-theorems verified by the maximum entropy approximate ones.

This chapter is organized as follows. In Section II we explain, and provide some examples, of the type of evolution equations that we are going to consider in this work. Some properties of the exact time dependent solutions to this equations are derived in Section III. A maximum entropy formalism to treat these equations is implemented in Section IV, where some of its main features are investigated. In Section V some examples are considered, in order to illustrate the results obtained in the previous sections. Finally, some conclusions are drawn in Section VI.

3.2 Evolution Equations with Source Terms

In references (Baker-Jarvis *et al.* (1989); El-Wakil *et al.* (2003); Hick and Stevens (1987); Malaza *et al.* (1999); Plastino and Plastino (1997); Plastino *et al.* (1997a,b,c)) the maximum entropy principle has been used with reference to the study of *equations of evolution exhibiting the form of continuity equations*. We may mention, for instance, the Liouville equation, the Fokker-Planck equation, diffusion equations, the Von Neumann equation in quantum mechanics, etc. The evolution equations that we are going to investigate here comprise a continuity-like equation *plus an extra term K describing a source or a sink*. Let us consider a classical system described by a time dependent density distribution $F(\mathbf{z}, t)$ evolving according to the partial differential equation

$$\frac{\partial F}{\partial t} + \nabla \cdot \mathbf{J} = K, \tag{3.1}$$

where $\mathbf{z}$ denotes a point in the relevant N-dimensional phase space, $\mathbf{J}$ is the flux vector, and K represents a source-term ($\mathbf{J}$ and K may depend on the distribution F). As examples we have:

- The one dimensional diffusion equation with a source term,

$$\frac{\partial F}{\partial t} - Q\frac{\partial^2 F}{\partial x^2} = K, \tag{3.2}$$

where Q denotes the diffusion coefficient, and the flux is given by

$$J = -Q\frac{\partial F}{\partial x}. \tag{3.3}$$

- The general Liouville equation with a source term K

$$\frac{\partial F}{\partial t} + \nabla \cdot (F\,\mathbf{w}) = K, \tag{3.4}$$

with flux

$$\mathbf{J} = F\,\mathbf{w}. \tag{3.5}$$

If $K = 0$ we recover the standard (general) Liouville equation (Andrey (1985); Liouville (1838); van Kampen (1992)). The Liouville equation describes the evolution of an ensemble of classical, deterministic dynamical systems evolving according to the equations of motion

$$\frac{d\mathbf{z}}{dt} = \mathbf{w}(\mathbf{z}), \tag{3.6}$$

where $\mathbf{z}$ denotes a point in the concomitant N-dimensional phase space.

- Hamiltonian ensemble dynamics with sources, a particular instance of the Liouville equations (3.6). For Hamiltonian systems with n degrees of freedom we have

1. $N = 2n$,
2. $\mathbf{z} = (q_1, \ldots, q_n, p_1, \ldots, p_n)$,
3. $w_i = \partial H/\partial p_i, \ \ (i = 1, \ldots, n), \quad$ and
4. $w_{i+n} = -\partial H/\partial q_i, \ \ (i = 1, \ldots, n)$,

where the q_i and the p_i stand for generalized coordinates and momenta, respectively.

With reference to the last item note that Hamiltonian dynamics exhibits the important feature of being divergence-less

$$\nabla \cdot \mathbf{w} = 0. \tag{3.7}$$

For it the Liouville equation simplifies to

$$\frac{\partial F}{\partial t} + \mathbf{w} \cdot \nabla F = K, \tag{3.8}$$

which is equivalent to a relationship obeyed by the total time derivative

$$\frac{dF}{dt} = K, \tag{3.9}$$

that is computed along an individual phase-space's orbit.

As mentioned in section (3.1), there are several problems in physics, astronomy, and biology where the evolution equations with sources arise naturally. When studying diffusion problems we can include explicitly, in the description of the diffusion process, the sources of the diffusing particles. In that case, the most natural kind of source term $K(\mathbf{z}, t)$ is given by a positive function of $\mathbf{z}$ and t (if the source is time dependent) not depending on F itself. Another type of situation leading naturally to a source term is given by the diffusion of particles that undergo a certain decay process. In such a case, the changes in the evolving density $F(\mathbf{z}, t)$ have two different origins. On the one hand, the diffusion process itself, and on the other hand, the decay process. This last factor gives rise to a negative source-like term (that is, a sink-like term) proportional to $F(\mathbf{z}, t)$ itself,

$$K \ = \ -\, qF, \tag{3.10}$$

where the (positive) constant q is related to the mean life τ of the decaying particles.

3.3 Evolution of the Entropy and the Relevant Mean Values

In order to implement the maximum entropy method, we need to re-formulate our problem in terms of a density $f(\mathbf{z}, t)$ that is normalized to unity and therefore can be regarded as a probability density. Consequently, it will prove convenient to re-cast the density distribution $F(\mathbf{z}, t)$ under the guise,

$$F(\mathbf{z}, t) \;=\; \mathcal{N}(t)\, f(\mathbf{z}, t), \tag{3.11}$$

with

$$\int F(\mathbf{z}, t)\, d^N z \;=\; \mathcal{N}(t), \tag{3.12}$$

and

$$\int f(\mathbf{z}, t)\, d^N z \;=\; 1. \tag{3.13}$$

The evolution equations for $\mathcal{N}$ and f are, respectively,

$$\frac{d\mathcal{N}}{dt} \;=\; \int K\, d^N z, \tag{3.14}$$

and

$$\begin{aligned}
\frac{\partial F}{\partial t} + \nabla \cdot \mathbf{J} \;&=\; K \\
\frac{\partial \mathcal{N} f}{\partial t} + \nabla \cdot \frac{\mathbf{j}}{\mathcal{N}} \;&=\; \frac{k}{\mathcal{N}} \\
\frac{\partial \mathcal{N}}{\partial t} f + \frac{\partial f}{\partial t}\mathcal{N} + \nabla \cdot \frac{\mathbf{j}}{\mathcal{N}} \;&=\; \frac{k}{\mathcal{N}} \\
\frac{\partial f}{\partial t} + \nabla \cdot \mathbf{j} \;&=\; k - \frac{\dot{\mathcal{N}}}{\mathcal{N}} f,
\end{aligned} \tag{3.15}$$

where we have introduced the abbreviations

$$\mathbf{j} = \frac{\mathbf{J}}{\mathcal{N}}, \tag{3.16}$$

and

$$k = \frac{K}{\mathcal{N}}. \tag{3.17}$$

3.3.1 Evolution of the Entropy

Since the density f is properly normalized, we can consider its (time-dependent) Shannon entropy,

$$S[f] = -\int f \ln f \, d^N z, \tag{3.18}$$

whose time derivative is given by (cf. Eq.(3.15))

$$\begin{aligned} \frac{dS}{dt} &= \int \left[\frac{\dot{\mathcal{N}}}{\mathcal{N}} f + \nabla \cdot \mathbf{j} - k \right] \ln f \, d^N z \\ &= -\frac{\dot{\mathcal{N}}}{\mathcal{N}} S + \int \left[\nabla \cdot \mathbf{j} - k \right] \ln f \, d^N z. \end{aligned} \tag{3.19}$$

The following alternative (but equivalent) expression for the time derivative of the entropy is also useful,

$$\frac{dS}{dt} = -\frac{\dot{\mathcal{N}}}{\mathcal{N}} S - \int k \ln f \, d^N z + \left\langle \nabla \cdot \left(\frac{\mathbf{j}}{f} \right) \right\rangle. \tag{3.20}$$

If the source $k(\mathbf{z})$ has a definite sign we can introduce the function

$$g(\mathbf{z}, t) = \frac{\mathcal{N}}{\dot{\mathcal{N}}} k(\mathbf{z}), \tag{3.21}$$

which verifies,

$$\begin{aligned} g(\mathbf{z}, t) &\geq 0, \\ \int g(\mathbf{z}, t) \, d^N z &= 1, \end{aligned} \tag{3.22}$$

and can thus be interpreted as a probability density function associated with the source term. Now, adding and subtracting the integral

$$\int k \ \ln \left(\frac{\mathcal{N}k}{\dot{\mathcal{N}}} \right) \, d^N z, \tag{3.23}$$

from equation (3.20), we get

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\dot{\mathcal{N}}}{\mathcal{N}} S[f] - \int k \ \ln f \, d^N z + \int k \ \ln g \, d^N z - \int k \ \ln g \, d^N z + \left\langle \nabla \cdot \left(\frac{\mathbf{j}}{f} \right) \right\rangle \\ &= \frac{\dot{\mathcal{N}}}{\mathcal{N}} \left[-S[f] - \int g \, (\ln f - \ln g) \, d^N z - \int g \ \ln g \, d^N z \right] + \left\langle \nabla \cdot \left(\frac{\mathbf{j}}{f} \right) \right\rangle \\ &= \frac{\dot{\mathcal{N}}}{\mathcal{N}} \Big\{ S[g] \ - \ S[f] \ - \ I[g, f] \Big\} \ + \ \left\langle \nabla \cdot \left(\frac{\mathbf{j}}{f} \right) \right\rangle, \end{aligned} \tag{3.24}$$

where

$$I[g, f] \ = \ \int g \ln(f/g) \, d^N z, \tag{3.25}$$

denotes the Kullback distance (Kullback (1959)) between the probability densities g and f.

An interesting particular instance of equation (3.24) is obtained when we have a source term proportional to F itself,

$$K \ = \ q \, F, \tag{3.26}$$

with q constant. If $q < 0$ we can interpret this source term as describing the flow of particles that undergo a decay process. With a term like (3.26) we have $g = f$ and

$$\frac{dS}{dt} \ = \ \left\langle \nabla \cdot \left(\frac{\mathbf{j}}{f} \right) \right\rangle. \tag{3.27}$$

In the particular case of a Liouville equation with a source like (3.26) we get,

$$\frac{dS}{dt} \ = \ \langle \nabla \cdot \mathbf{w} \rangle, \tag{3.28}$$

which coincides with the expression for the time derivative of the entropy for the standard, norm preserving Liouville equation (Andrey (1985); Ruelle (2004)).

3.3.2 Evolution of the Relevant Mean Values

Another important ingredient of the maximum entropy approach is given by the set of mean values

$$\langle A_i \rangle \;=\; \int A_i \, F \, d^N z, \tag{3.29}$$

of M relevant quantities A_i, $(i = 1, \ldots, M)$. These M quantities are going to play the role of the prior information used to construct the maximum entropy ansatz. We are going to assume that these M mean values are known at an initial time t_0 (more on this later).

The time derivatives of the relevant mean values (3.29) are

$$\begin{aligned} \frac{d}{dt}\langle A_i \rangle \;&=\; \int A_i \frac{d}{dt} F \, d^N z, \\ &=\; \int \Big[-A_i \, \nabla \cdot \mathbf{J} + A_i \, K \Big] \, d^N z, \quad (i = 1, \ldots M), \end{aligned} \tag{3.30}$$

Integrating by parts and making the usual assumption that $\mathbf{J} \to 0$ rapidly enough as $|z| \to \infty$, surface terms vanish (as they do in most physics problems) and we finally obtain

$$\frac{d}{dt}\langle A_i \rangle \;=\; \int \Big[\mathbf{J} \cdot \nabla A_i + \, A_i \, K \Big] \, d^N z, \;\; (i = 1, \ldots, M). \tag{3.31}$$

We are also going to need, and thus introduce now, the "re-scaled" mean values,

$$a_i \;=\; \frac{1}{\mathcal{N}} \langle A_i \rangle. \tag{3.32}$$

3.4 Maximum Entropy Ansatz for the Evolution Equation

3.4.1 Preliminaries

A central point for our present discussion is that of considering a specially important ansatz for solving the evolution equation (3.1), namely, the maximum entropy one,

$$F(\mathbf{z}, t) \;=\; \mathcal{N} f_{ME}(\mathbf{z}, t) \;=\; \frac{\mathcal{N}}{Z} \exp\left[-\sum_{i=1}^{M} \lambda_i A_i \right], \tag{3.33}$$

where the $A_i(\mathbf{z})$ are M appropriate quantities that are functions of the phase space location $\mathbf{z}$. The partition function Z is given by,

$$Z \;=\; \int \exp\left[-\sum_{i=1}^{M} \lambda_i A_i \right] \, d^N z. \tag{3.34}$$

The probability distribution f_{ME} appearing in (3.33) is the one that maximizes the entropy $S[f]$ under the constraints imposed by normalization and the relevant mean values $\langle A_i \rangle$ (or the $a_i = \langle A_i \rangle / \mathcal{N}$). The re-scaled relevant mean values a_i and the associated Lagrange multipliers λ_i are related by the celebrated Jaynes' relations (Katz (1967) see also equations (1.16), (1.17), (1.15) and (1.18))

$$\lambda_i \;=\; \frac{\partial S}{\partial a_i}, \tag{3.35}$$

$$a_i = \frac{\langle A_i \rangle}{\mathcal{N}} \;=\; -\frac{\partial}{\partial \lambda_i} (\ln Z), \tag{3.36}$$

$$S \;=\; \ln Z \,+\, \sum_i \lambda_i a_i, \tag{3.37}$$

and

$$\frac{\partial \lambda_i}{\partial a_j} = \frac{\partial^2 S}{\partial a_i \partial a_j} = \frac{\partial \lambda_j}{\partial a_i}. \tag{3.38}$$

As already mentioned all the basic equations of equilibrium thermodynamics are particular instances of (3.35-3.38), or can be derived from special instances of (3.35-3.38). *This fact alone provides already a strong motivation* for studying in detail the interplay between the various quantities appearing in Jaynes' relations, when applying the maximum entropy principle to diverse physical scenarios. Indeed, a special instance of this line of enquiry constitutes one of our main focus of attention here.

All the time dependence of the maximum entropy distribution f_{ME} appearing in the ansatz (3.33) is contained in the Lagrange multipliers $\lambda_i(t)$, which are assumed to be time dependent. The Lagrange multipliers (and the normalization factor $\mathcal{N}$) change in time in order to accommodate the evolving mean values $\langle A_i \rangle$ (and the evolving norm of $F(\mathbf{z}, t)$). We assume that the mean values of the M relevant quantities A_i at an initial time t_0,

$$\Big\{ \langle A_1 \rangle_{t_0}, \ldots, \langle A_M \rangle_{t_0} \Big\}, \tag{3.39}$$

as well as the initial value $\mathcal{N}_{t_0}$, are known. They constitute our prior information. On the basis of these initial data we determine the initial values of the Lagrange multipliers λ_i and the partition function Z. Then, on the basis of an appropriate set of equations of motion for the relevant mean values (constructed using the evolving maximum entropy ansatz) we determine the (approximate) time evolution of the $\langle A_i \rangle$. Now, in general, the time derivatives of the aforementioned mean values are given by equation (3.31), that is re-written here for convenience,

$$\frac{d}{dt} \langle A_i \rangle = \int [\mathbf{J} \cdot \nabla A_i + A_i K] \, d^N z, \quad (i = 1, \ldots, M). \tag{3.40}$$

The integrals appearing in the right hand sides of these equations generally involve, unfortunately, new mean values not included in the original set $\langle A_i \rangle$ $(i = 1, \ldots, M)$ (remember that the flux $\mathbf{J}$ depends on the distribution f). One way to implement the maximum entropy approach to solve the evolution equation (3.1) is to evaluate, at each instant of time, the right hand sides of (3.40) using the maximum entropy ansatz (3.33). In this way, the system of equations (3.40) can be translated into a closed system of equations of motion for the Lagrange multipliers λ_i. This (time dependent self-consistent) approach will yield either exact solutions, or only approximate solutions, depending on the specific form of the evolution equation (3.1) (such is also the case, of course, for continuity equations. See Malaza *et al.* (1999); Plastino and Plastino (1997); Plastino *et al.* (1997a,b,c) and references therein).

3.4.2 Time Evolution

We discuss now specific details of the temporal evolution, beginning with that of the Lagrange multipliers. Regarding the set of quantities a_i, $(i = 1, \ldots, M)$ as the set of independent parameters characterizing f_{ME}, we get

$$
\begin{aligned}
\frac{d\lambda_i}{dt} &= \sum_{j=1}^{M} \left(\frac{\partial \lambda_i}{\partial a_j}\right)\left(\frac{da_i}{dt}\right) \\
&= \sum_{j=1}^{M} \left(\frac{\partial \lambda_j}{\partial a_i}\right)\left(\frac{da_i}{dt}\right) \\
&= \frac{\partial}{\partial a_i}\left(\sum_{j=1}^{M} \lambda_j \left(\frac{da_j}{dt}\right)\right) - \sum_{j=1}^{M} \lambda_j \frac{\partial}{\partial a_i}\left(\frac{da_j}{dt}\right).
\end{aligned}
\tag{3.41}
$$

Now, since $\langle A_i \rangle = \mathcal{N} a_i$ (equation (3.32)), we have

$$
\begin{aligned}
\frac{da_i}{dt} &= \frac{1}{\mathcal{N}}\frac{d\langle A_i\rangle}{dt} - \frac{\langle A_i\rangle}{\mathcal{N}^2}\frac{d\mathcal{N}}{dt} \\
&= \frac{1}{\mathcal{N}}\int \left[\mathbf{J}\cdot\nabla A_i + A_i K - \frac{\dot{\mathcal{N}}}{\mathcal{N}} F A_i\right] d^N z \\
&= \frac{1}{\mathcal{N}}\int \left[\mathbf{J}\cdot\nabla A_i + A_i K - \dot{\mathcal{N}} f A_i\right] d^N z,
\end{aligned}
\tag{3.42}
$$

and, as a consequence,

$$
\sum_{i=1}^{M} \lambda_i \left(\frac{da_i}{dt}\right) = \frac{1}{\mathcal{N}}\int \left(\mathbf{J}\cdot\nabla \sum_i \lambda_i A_i + K \sum_i \lambda_i A_i - \dot{\mathcal{N}} f \sum_i \lambda_i A_i\right) d^N z.
\tag{3.43}
$$

Substituting now the maximum entropy ansatz (3.33) for f (remember that we have defined $\mathbf{j} = \mathbf{J}/\mathcal{N}$) one gets

$$
\begin{aligned}
\sum_{i=1}^{M} \lambda_i \left(\frac{da_i}{dt}\right) &= \frac{1}{\mathcal{N}}\left[-\int f\,(\mathcal{N}\mathbf{j})\cdot\nabla[\ln(fZ)]\,d^N z + \int [\dot{\mathcal{N}} f - K]\,[\ln(fZ)]\,d^N z\right] \\
&= \frac{1}{\mathcal{N}}\left[\int F\,\nabla\cdot(\mathbf{j}/f)\,d^N z + \int [\dot{\mathcal{N}} f - K]\,[\ln(fZ)]\,d^N z\right] \\
&= \frac{1}{\mathcal{N}}\left(\langle\nabla\cdot(\mathbf{j}/f)\rangle + \int [\dot{\mathcal{N}} f - K]\,\ln f\,d^N z\right),
\end{aligned}
\tag{3.44}
$$

where the fact has been used that (3.14) implies $\int [\dot{\mathcal{N}} f - K]\,[\ln(Z)]\,d^N z = 0$. Finally,

$$\begin{aligned}\frac{d\lambda_i}{dt} &= \frac{\partial}{\partial a_i}\int \left[f\,\nabla\cdot\left(\frac{\mathbf{j}}{f}\right) + \left(\frac{\dot{\mathcal{N}}}{\mathcal{N}}f - k\right)\ln f\right] d^N z \\ &\quad -\sum_{j=1}^{M}\lambda_j \frac{\partial}{\partial a_i}\int\left[\mathbf{j}\cdot\nabla A_j + A_j\,k - \frac{\dot{\mathcal{N}}}{\mathcal{N}}f A_j\right] d^N z \end{aligned} \tag{3.45}$$

3.4.3 Evolution of the Entropy

Now we are going to consider the time derivative of the entropy evaluated on the maximum entropy solution: $S[f_{ME}]$. From equations (3.36) and (3.37) we have,

$$\begin{aligned}\frac{d}{dt}S[f_{ME}] &= \frac{d}{dt}(\ln Z) + \frac{d}{dt}\left(\sum_{i=1}^{M}\lambda_i a_i\right) \\ &= \sum_i \frac{d\lambda_i}{dt}\frac{\partial}{\partial\lambda_i}(\ln Z) + \sum_i \frac{d\lambda_i}{dt}a_i + \sum_i \lambda_i \frac{da_i}{dt} \\ &= \sum_i \lambda_i \frac{da_i}{dt}, \end{aligned} \tag{3.46}$$

and, now using equation (3.43), we find the important relation

$$\begin{aligned}\frac{d}{dt}S[f_{ME}] &= \frac{1}{\mathcal{N}}\int\left[\mathbf{J}\cdot\nabla\sum_i \lambda_i A_i + K\sum_i \lambda_i A_i - \dot{\mathcal{N}} f_{ME}\sum_i \lambda_i A_i\right] d^N z \\ &= \int\left[-(\nabla\cdot\mathbf{j})\left(\sum_i \lambda_i A_i\right) + k\sum_i \lambda_i A_i - \frac{\dot{\mathcal{N}}}{\mathcal{N}} f_{ME}\sum_i \lambda_i A_i\right] d^N z \\ &= \int\left[(\nabla\cdot\mathbf{j})\,(\ln Z + \ln f_{ME}) + \left(\frac{\dot{\mathcal{N}}}{\mathcal{N}} f_{ME} - k\right)(\ln Z + \ln f_{ME})\right] d^N z \\ &= \int\left(\nabla\cdot\mathbf{j} + \frac{\dot{\mathcal{N}}}{\mathcal{N}} f_{ME} - k\right)\ln f_{ME}\, d^N z \\ &= -\frac{\dot{\mathcal{N}}}{\mathcal{N}} S[f_{ME}] + \int(\nabla\cdot\mathbf{j} - k)\ln f_{ME}\, d^N z. \end{aligned} \tag{3.47}$$

Comparing now the expression for the entropy's time derivative corresponding to the exact solutions (cf. equation(3.19)) with the expression just derived (3.47) for the maximum entropy ansatz, we can reach an important conclusion:

our present maximum entropy scheme always (even in the case of approximate solutions) preserves the exact functional relationship between the time derivative of the entropy and the time dependent solutions of the evolution equation. Consequently, any H-theorem verified when evaluating the entropy functional upon the exact solutions is also verified when evaluating the entropy upon the maximum entropy approximate treatments. This is of considerable relevance in connection with the consistency of the method as a maximum entropy approach.

3.5 Examples

3.5.1 Liouville Equation with Constant Sources

According to equation (3.31), and remembering that, for the Liouville equation, the flux is given by $\mathbf{J} = F\mathbf{w}$, the temporal evolution of the mean values of the dynamical quantities A_i is

$$\begin{aligned} \frac{d\langle A_i\rangle}{dt} &= \int \Big[F\,\mathbf{w}\cdot\nabla A_i \,+\, A_i K \Big]\, d^N z \\ &= \langle \mathbf{w}\cdot\nabla A_i\rangle \,+\, B_i, \quad (i=1,\ldots,M)\,, \end{aligned} \tag{3.48}$$

where

$$B_i \,=\, \int A_i K\, d^N z, \quad (i=1,\ldots,M)\,. \tag{3.49}$$

Here we are going to assume that f is given by the ansatz (3.33)-(3.34). We can then regard the quantities Z, f, and λ_i's as functions of the set $a_1,\ldots,a_M$. Alternatively, it is also possible to regard all relevant quantities as functions of the λ_i's instead.

Let us consider the important particular case where the following closure relationship holds,

$$\mathbf{w}\cdot\nabla A_i \,=\, \sum_j^M C_{ij}\, A_j, \;\; (i=1,\ldots,M), \tag{3.50}$$

where the C_{ij} constitute a set of (structure) constants. This entails that

$$\frac{d\langle A_i\rangle}{dt} = \sum_j^M C_{ij}\,\langle A_j\rangle \;+\; B_i, \;\; (i = 1, \ldots, M). \tag{3.51}$$

It is useful also to introduce the quantity,

$$B_0 \;=\; \int K\, d^N z. \tag{3.52}$$

The general solution of the equations of motion for the mean values is then seen to be of the form

$$\langle A_i\rangle(t) \;=\; \langle A_i\rangle_{\text{inhom.}} \;+\; \langle A_i\rangle_{\text{hom.}}, \tag{3.53}$$

where $\langle A_j\rangle_{\text{inhom.}}$ complies with

$$\sum_{j=1}^N C_{ij}\langle A_j\rangle_{\text{inhom.}} \;+\; B_i \;=\; 0, \tag{3.54}$$

and is a particular solution of the (inhomogeneous) set of linear differential equations, while $\langle A_i\rangle_{\text{hom.}}$ is the general solution of the homogeneous set of equations

$$\frac{d\langle A_i\rangle}{dt} = \sum_j^M C_{ij}\,\langle A_j\rangle \;\; (i = 1, \ldots, M). \tag{3.55}$$

Now, if $\nabla\cdot\mathbf{w} = 0$ (that is, if the flux $\mathbf{w}$ is divergenceless) the temporal evolution of the Lagrange multiplier is given by,

$$\begin{aligned}
\frac{d\lambda_i}{dt} \;&=\; \frac{\partial}{\partial a_i}\int\left[f\,\nabla\cdot\left(\frac{\mathbf{j}}{f}\right) \;+\; \left(\frac{\dot{\mathcal{N}}}{\mathcal{N}}\,f - k\right)\,\ln f\right]\,d^N z \\
&\quad - \sum_{j=1}^M \lambda_j\frac{\partial}{\partial a_i}\int\left[\mathbf{j}\cdot\nabla A_j \;+\; A_j\,k \;-\; \frac{\dot{\mathcal{N}}}{\mathcal{N}}\,f A_j\right]\,d^N z \\
&=\; \frac{\partial}{\partial a_i}\int\left[f\,\nabla\cdot\mathbf{w} \;+\; \frac{1}{\mathcal{N}}\left(\dot{\mathcal{N}}f \;-\; K\right)\,\ln f\right]\,d^N z \\
&\quad - \sum_{j=1}^M \lambda_j\frac{\partial}{\partial a_i}\int\left[f\mathbf{w}\cdot\nabla A_j \;+\frac{1}{\mathcal{N}}(\,A_j\,K \;-\; \dot{\mathcal{N}}\,f A_j)\right]\,d^N z \\
&=\; -\,\frac{\dot{\mathcal{N}}}{\mathcal{N}}\frac{\partial S}{\partial a_i} - \frac{1}{\mathcal{N}}\frac{\partial}{\partial a_i}\int K\,\ln f\, d^N z
\end{aligned}$$

$$
\begin{aligned}
& - \sum_{j=1}^{M} \lambda_j \frac{\partial}{\partial a_i} \int \left[f \left(\sum_{k}^{M} C_{jk}\, A_k \right) + \frac{1}{\mathcal{N}} (A_j\, K - \dot{\mathcal{N}}\, f\, A_j) \right] d^N z \\
= & - \frac{\dot{\mathcal{N}}}{\mathcal{N}} \lambda_i - \frac{1}{\mathcal{N}} \frac{\partial}{\partial a_i} \int K \,\ln f \, d^N z \\
& - \sum_{j=1}^{M} \lambda_j \frac{\partial}{\partial a_i} \left[\left(\sum_{k}^{M} C_{jk}\, a_k \right) + \frac{1}{\mathcal{N}} \left(\int A_j\, K \, d^N z \right) - \frac{\dot{\mathcal{N}}}{\mathcal{N}} a_j \right] \\
= & - \frac{\dot{\mathcal{N}}}{\mathcal{N}} \lambda_i - \frac{1}{\mathcal{N}} \frac{\partial}{\partial a_i} \int K \,\ln f \, d^N z \\
& - \sum_{j=1}^{M} \lambda_j \left[C_{ji} + \frac{1}{\mathcal{N}} \frac{\partial}{\partial a_i} \left(\int A_j\, K \, d^N z \right) - \frac{\dot{\mathcal{N}}}{\mathcal{N}} \delta_{ij} \right] \\
= & - \sum_{j=1}^{M} \lambda_j \left[C_{ji} + \frac{1}{\mathcal{N}} \frac{\partial}{\partial a_i} \left(\int A_j\, K \, d^N z \right) \right] \\
& - \frac{1}{\mathcal{N}} \frac{\partial}{\partial a_i} \int K \,\ln f \, d^N z.
\end{aligned} \tag{3.56}
$$

In the particular case where the source term $K(\bar{z}, t)$ does not depend explicitly on the distribution F this equation reduces to

$$
\frac{d\lambda_i}{dt} = - \left(\sum_{j=1}^{M} C_{ji} \lambda_j \right) - \frac{1}{\mathcal{N}} \frac{\partial}{\partial a_i} \int K \,\ln f \, d^N z. \tag{3.57}
$$

3.5.2 A Collisional Vlasov Equation with Sources

We are going to consider the following collisional Vlasov equation with sources,

$$
\frac{\partial F}{\partial t} + v \frac{\partial F}{\partial x} - \left[\frac{\partial \phi}{\partial x} + \gamma v \right] \frac{\partial F}{\partial v} - \gamma \alpha \frac{\partial^2 F}{\partial v^2} - \gamma F = \left[\beta_0 + \beta_1 x^2 \right] F, \tag{3.58}
$$

where γ, α, β_0, and β_1 are constants (γ and α are positive) and the potential ϕ is of a quadratic form,

$$
\phi(x) = \frac{1}{2} \phi_2 x^2. \tag{3.59}
$$

Here we are also going to assume that $\phi_2 > 0$. Equation (3.58) is a generalization of the source-free equation studied by El-Wakil *et al.* (2003). Let us now consider a maximum entropy ansatz of the form

$$F(x,v,t) = \exp[-\lambda_0 - \lambda_1 x - \lambda_2 v - \lambda_3 x^2 - \lambda_4 xv - \lambda_5 v^2], \\ = \frac{\mathcal{N}}{Z}\exp[-\lambda_1 x - \lambda_2 v - \lambda_3 x^2 - \lambda_4 xv - \lambda_5 v^2] \\ = \mathcal{N} f, \tag{3.60}$$

where the λ_i, $\quad i = 0, \ldots, 5$ are appropriate Lagrange multipliers and

$$Z = \int \exp[-\lambda_1 x - \lambda_2 v - \lambda_3 x^2 - \lambda_4 xv - \lambda_5 v^2]\, dxdv. \tag{3.61}$$

The (normalized) distribution f appearing in (3.60) maximizes the Boltzmann-Gibbs entropic functional,

$$S[f] = -\int f(x,v,t)\, \ln f(x,v,t)\, dxdv, \tag{3.62}$$

under the constraints imposed by normalization and the instantaneous mean values of the quantities $B_1 = x$, $B_2 = v$, $B_3 = x^2$, $B_4 = xv$, and $B_5 = v^2$. All the time dependence of the ansatz (3.60) is expressed through the Lagrange multipliers λ_i, which are time dependent. Inserting the ansatz (3.60) into the partial differential equation (3.58), and equating to zero, separately, terms proportional to $x^i v^j$ with different exponents i, j, it is possible to prove that the ansatz (3.60) constitutes an exact solution to (3.58), provided that the Lagrange multipliers comply with the set of coupled ordinary differential equations,

$$\frac{d\lambda_0}{dt} = -\gamma\alpha\lambda_2^2 + 2\gamma\alpha\lambda_5 - \gamma - \beta_0, \tag{3.63}$$

$$\frac{d\lambda_1}{dt} = \phi_2\lambda_2 - 2\gamma\alpha\lambda_4\lambda_2, \tag{3.64}$$

$$\frac{d\lambda_2}{dt} = -\lambda_1 + \gamma\lambda_2 - 4\gamma\alpha\lambda_2\lambda_5, \tag{3.65}$$

$$\frac{d\lambda_3}{dt} = \phi_2\lambda_4 - \gamma\alpha\lambda_4^2 - \beta_1, \tag{3.66}$$

$$\frac{d\lambda_4}{dt} = -2\lambda_3 + 2\phi_2\lambda_5 + \gamma\lambda_4 - 4\gamma\alpha\lambda_4\lambda_5, \tag{3.67}$$

and

$$\frac{d\lambda_5}{dt} = -\lambda_4 - 4\gamma\alpha\lambda_5^2 + 2\gamma\lambda_5. \tag{3.68}$$

Alternatively, we can focus our attention on the set of ordinary differential equations governing the evolution of the selected set of relevant mean values,

$$\frac{d}{dt}\langle x\rangle = \langle v\rangle + \beta_0\langle x\rangle + \beta_1\langle x^3\rangle, \tag{3.69}$$

$$\frac{d}{dt}\langle v\rangle = -\phi_2\langle x\rangle - \gamma\langle v\rangle + \beta_0\langle v\rangle + \beta_1\langle x^2 v\rangle, \tag{3.70}$$

$$\frac{d}{dt}\langle x^2\rangle = 2\langle xv\rangle + \beta_0\langle x^2\rangle + \beta_1\langle x^4\rangle, \tag{3.71}$$

$$\frac{d}{dt}\langle xv\rangle = -\phi_2\langle x^2\rangle - \gamma\langle xv\rangle + \langle v^2\rangle + \beta_0\langle xv\rangle + \beta_1\langle x^3 v\rangle, \tag{3.72}$$

and

$$\frac{d}{dt}\langle v^2\rangle = -2\phi_2\langle xv\rangle - 2\gamma\langle v^2\rangle + 2\alpha\gamma + \beta_0\langle v^2\rangle + \beta_1\langle x^2 v^2\rangle. \tag{3.73}$$

This example exhibits the peculiarity that, in spite of the fact that the maximum entropy ansatz (3.60) provides exact time dependent solutions to the equation (3.58), the equations of motion (3.69-3.73) for the five relevant mean values do not constitute a closed set of differential equations of motion for these quantities.

3.6 Conclusions

A maximum entropy approach to construct approximate, time dependent solutions to evolution equations endowed with source terms was considered. We have shown that in some particular cases the method leads to exact time dependent solutions. By construction our present implementation of the maximum entropy prescription complies with the exact equations of motion of the relevant mean values. Moreover, it always (even in the case of approximate solutions) preserves the exact functional relationship between the time derivative of the entropy and the time dependent solutions of the evolution equation. This means that any

H-theorem verified when evaluating the entropy functional upon the exact solutions is also verified when evaluating the entropy upon the maximum entropy approximate treatments. This is of considerable relevance in connection with the consistency of the method as a maximum entropy approach. Other features exhibited by the maximum entropy solutions and some illustrative examples were also discussed.

Chapter 4

Maximum Entropy Principle, Evolution Equations and Physics Education

4.1 Introduction

There is some repetition in this chapter of things already covered in previous chapters but it is deemed necessary for clarity and in order for this chapter to be self-contained.

The contents and structure of the physics curriculum have been in continuous evolution since the last quarter of the 19th century, when physics finally acquired, as a consolidated independent discipline and as a professional career, a form that would be (at least barely) recognizable by a physics student today. However, the pace of change of the physics curriculum has not been uniform. The first half of last century witnessed deep and rapid changes arising from the relativity and the quantum revolutions. On the other hand, during the second half of the 20th century the changes made on the physics curriculum have not been that dramatic. This (relatively speaking) "stationary state" had the psychological consequence that some physicists seem to believe that we have already reached "the end of History", as far as the physics curriculum is concerned. Far from the truth. Physics is nowadays experiencing profound changes both in terms of the contents of physics as a discipline, and in terms of the activities developed by professional physicists involved either in pure research or in the practical applications of the physical science. Two of the main sources behind these deep changes are (i) the fundamental new role played by the concept of information in some of the currently most active branches of theoretical physics and (ii) the increasing importance of the multidisciplinary areas of research (particularly concerning the application of methods and ideas from physics to biology, economics, sociology, etc.).

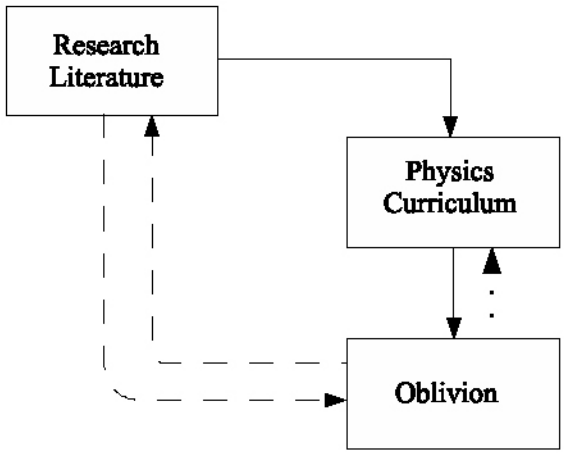

Figure 4.1: The flow of physical knowledge

Of course, the physics curriculum must have a finite length. Consequently, it is not possible to incorporate new contents to the curriculum without doing at the same time an appropriate re-organization of the traditional contents. The way to do this is to focus on the teaching of the general, unifying principles, concepts, methods and techniques. Consequently, there should be a flow (see figure (4.1)) of these "grand themes", originating in the physics research literature, to be integrated into the physics curriculum. Conversely, one should also expect some of the old, more specific contents to move away from the physics curriculum into what we may call "oblivion". This "flow" out of the physics curriculum has been taking place all the time (just compare a general physics textbook written before 1940 with one written at the end of the 20th century). There is also a continuous flow of themes out of the current research literature into "oblivion" (dashed lines in figure (4.1)). But to fall into oblivion from the research literature is less dramatic than to fall from the physics curriculum. Research interests and fashions change all the time, and a subject that fell into "oblivion" may come back at any time. But if something was once part of the physics curriculum, it means that there was once a consensus that it was among the most fundamental topics in physics. And when something falls from the curriculum, it almost never comes back.

The maximum entropy principle constitutes one of the alluded general, unifying, ideas that plays an important role in current research. It is, undoubtedly, one of the most fundamental tools in statistical physics, both from the conceptual and the practical points of view. It was first mentioned by Gibbs himself in his famous book on statistical mechanics (Gibbs (1902)). In that book Gibbs noticed that his canonical distribution is the one that maximizes the entropy under the constraints imposed by the mean energy and normalization. However, it was Jaynes who, inspired by ideas from information theory, elevated the maximum entropy principle to the status of the basic postulate of statistical mechanics (Jaynes (1983)). There are already several textbooks on equilibrium statistical mechanics that develop this subject taking as its basis the maximum entropy principle (Baierlein (1971); Katz (1967); Tribus (1961); Wyllie (1970)). However, the scientific relevance of the maximum entropy principle (and the information theoretical ideas behind it) goes well beyond the study of equilibrium statistical mechanics. The large number of applications of the maximum entropy principle to diverse areas of science attest to this. One of the first places in which this was explored is the classical work by Brillouin (1962). It is impossible to review here all the applications of the maximum entropy principle. To give an idea of the richness of its scope we mention now some recent applications.

- In Agrawal *et al.* (2005) the principle of maximum entropy yields a conditional probability distribution model for estimating the run-off for the catchment (watershed) of the Matatila dam in India. The model predicts run-off, subject to the selected constraints, in response to a given rainfall, in a rather adequate fashion.

- In Lukacs and Papp (2004) a maximum entropy method is applied directly to experimental kinetic absorption data in order to select between possible photocycle kinetics. No assumption is needed for the number of intermediate states taking part in the photocycle.

- In Blokhin *et al.* (2004), based on the maximum entropy principle, the authors proved the asymptotic stability of the equilibrium state for the balance-equations of charge transport in semiconductors, in the non-linear approximation, for a typical one dimensional problem.

- In Gong *et al.* (2004) a maximum entropy model-based framework is developed to provide a platform capable of integrating multimedia features as well as their contextual information in a uniform fashion to automatically detect and classify baseball highlights. This model simplifies the training-data creation and the highlight-detection and classification tasks.

- In Shams *et al.* (2004) the authors found that for a particular choice of the set of parameters related to the strengths of the (i) mean field, (ii) anti-alignment, (iii) internal magnetic field, and (iv) hopping, a system could exhibit physical properties characteristic of the colossal magnetoresistance. This property has been investigated within the framework of the maximum entropy principle for a system described by a simplified version of the Hubbard-Anderson Hamiltonian.

- In Amemiya *et al.* (2003), making use of the maximum entropy method, it is possible to determine the resonant frequency of a mechanical oscillator from the stochastic time-series data.

- In Israel *et al.* (2003) highly resolved electron density maps for LiF and NaF have been elucidated using reported X-ray structure factors. Here, the bonding electron density distribution is clearly revealed, both qualitatively and quantitatively, using the maximum entropy method.

- In Kim and Lee (2002) the maximum entropy method is introduced in order to build a robust formulation of the inverse problem. This method finds the solution which maximizes the entropy functional under the given temperature measurements.

- In Clowser and Strouthos (2002) the maximum entropy method is applied to dynamical fermion simulations of a Nambu-Jona-Lasinio model. The authors present results on large lattices for the spectral functions of the elementary fermion, the pion, the sigma, the massive pseudo-scalar meson, and the symmetric phase resonances.

- In Elgarayhi (2002) the method of maximum entropy is used for the solution of the aerosol dynamic equation so as to get physical insights into the role of coagulation, condensation, and removal processes.

- In Raychaudhuri *et al.* (2002) the possibility that statistical, natural-language processing techniques could be used to assign Gene-Ontology codes is explored. It is shown that maximum entropy modelling outperforms other methods for associating a set of GO codes (for biological processes) to literature-abstracts and thus to the genes associated with the abstracts.

- In El-Wakil *et al.* (2001) the maximum entropy approach is used to find the exact solution of the one-dimensional Fokker-Planck equation with variable coefficients. They consider three examples: the well-known Ornstein-Uhlenbeck differential equation, the Lamm equation and the Fokker-Planck equation for the linear Brownian motion.

The aim of this chapter's work is to provide some hints on how the maximum entropy principle can be incorporated into the teaching of those aspects of theoretical physics related to, but not restricted to, statistical mechanics. We are going to focus our attention on the study of maximum entropy solutions to evolution equations that exhibit the form of continuity equations. Such equations include, for instance, the Liouville equation, the diffusion equation, the Fokker-Planck equation, etc.

4.2 Brief Review of the Maximum Entropy Ideas

The second law of thermodynamics (Callen (1960); Desloge (1968)) is one of physics' most important statements. Together with the first law, they constitute strong pillars of our understanding of Nature. In statistical mechanics an underlying microscopic substratum is added that is able to explain not only these laws but the whole of thermodynamics itself (Katz (1967); Pathria (1993); Reif (1965); Sakurai (1985)). The most basic ingredient of such an explanation is a microscopic probability distribution that controls the population of microstates of the system under consideration (Pathria (1993)). Primarily, the maximum entropy approach, is an algorithm designed to obtain this probability distribution. In order to make sense of it, however, we must consider the concept of entropy in a more general information theoretic sense (Jaynes (1983); Katz (1967); Scalapino (1993), see also section 1.1).

4.2.1 A Derivation of Thermodynamics' First Law from the Maximum Entropy Principle

As a physical example of a maximum entropy application, let us tackle deriving the first law of thermodynamics from it in a special case: that in which we are concerned *only with changes that affect exclusively the microstate-population.* Thus, one considers a system whose possible atomic energy-levels are labelled by a set of quantum numbers collectively denoted by i that can be occupied with probabilities p_i. The way in which the variations dp_i are related to changes in a system's extensive quantities can be interpreted as one of the essential aspects of the first law (Reif (1965)). Consequently, one has to show that for any system described by a microscopic probability distribution $\{p_i\}$ with

- a concave entropic form (or information measure) S,
- a mean internal energy U,
- mean values $A_\nu \equiv \langle \mathcal{A}_\nu \rangle$, $(\nu = 1, \ldots, M)$ of M extensive quantities $\mathcal{A}_\nu$,
- a temperature T, and
- assuming a reversible process via $p_i \rightarrow p_i + dp_i$,

(**Thesis**): If a normalized probability distribution $\{p_i\}$ maximizes S, with the numerical values of U and the M A_ν as constraints, it entails that

$$dU = TdS - \sum_{\nu=1}^{M} \gamma_\nu \, dA_\nu$$
$$\text{First Law of Thermodynamics.} \tag{4.1}$$

.

4.2.1.1 Proof

Consider a quite *general* information measure (Plastino and Curado (2005); Plastino and Plastino (1997)) of the form

$$S = k \sum_i p_i \, f(p_i), \tag{4.2}$$

where, for simplicity's sake, Boltzmann's constant k_B is denoted here just by k. The sum runs over a set of quantum numbers, collectively denoted by i (characterizing levels of energy ϵ_i), that specify an appropriate basis in Hilbert space and $\mathcal{P} = \{p_i\}$ is an (as yet unknown) normalized probability distribution such that

$$\sum_i p_i = 1. \tag{4.3}$$

Let f be an arbitrary smooth function of the p_i. Further, consider M quantities A_ν that represent mean values of the extensive physical quantities $\mathcal{A}_\nu$. These take, for the state i, the value a_i^ν with probability p_i.

The mean energy U and the A_ν are given by

$$\begin{aligned} U &= \sum_i \epsilon_i \, p_i, \\ A_\nu &= \sum_i a_i^\nu \, p_i. \end{aligned} \tag{4.4}$$

Assume now that the set $\mathcal{P}$ changes in the fashion

$$p_i \to p_i + dp_i, \tag{4.5}$$

with $\sum_i dp_i = 0$ (cf. equation (4.3)), which in turn generates corresponding changes dS, dA_ν and dU in, respectively, S, the A_ν, and U. We wish to extremise S subject to the constraint of fixed i) U and ii) the M values A_ν. This is achieved via Lagrange multipliers i) β and ii) γ_ν ($\nu = 1, \ldots, M$). We need also a normalization Lagrange multiplier ξ.

$$\delta_{\{p_i\}} \left[S - \beta U - \sum_{\nu=1}^{M} \gamma_\nu A_\nu - \xi \sum_i p_i \right] = 0, \tag{4.6}$$

leading, with $\gamma_\nu = \beta \lambda_\nu$, to

$$0 = \delta_{p_m} \sum_i p_i f(p_i) - \delta_{p_m} \left[\sum_i \beta p_i \left(\sum_{\nu=1}^{M} \lambda_\nu \, a_i^\nu + \epsilon_i \right) - \xi \right], \tag{4.7}$$

so that

$$0 = f(p_i) + p_i f'(p_i) - \beta \left(\sum_{\nu=1}^{M} \lambda_\nu \, a_i^\nu + \epsilon_i \right) - \xi,$$

that after setting $\xi = \beta K$ becomes

$$0 = f(p_i) + p_i f'(p_i) - \beta \left[\left(\sum_{\nu=1}^{M} \lambda_\nu \, a_i^\nu + \epsilon_i \right) - K \right]. \tag{4.8}$$

To see that this equation leads to the first law (Plastino and Curado (2005)) we go back to the expression for the first law

$$dU - TdS + \sum_{\nu=1}^{M} dA_\nu \lambda_\nu = 0, \tag{4.9}$$

with T the temperature and see what happens when the p_i vary in the fashion $p_i \to p_i + dp_i$. A little algebra yields, up to first order in the dp_i

$$\sum_i \left[C_i^1 + C_i^2 \right] dp_i \equiv \sum K_i dp_i = 0$$

$$C_i^1 = \left[\sum_{\nu=1}^{M} \lambda_\nu \, a_i^\nu + \epsilon_i \right]$$

$$C_i^2 = -kT\left[f(p_i) + p_i\, f'(p_i)\right], \tag{4.10}$$

where the primes indicate derivative with respect to p_i. We proceed to show now that all the K_i are equal. Indeed, select just two of the dp's $\neq 0$, say, dp_i *and* dp_j with the remaining $dp_k = 0$ for $k \neq j$ and $k \neq i$, which entails $dp_i = -dp_j$. In these circumstances, for equation (4.10) to hold we necessarily have $K_i = K_j$. But, since i and j have been arbitrarily chosen, a posteriori we find $K_i = constant = K$ for all i. The value of K will be determined by the normalization condition on the probability distribution, to be determined by the relation:

$$K = D_i^1 + D_i^2 \tag{4.11}$$

$$D_i^1 = \left[\sum_{\nu=1}^{M} \lambda_\nu\, a_i^\nu + \epsilon_i\right]$$
$$D_i^2 = -kT\left[f(p_i) + p_i\, f'(p_i)\right], \tag{4.12}$$

so that we can recast equation (4.12) in the fashion

$$T_i^1 = -\beta\left[\left(\sum_{\nu=1}^{M} \lambda_\nu\, a_i^\nu + \epsilon_i\right) - K\right]$$
$$T_i^2 = f(p_i) + p_i\, f'(p_i), \tag{4.13}$$

which when $\beta \equiv 1/kT$ leads to

$$\sum_i \left[T_i^1 + T_i^2\right] = 0. \tag{4.14}$$

Equation(4.13) comes from the first law while eq. (4.8) comes from the maximum entropy principle. Since it is apparent that the two equations are identical, our proof is complete. In fact,

$$T_i^1 + T_i^2 = 0, \tag{4.15}$$

and can be solved for p_i in terms of the constraints. p_i would then be a maximum entropy distribution.

4.3 Why is the Maximum Entropy Method a Useful Teaching Tool?

There are thousands of maximum entropy applications in the most diverse fields of knowledge. Why is this useful for the teaching of Physics?

In elementary courses the maximum entropy principle illustrates in simple fashion the utility of Lagrange multipliers. These are seen in Calculus but scarcely illustrated in physics lectures, save for a brief mention in Analytical Mechanics. Some maximum entropy examples could already be taught in first year courses without any difficulty.

Of course, the maximum entropy principle should be examined in a more detailed vein in teaching Thermodynamics and Statistical Mechanics. Additionally, maximum entropy can be used with reference to the teaching of *equations of evolution exhibiting the form of continuity equations*. We can mention, for instance, the Liouville equation, the Fokker-Planck equation, Diffusion equations, the Von Neumann's equation in quantum mechanics, etc. This entails a change of perspective. In the preceding discussion we were concerned with discrete probabilities, while we need now continuous ones, i.e., probability densities $f(\mathbf{z})$ for the random (vector) variable $\mathbf{z}$. Let us thus consider a classical system described by a time dependent probability distribution $f(\mathbf{z}, t)$ evolving according to the continuity equation

$$\frac{\partial f}{\partial t} + \nabla \cdot \mathbf{J} = 0, \tag{4.16}$$

where $\mathbf{z}$ denotes a point in the relevant N-dimensional phase space and $\mathbf{J}$ is the flux vector (which, in general, depends on the distribution f). As examples we have:

- i) The one dimensional diffusion equation,

$$\frac{\partial f}{\partial t} - Q\frac{\partial^2 f}{\partial x^2} = 0, \tag{4.17}$$

where Q denotes the diffusion coefficient, and the flux is given by

$$J \,=\, -\,Q\,\frac{\partial f}{\partial x}. \tag{4.18}$$

- ii) The general Liouville equation

$$\frac{\partial f}{\partial t} \,+\, \nabla\cdot(f\,\mathbf{w}) \,=\, 0, \tag{4.19}$$

  with flux

$$\mathbf{J} \,=\, f\,\mathbf{w}. \tag{4.20}$$

  The Liouville equation describes the evolution of an ensemble of classical, deterministic dynamical systems evolving according to the equations of motion

$$\frac{d\mathbf{z}}{dt} \,=\, \mathbf{w}(\mathbf{z}), \tag{4.21}$$

  where $\mathbf{z}$ denotes a point in the concomitant N-dimensional phase space.

- Hamiltonian ensemble dynamics, a particular instance of the Liouville equations (4.21). For Hamiltonian systems with n degrees of freedom we have

  1. $N = 2n$,
  2. $\mathbf{z} = (q_1, \ldots, q_n, p_1, \ldots, p_n)$,
  3. $w_i = \partial H/\partial p_i,\ \ (i = 1, \ldots, n),$ and
  4. $w_{i+n} = -\partial H/\partial q_i,\ \ (i = 1, \ldots, n),$

  where the q_i and the p_i stand for generalized coordinates and momenta, respectively.

With reference to the last item note that Hamiltonian dynamics i) exhibits the important feature of being divergence-free

$$\nabla\cdot\mathbf{w} = 0, \tag{4.22}$$

and ii) for it the Liouville equation simplifies to

$$\frac{\partial f}{\partial t} + \mathbf{w} \cdot \nabla f = 0, \tag{4.23}$$

equivalent to a relationship obeyed by the total time derivative

$$\frac{df}{dt} = 0, \tag{4.24}$$

that is computed along an individual phase-space's orbit. This last form of Liouville-equation for divergenceless systems has an important consequence: if $f(\mathbf{z}, t)$ is a solution of equations (4.23)-(4.24), so is any function $g[f(\mathbf{z}, t)]$.

4.4 Maximum Entropy Ansatz for the Continuity Equation

A central point for our present discussion is that of considering a specially important ansatz for solving the equation of continuity (4.16), namely, the maximum entropy one, that writes

$$f_{ME} = \frac{1}{Z} \exp\left[-\sum_{i=1}^{M} \lambda_i A_i \right], \tag{4.25}$$

where the $A_i(\mathbf{z})$ are M appropriate quantities that are functions of the phase space location $\mathbf{z}$, and the partition function Z (normalization constant) is given by,

$$Z = \int \exp\left[-\sum_{i=1}^{M} \lambda_i A_i \, d^N z \right]. \tag{4.26}$$

The probability distribution (4.25) is the one that maximizes the entropy (here we are dealing with continuous probability distributions, and the summations appearing in previous sections are replaced by integrals),

$$S[f] = -\int f \ln f \, d^N z, \tag{4.27}$$

under the constraints imposed by normalization and the relevant mean values,

$$\langle A_i \rangle = \int A_i \, f \, d^N z. \tag{4.28}$$

The relevant mean values $\langle A_i \rangle$ and the associated Lagrange multipliers λ_i are related by the celebrated Jaynes' relations (see equations (1.15 and 1.17))

$$\lambda_i = \frac{\partial}{\partial \langle A_i \rangle} S, \tag{4.29}$$

and

$$\langle A_i \rangle = -\frac{\partial}{\partial \lambda_i} (\ln Z). \tag{4.30}$$

All the time dependence of the maximum entropy distribution (4.25) is contained in the Lagrange multipliers $\lambda_i(t)$, which are assumed to be time dependent. The Lagrange multipliers change in time, in order to accommodate to the evolving mean values $\langle A_i \rangle$. Now, in general, the time derivatives of the aforementioned mean values are

$$\frac{d}{dt} \langle A_i \rangle = -\int A_i \nabla \cdot \mathbf{J} \, d^N z, \quad i = 1, \ldots M. \tag{4.31}$$

Integrating by parts and making the usual assumption that $\mathbf{J} \to 0$ quickly enough as $|z| \to \infty$, surface terms vanish (they do in most physics problems) and we finally obtain

$$\frac{d}{dt} \langle A_i \rangle = \int \mathbf{J} \cdot \nabla A_i \, d^N z, \ \ (i = 1, \ldots, M). \tag{4.32}$$

The integrals appearing in the right hand sides of these equations generally involve, unfortunately, new mean values not included in the original set $\langle A_i \rangle$ $(i = 1, \ldots, M)$ (remember that the flux $\mathbf{J}$ depends on the distribution f). One way to implement the maximum entropy approach to solve the evolution equation (4.16) is to evaluate, at each instant of time, the right hand sides of (4.31) using the maximum entropy anzats (4.25). In this way, the system of equations (4.31) can be translated into a system of equations of motion for the Lagrange multipliers λ_i. This approach will yield exact solutions, or only approximate solutions, depending on the specific form of the evolution equation (4.16) (Frank

(2005); Malaza *et al.* (1999); Plastino (2001); Plastino *et al.* (1997b); Plastino and Plastino (1998)).

4.5 Maximum Entropy Solution to the Liouville Equation

According to equation (4.32), and remembering that, for the Liouville equation, the flux is given by $\mathbf{J} = f\mathbf{w}$, the temporal evolution of the mean values of the dynamical quantities A_i is

$$\frac{d\langle A_i\rangle}{dt} = \int f\mathbf{w}\cdot\nabla A_i\, d^N z = \langle \mathbf{w}\cdot\nabla A_i\rangle\,, \quad (i = 1,\ldots,M)\,. \tag{4.33}$$

Here we are going to assume that f is given by the ansatz (4.25)-(4.26). We can then regard the quantities Z, f, and λ_i's as functions of the set $\langle A_1\rangle,\ldots,\langle A_M\rangle$. Alternatively, it is also possible to regard all relevant quantities as functions of the λ_i's instead. Making use of the Jaynes' relation (1.18),

$$\begin{aligned}\frac{\partial\lambda_i}{\partial\langle A_j\rangle} &= \frac{\partial^2 S}{\partial\langle A_j\rangle\,\partial\langle A_i\rangle}\\ &= \frac{\partial\lambda_j}{\partial\langle A_i\rangle},\end{aligned} \tag{4.34}$$

the time derivative of the Lagrange multipliers reads

$$\begin{aligned}\frac{d\lambda_i}{dt} &= \sum_{j=1}^{M}\frac{\partial\lambda_i}{\partial\langle A_j\rangle}\frac{d\langle A_j\rangle}{dt}\\ &= \sum_{j=1}^{M}\frac{\partial\lambda_j}{\partial\langle A_i\rangle}\frac{d\langle A_j\rangle}{dt}\\ &= \frac{\partial}{\partial\langle A_i\rangle}\left\{\sum_{j=1}^{M}\lambda_j\frac{d\langle A_j\rangle}{dt}\right\} - \sum_{j=1}^{M}\lambda_j\frac{\partial}{\partial\langle A_i\rangle}\frac{d\langle A_j\rangle}{dt}.\end{aligned} \tag{4.35}$$

Now, from (4.33), the form of f_{ME} (4.25) and since as $|\bar{z}|\to\inf$, $f\to 0$ in rapid enough fashion we find that

$$\begin{aligned}
\sum_{j=1}^{M} \lambda_j \frac{d\langle A_j\rangle}{dt} &= \sum_{j=1}^{M} \lambda_j \langle \mathbf{w}\cdot\nabla A_j\rangle \\
&= \left\langle \mathbf{w}\cdot\nabla\left(\sum_{j=1}^{M}\lambda_j A_j\right)\right\rangle \\
&= -\int f\nabla(\ln f)\cdot\mathbf{w}\, d^N z \\
&= -\int \nabla f\cdot\mathbf{w}\, d^N z \\
&= -\nabla\int \mathbf{w} f\, d^N z + \int f\nabla\cdot\mathbf{w}\, d^N z \\
&= \langle\nabla\cdot\mathbf{w}\rangle . \qquad (4.36)
\end{aligned}$$

The equation of motion for the Lagrange multipliers then becomes

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^{M}\left[\lambda_j\int\frac{\partial f}{\partial\langle A_i\rangle}\mathbf{w}\cdot\nabla A_j\, d^N z - \frac{\partial\langle\nabla\cdot\mathbf{w}\rangle}{\partial\langle A_i\rangle}\right]. \qquad (4.37)$$

Note that, for the important instance of a divergenceless flow, which implies that $\nabla\cdot\mathbf{w}=0$, equation (4.37) specializes to

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^{M}\lambda_j\frac{\partial}{\partial\langle A_i\rangle}\frac{d\langle A_j\rangle}{dt}. \qquad (4.38)$$

It is often the case that we deal with a set of relevant quantities A_i, $(i = 1,\ldots,M)$ entering equations (4.25)-(4.26) such that

$$\mathbf{w}\cdot\nabla A_i = \sum_{j}^{M} C_{ij}\, A_j, \quad (i=1,\ldots,M), \qquad (4.39)$$

where the C_{ij} constitute a set of (structure) constants. This entails, remembering that $d\langle A_i\rangle/dt = \langle\mathbf{w}\cdot\nabla A_i\rangle$,

$$\frac{d\langle A_i\rangle}{dt} = \sum_{j}^{M} C_{ij}\,\langle A_j\rangle, \quad (i=1,\ldots,M). \qquad (4.40)$$

Now, if $\nabla\cdot\mathbf{w}=0$, we have, for temporal evolution of the Lagrange multipliers in equations (4.25)-(4.26)

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^{M} \lambda_j \frac{\partial}{\partial \langle A_i \rangle} \frac{d\langle A_j \rangle}{dt}, \tag{4.41}$$

so that

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^{M} \lambda_j \frac{\partial}{\partial \langle A_i \rangle} \left[\sum_k C_{jk} \langle A_k \rangle \right], \tag{4.42}$$

which yields the equation of motion for the Lagrange multipliers in the fashion

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^{M} C_{ji} \lambda_j. \tag{4.43}$$

We can now study the time-evolution of $\sum_{i=1}^{M} \lambda_i A_i$ using equations (4.39)-(4.43) and the fact that this dependence is entirely contained in the Lagrange multipliers. Thus,

$$\begin{aligned} \frac{\partial}{\partial t} \left(\sum_{i=1}^{M} \lambda_i A_i \right) &= \sum_i \frac{d\lambda_i}{dt} A_i \\ &= -\sum_i A_i \left[\sum_{j=1}^{M} C_{ji} \lambda_j \right], \end{aligned} \tag{4.44}$$

that, after interchanging sums over i and j yields

$$\begin{aligned} \frac{\partial}{\partial t} \left(\sum_{i=1}^{M} \lambda_i A_i \right) &= -\sum_j \lambda_j \left[\sum_{i=1}^{M} C_{ji} A_i \right] \\ &= -\sum_j \lambda_j (\mathbf{w} \cdot \nabla A_j) \\ &= -\mathbf{w} \cdot \nabla \sum_j \lambda_j A_j, \end{aligned} \tag{4.45}$$

i.e.,

$$\frac{\partial}{\partial t} \left(\sum_{i=1}^{M} \lambda_i A_i \right) + \mathbf{w} \cdot \nabla \left(\sum_{i=1}^{M} \lambda_i A_i \right) = 0, \tag{4.46}$$

which entails that $\sum_{i=1}^{M} \lambda_i A_i$ is an *exact* solution of Liouville's equation for divergenceless systems (eg. Hamiltonian systems), and so is (because of equation (4.24)) any function of this quantity like the one that interest us here, this is, the maximum entropy ansatz (4.25)-(4.26).

4.5.1 Example: Application to the Harmonic Oscillator

As a simple illustration of the above ideas we are going to consider maximum entropy solutions of the Liouville equation associated with a one dimensional harmonic oscillator with time dependent frequency $\omega(t)$. Given the harmonic oscillator Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2} m w^2(t) q^2, \tag{4.47}$$

we have to deal with the following observables, that take here the place of the $\langle A_i \rangle$'s, namely,

$$\langle p \rangle, \langle q \rangle, \langle p^2 \rangle, \langle q^2 \rangle, \langle pq \rangle. \tag{4.48}$$

Making use of Hamilton's equations we find

$$\frac{d\langle p \rangle}{dt} = \langle \dot{p} \rangle = \left\langle -\frac{\partial H}{\partial q} \right\rangle = -m w^2(t) \langle q \rangle, \tag{4.49}$$

$$\frac{d\langle q \rangle}{dt} = \langle \dot{q} \rangle = \left\langle \frac{\partial H}{\partial p} \right\rangle = \frac{\langle p \rangle}{m}, \tag{4.50}$$

$$\frac{d\langle p^2 \rangle}{dt} = \left\langle \frac{d}{dt} p^2 \right\rangle = 2 \langle p\dot{p} \rangle = -2 m w^2(t) \langle pq \rangle, \tag{4.51}$$

$$\frac{d\langle q^2 \rangle}{dt} = \left\langle \frac{d}{dt} q^2 \right\rangle = 2 \langle q\dot{q} \rangle = \frac{2\langle pq \rangle}{m}, \tag{4.52}$$

and

$$\frac{d\langle pq \rangle}{dt} = \left\langle \frac{d}{dt} pq \right\rangle = \langle \dot{p} q \rangle + \langle p\dot{q} \rangle = -m w^2(t) \langle q \rangle^2 + \frac{\langle p \rangle^2}{m}. \tag{4.53}$$

In the harmonic oscillator case we have a divergenceless flow so that equation (4.38) applies,

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^{M} \lambda_j \frac{\partial}{\partial \langle A_i \rangle} \frac{d\langle A_j \rangle}{dt}, \tag{4.54}$$

wherefrom we find:

$$\frac{d\lambda_p}{dt} = -\frac{\lambda_q}{m}, \tag{4.55}$$

$$\frac{d\lambda_q}{dt} = \lambda_p m w^2(t), \tag{4.56}$$

$$\frac{d\lambda_{p^2}}{dt} = -\frac{\lambda_{pq}}{m}, \tag{4.57}$$

$$\frac{d\lambda_{q^2}}{dt} = \lambda_{pq} m w^2(t), \tag{4.58}$$

and

$$\frac{d\lambda_{pq}}{dt} = \lambda_{p^2} 2 m w^2(t) - \frac{\lambda_{q^2} 2}{m}. \tag{4.59}$$

The system of linear differential equations (4.55-4.59) for the Lagrange multipliers λ_i can be solved (given a specific form of $w(t)$) by a variety of standard methods. Given a particular solution $\lambda_i(t)$, the maximum entropy ansatz (remember that all the time dependence of $f(q,p,t)$ is through the Lagrange multipliers λ_i)

$$f(q,p,t) \;=\; \frac{1}{Z}\exp(-\lambda_q\, q - \lambda_p\, p - \lambda_{q^2}\, q^2 - \lambda_{qp}\, qp - \lambda_{p^2}\, p^2), \tag{4.60}$$

with

$$Z \;=\; \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp(-\lambda_q\, q - \lambda_p\, p - \lambda_{q^2}\, q^2 - \lambda_{qp}\, qp - \lambda_{p^2}\, p^2)\, dq\, dp, \tag{4.61}$$

constitutes an exact time dependent solution to the Liouville equation of the harmonic oscillator. It is worth mentioning that standard textbooks on classical dynamics rarely provide examples of exact, time dependent solutions of Liouville equation.

4.6 Conclusions

This effort has revolved around the idea of giving the Maximum Entropy Methodology a more important place in the physics curriculum than it has now. The following points have been emphasized:

1. The maximum entropy principle constitutes an interesting application of the Lagrange multipliers technique, and some aspects of it could already be taught in elementary Calculus courses.

2. The maximum entropy principle provides the foundation of statistical mechanics, not only in its equilibrium version but also in its non-equilibrium one.

3. The maximum entropy method constitutes a useful didactic tool for comfortably tackling *other* aspects of theoretical physics, as it provides a simple and elegant method for obtaining *analytical* solutions of several evolution equations, like the Liouville, diffusion, and Fokker-Planck ones.

4. The maximum entropy method today is an indispensable tool in Physics, Chemistry, Engineering, etc., with which to confront "real world" problems.

Of course, all these points are inextricably intertwined. In this contribution we have focused attention on point 3, providing a simple and informative application that any attentive student of physics should understand.

Chapter 5

Conclusions

The subject of the present work was to investigate the application of maximum entropy schemes for solving evolution equations describing time dependent processes. In particular we considered:

1. A collisional Vlasov equation with a Fokker-Planck like collision term. (Schönfeldt *et al.*, Physica A, 369 (2006) 408-416).

2. Evolution equations with the form of continuity equations with sources. (Schönfeldt *et al.*, Physica A, 374 (2007) 573-584).

3. The use of simple illustrations of the afore mentioned techniques in physics education. (Schönfeldt *et al.*, Rev. Mex. Fis. E, 52 (2) (2006) 151-159).

We investigated a maximum entropy procedure for constructing solutions for a collisional Vlasov equation. We showed that the aforementioned equation admits exact maximum entropy solutions, as opposed to the alternative maximum entropy solutions previously advanced by El-Wakil *et al.* (2003) that were only approximate. We considered two different approaches to the exact maximum entropy solutions of the collisional Vlasov equation. In the first instance we identified an appropriate set of five relevant mean values that evolve according to a closed set of coupled , ordinary, linear differential equations. We then constructed exact maximum entropy solutions associated with that set of mean values. We investigated these solutions using both the equations of motion of the mean values, as well as the equations of motion of the corresponding Lagrange multipliers.

In the second instance we proved that it is possible to obtain exact solutions of the reduced equation considered by El-Wakil *et al.* (2003), if the zeroth-order moment of the solutions is explicitly taken into account. In order to do this, we proved that some of the partial differential equations considered by El-Wakil *et al.* (2003) admit exact solutions. The maximum entropy solutions considered by El-Wakil *et al.* (2003) assumed a constant normalization of 1. Where as the exact solutions, to the equation they considered, have a time-dependent normalization. Their solutions, neglecting this important piece of information, are thus bound to be only approximate.

We considered a maximum entropy approach to construct approximate, time dependent solution to evolution equations endowed with source terms. We showed that in some particular cases the method leads to exact time dependent solutions. By construction our present implementation of the maximum entropy prescription complies with the exact equations of motion of the relevant mean values. Moreover, it always (even in the case of approximate solutions) preserves the exact functional relationship between the time derivative of the entropy and the time dependent solutions of the evolution equation. This means that any H-theorem verified when evaluating the entropy functional upon the exact solutions is also verified when evaluating the entropy upon the maximum entropy approximate treatments.

We explored the idea of giving the maximum entropy methodology a more important place in the physics curriculum than it has now. The points emphasized were:

1. The maximum entropy principle constitutes an interesting application of the Lagrange multipliers technique.

2. The maximum entropy principle provides the foundation of statistical mechanics, not only in its equilibrium version but also in its non-equilibrium one.

3. The maximum entropy principle constitutes a useful didactic tool for comfortably tackling other aspects of theoretical physics, since it is an important and elegant method to obtain analytical solutions of several evolution equations.

4. The maximum entropy principle is an indispensable tool in Physics, Chemistry, Engineering, etc., for solving "real world" problems.

We focused our attention on point 3, providing simple and informative illustrations of the maximum entropy approach to some basic evolution equations of physics.

References

Agrawal, D., Singh, J.K., and Kumar, A. *Biosystems Engineering*, **90** (**2005**) 103.

Amemiya, K., Moriwaki, S., and Mio, N. *Physics Letters A*, **315** (**2003**) 184.

Andrey, L. *Physics Letters A*, **111** (**1985**) 45.

Baierlein, R. (Editor) *Atoms and Information Theory* (W.H. Freeman, San Francisco, **1971**).

Baker-Jarvis, J., Racine, M., and Alameddine, J. *J. Math. Phys.*, **30** (**1989**) 1459.

Beck, C. and Schlögl, F. *Thermodynamics of Chaotic Systems: An Introduction* (Cambridge University Press, Cambridge, **1993**).

Binney, J. and Tremaine, S. *Galactic Dynamics* (Princeton University Press, Princeton, **1987**).

Blokhin, A.M., Bushmanov, R.S., and Romano, V. *International Journal of Engineering Science*, **42** (**2004**) 915.

Boghosian, B.M. *Phys. Rev. E*, **53** (**1996**) 4754.

Borland, L. *Physica D*, **99** (**1996**) 175.

Borland, L. *Phys. Rev. Lett.*, **89** (**2002**) 098701.

Borland, L., Pennini, F., Plastino, A.R., and Plastino, A. *Eur. Phys. J. B*, **12** (**1999**) 285.

REFERENCES

Brillouin, L. *Science and Information Theory* (Academic Press Inc., New York, **1962**).

Callen, H.B. (Editor) *Thermodynamics* (J. Wiley, New York, **1960**).

Cercignani, C. *The Boltzmann Equation and Its Applications* (Springer-Verlag, New York, **1988**).

Clowser, J. and Strouthos, C. *Nuclear Physics B*, **106-107** (**2002**) 489.

da Silva, P.C., da Silva, L.R., Lenzi, E.K., Mendes, R.S., and Malacarne, L.C. *Physica A*, **342** (**2004**) 16.

Desloge, E.A. *Thermal Physics* (Holt, Rhinehart and Winston, New York, **1968**).

El-Hanbaly, A.M. and Elgarayhi, A. *J. Plasma Phys.*, **59** (**1998**) 169.

El-Wakil, S.A., Elhanbaly, A., and Abdou, M.A. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **69** (**2001**) 41.

El-Wakil, S.A., Elhanbaly, A., and Abdou, M.A. *Physica A*, **323** (**2003**) 213.

Elgarayhi, A. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **75** (**2002**) 1.

Frank, T.D. *Nonlinear Fokker-Planck Equations: Fundamentals and Applications* (Springer-Verlag, Berlin, **2005**).

Frieden, B.R. *Physics from Fisher Information* (Cambridge University Press, Cambridge, **1998**).

Frieden, B.R. *Science from Fisher Information* (Cambridge University Press, Cambridge, **2004**).

Frieden, B.R. and Soffer, B.H. *Phys. Rev. E*, **52** (**1995**) 2274.

Gell-Mann, M. and Tsallis, C. (Editors) *Nonextensive Entropy - Interdisciplinary Applications* (Oxford University Press, USA, **2004**).

Gibbs, J.W. *Elementary Principles in Statistical Mechanics* (C. Scribner's sons, New York, **1902**).

Gong, Y., Han, M., Hua, W., and Xu, W. *Computer Vision and Image Understanding*, **96** (**2004**) 181.

Grandy, W.T. and Milonni, P.W. (Editors) *Physics and Probability. Essays in Honor of Edwin T. Jaynes* (Cambridge University Press, New York, **1993**).

Havrda, J.H. and Charvat, F. *Kybernetica*, **3** (**1967**) 30.

Hick, P. and Stevens, G. *Astron. Astrophys.*, **172** (**1987**) 350.

Israel, S., Saravanan, R., Srinivasan, N., and Rajaram, R.K. *Journal of Physics and Chemistry of Solids*, **64** (**2003**) 43.

Jaynes, E.T. *Phys. Rev.*, **106** (**1957**) 620.

Jaynes, E.T. In *Statistical Physics*, edited by Ford, K. (Benjamin, New York, **1963**).

Jaynes, E.T. In *E.T. Jaynes: Papers on Probability, Statistics and Statistical Physics*, edited by Rosenkrantz, R.D. (Kluwer Academc Publishers, Dordrecht, **1983**).

Jaynes, E.T. *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, **2003**).

Kapur, A. *Maximum-Entropy Models in Science and Engineering* (Wiley Eastern Limited, New Delhi, **1989**).

Katz, A. *Principles of Statistical Mechanics* (W.H. Freeman and Co., San Francisco, **1967**).

Khinchin, A.I. *Mathematical Foundations of Information Theory* (Dover, New York, **1957**).

Kim, S.K. and Lee, W.I. *International Journal of Heat and Mass Transfers*, **45** (**2002**) 381.

REFERENCES

Kullback, S. *Information Theory and Statistics* (Wiley, New York, **1959**).

Liouville, J. *J. Math. Pure. Appl.*, **3** (**1838**) 342.

Lukacs, A. and Papp, E. *J Photochem Photobiol B*, **77** (**2004**) 1.

Malaza, E.D., Miller, H.G., Plastino, A.R., and Solms, F. *Physica A*, **265** (**1999**) 224.

Nielsen, M.A. and Chuang, I.L. *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, **2000**).

Pathria, R.K. *Statistical Mechanics* (Pergamon Press, Exeter, **1993**).

Plastino, A. and Curado, E. *Phys. Rev. E*, **72** (**2005**) 047103.

Plastino, A. and Plastino, A.R. *Physics Letters A*, **226** (**1997**) 257.

Plastino, A., Plastino, A.R., and Miller, H.G. *Physics Letters A*, **232** (**1997**a) 349.

Plastino, A.R. In *NonExtensive Statistical Mechanics and Its Applications, Lecture Notes in Physics.*, edited by Abe, S. and Okamoto, Y. (Springer-Verlag, Berlin, **2001**).

Plastino, A.R., Miller, H.G., and Plastino, A. *Phys. Rev. E*, **56** (**1997**b) 3927.

Plastino, A.R., Miller, H.G., Plastino, A., and Yen, G.D. *J. Math. Phys.*, **38** (**1997**c) 6675.

Plastino, A.R. and Plastino, A. *Physica A*, **258** (**1998**) 429.

Raychaudhuri, S., Chang, J.T., Sutphin, P.D., and Altman, R.B. *Genome Research*, **12** (**2002**) 203.

Reif, F. *Statistical and Thermal Physics* (McGraw-Hill, New York, **1965**).

Rényi, A. *Proc. 4th Berkeley Symp. Math. Stat. and Prob*, **1** (**1961**) 547.

Ruelle, D. *Physics Today*, **57** (**2004**) 48.

Sakurai, J. *Modern Quantum Mechanics* (Benjamin, Menlo Park, Ca., **1985**).

Scalapino, D.J. In *Physics and Probability. Essays in Honor of Edwin T. Jaynes*, edited by Grandy, W.T. and Milonni, P.W. (Cambridge University Press, Cambridge, **1993**).

Schönfeldt, J.-H., Jimenez, N., Plastino, A.R., Plastino, A., and Casas, M. *Physica A*, **374** (**2007**) 573.

Schönfeldt, J.-H. and Plastino, A.R. *Physica A*, **369** (**2006**) 408.

Schönfeldt, J.-H., Roston, G.B., Plastino, A.R., and Plastino, A. *Rev. Mex. Fis. E*, **52** (**2006**) 151.

Shams, A.A., Picozzi, S., and Malik, F.B. *Physica B*, **352** (**2004**) 269.

Shannon, C.E. *Bell Syst. Tech. J.*, **27** (**1948**) 379.

Sieniutycz, S. and Farkas, H. (Editors) *Variational and Extremum Principles in Macroscopic Systems* (Elsevier, Amsterdam, **2005**).

Steinmetz, M. *Astrophysics and Space Science*, **269-270** (**1999**) 513.

Tribus, M. *Thermostatics and Thermodynamics* (Van Nostrand, New York, **1961**).

Tsallis, C. and Bukman, D. *Phys. Rev. E*, **54** (**1996**) 2197.

Uhlenbeck, G.E. In *Probability and Related Topics in Physical Sciences*, edited by Kac, M. (Interscience Publishers, Inc., London, **1957**).

van Kampen, N.G. *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, **1992**).

Wyllie, G.A.P. *Elementary Statistical Mechanics* (Hutchinson, London, **1970**).

Yamano, T. *Eur. Phys. J. B*, **18** (**2000**) 103.