Maximum entropy principle, evolution equations, and physics education principio de máxima entropía como herramienta didáctica para discutir ecuaciones de evolución temporal

J-H. Schonfeldt[a], G.B. Roston[a], A.R. Plastino[a], and A. Plastino[a,b]
[a]*Department of Physics, University of Pretoria, Pretoria 0002, South Africa.*
[b]*Facultad de Ciencias Exactas, Universidad Nacional de La Plata, IFLP-CONICET, C.C. 727, 1900 La Plata, Argentina.*

Recibido el 27 de 09 de 2005; aceptado el 22 de 03 de 2006

The landscape of Physics is in a constant state of change and the structure of the University level Physics Curriculum needs to be adapted to this state of affairs. One of the most interesting current features of physics is the increasing importance of multidisciplinary studies. Methods and ideas from physics are being applied to diverse areas of science ranging from biology and economics to sociology and linguistics. Statistical Physics (SP) provides the most fertile set of methods for these kind of applications. The aim of the present contribution is to show how a powerful idea from SP that is widely applied in many fields, the maximum entropy principle (MaxEnt), can be integrated into the physics curriculum. First of all, the constrained maximization of an entropic measure provides an important illustration of the Lagrange multipliers technique, which is part of the standard calculus course for physics students. Secondly, MaxEnt provides the basis for an alternative foundation for statistical mechanics, which is nowadays being considered in some modern textbooks on SP. In point of fact, the main role usually assigned to MaxEnt (as a tool for teaching theoretical physics) is in connection with the Gibbs canonical and grand canonical ensembles. However, as we shall here explain, MaxEnt also constitutes a useful tool in the teaching of other aspects of theoretical physics: it provides an elegant and simple method for obtaining analytical solutions for several evolution equations, like the Liouville equation, the diffusion equation, and the Fokker-Planck equation. Last but certainly not least, MaxEnt belongs to the tool-kit that physicist use to solve concrete "real-world" problems.

Keywords: Maximum entropy principle; continuity equations; Liouville equation.

El panorama de la física contemporánea se encuentra en un estado de continuo cambio y por ende la estructura de los planes de estudio del área necesita adaptarse a tal situación. La creciente importancia de la multidisciplinariedad es hoy faceta típica de la actividad en física. Técnicas e ideas de origen físico están siendo aplicados con éxito en áreas diversas. Biología y economía constituyen ejemplos importantes. La física estadística (FE) es la principal proveedora de métodos para este tipo de aplicaciones. Este trabajo se refiere a una idea muy fecunda (y de amplia aplicación) de la FE, el llamado "principio de máxima entropía" (PME). Pretendemos aquí mostrar *como* puede el PME ser integrado con provecho en la currícula de la física. Se verá que los cursos de mecánica estadística no son los únicos donde este principio puede ser exitosamente incorporado. En particular, ilustraremos como el PME puede ser empleado para construir soluciones analíticas relativamente sencillas para ecuaciones de evolución muy importantes, tales como las de Liouville y Fokker-Planck.

Descriptores: Principio de maxima entropía; ecuaciones de continuidad; ecuación de Liouville.

PACS: 05.40.-a; 05.20.Gg

1. Introduction

The contents and structure of the physics curriculum have been in continuous evolution since the last quarter of the XIX 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 the last century witnessed deep, rapid changes arising from relativity and quantum revolutions. On the other hand, during the second half of the 20th century, the changes made to the physics curriculum werw not that dramatic. This (relatively speaking) "stationary state" had the psychological consequence that some physicists seem to believe that we had 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 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.).

Of course, the physics curriculum must have a finite length. Consequently, it is not possible to incorporate new

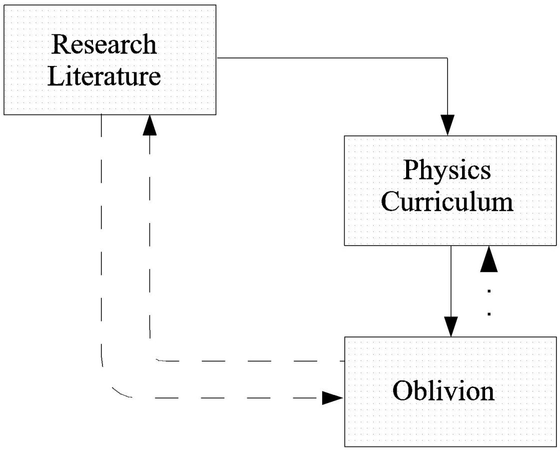

FIGURE 1. The flow of physical knowledge.

contents in to the curriculum without at the same time effecting 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 Fig. 1) of these "great topis", 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 might 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 topis out of the current research literature into "oblivion" (dashed lines in Fig. 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 has fallen 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 these general, unifying ideas that play an important role in current research. It is, undoubtedly, one of the most fundamental tools in statistical physics, from both the conceptual and the practical points of view. It was first mentioned by Gibbs himself in his famous book on statistical mechanics [1]. 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 [2]. There are already several textbooks on equilibrium statistical mechanics that develop this subject taking as its basis the maximum entropy principle [3–6]. However, the scientific relevance of the maximum entropy principle (as wellas the information theory concepts behind it) goes well beyond the study of equilibrium statistical mechanics. One of the first places in which this was explored is the classic work by Brillouin [7]. The great number of applications of the maximum entropy principle to various areas of science attest to this. 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 shall mention some now recent applications.

- In Ref. 8, the principle of maximum entropy (MaxEnt) yields a conditional probability distribution model for estimating the runoff for the catchment (watershed) of the Matatila dam in India. The model predicts runoff, subject to the selected constraints, in response to a given rainfall, in a rather adequate fashion.

- In Ref. 9, 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 Ref. 10, 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 Ref. 11, 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 Ref. 12, 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 Hubbard-Anderson Hamiltonian.

- In Ref. 13, 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 Ref. 14, 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 MaxEnt.

- In Ref. 15, 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 Ref. 16, MaxEnt 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 pseudoscalar meson, and the symmetric phase resonances.

- In Ref. 17, 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 Ref. 18, the possibility that statistical, natural-language processing techniques could be used to assign Gene-Ontology codes is explored. It is shown that maximum entropy modeling outperforms other methods for associating a set of GO codes (for biological process) to literature-abstracts and thus to the genes associated with the abstracts.

- In Ref. 19, the MaxEnt 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 article is to provide some hints on how the maximum entropy principle can be incorporated in to 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.

2. Brief review of the MaxEnt ideas

The second law of thermodynamics [20] is one of physics' most important statements. In fact the first and second laws, 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 [21–24]. The most basic ingredient of such an explanation is a microscopic probability distribution (PD) that controls the population of microstates of the system under consideration [21]. Primarily, the maximum entropy approach, or MaxEnt, is an algorithm designed to obtain this PD. In order to make sense of it, however, we must first of all extract the concept of entropy from its natural thermodynamic surroundings and give it a more general meaning [2, 24, 25].

2.1. Entropy as ignorance

Jaynes [2] pioneered MaxEnt based on the idea of regarding entropy as a measure of our ignorance in a given situation [24]. The following example is quite instructive and conveys the essentials of this interpretation of entropy [24]. If you have N *identical* cells and in one of them a specific feature can be encountered, but you do not know *in which* of them you have to look for it, then your ignorance I in such a situation (A) depends only on N ($I = I(N)$), (B) increases monotonically with N, and (C) reduces to zero for $N = 1$. Further, if each cell is subdivided into M sub-cells, your ignorance is now

- (1) that of selecting among the N possibilities, namely, $I(N)$ plus

- (2) that of choosing among the M subdivisions, that is, $(I(M))$.

(D) Ignorance is additive: if Alice has information about (1) only, and Bob about (2) only, you can overcome it by speaking to both of them to get $I(N + M) = I(N) + I(M)$. The only function compatible with A-D is the logarithmic one, *i.e.*

$$I = k \ln N, \tag{1}$$

with k a constant that we can use as the unit of information. It is easy to show [24] that for N non-identical cells, each endowed with a probability $p_i, (i = 1, \ldots, N)$, (5) generalizes to

$$I = -k \sum_{i=1}^{N} p_i \ln p_i, \tag{2}$$

a quantity that measures the amount of ignorance attached to a probability distribution. I measures what we still have to learn in order to reach the ideal situation of certainty, namely

$$p_i = 1 \qquad p_{m\neq i} = 0. \tag{3}$$

Equation (2), with $k = k_B$, the Boltzmann constant, yields the thermodynamic entropy in statistical mechanics [22], but we clearly appreciate the fact that this equation has a much wider span.

2.2. A derivation of thermodynamics' first law from MaxEnt

As a physical example of MaxEnt 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 exclusively affect 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 p_i 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 [22]. Consequently, one must show that,

- given a concave entropic form (or information measure (IM)) 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
- for any system described by a microscopic probability distribution (PD) $\{p_i\}$,
- assuming a reversible process via $p_i \to p_i + dp_i$,
- (**Thesis**) if a normalized PD $\{p_i\}$ maximizes S, with the numerical values of U and the M A_ν as constraints, this entails

$$dU = TdS - \sum_{\nu=1}^{M} \gamma_\nu \, dA_\nu$$

First Law of Thermodynamics. (4)

2.2.1. Proof

Consider [26, 27] a quite *general* information measure of the form

$$S = k \sum_i p_i \, f(p_i), \tag{5}$$

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 the Hilbert space and $\mathcal{P} = \{p_i\}$ is an (as yet unknown) normalized probability distribution such that

$$\sum_i p_i = 1. \tag{6}$$

Let f be an arbitrary smooth function of the p_i. Further, consider M quantities A_ν that represent mean values of extensive physical quantities $\mathcal{A}_\nu$. These take on, for the state i, the value a_i^ν with probability p_i. Also, we suppose that g is another arbitrary smooth, monotonic function of the p_i such that $g(0) = 0$ and $g(1) = 1$. We do not require the condition

$$\sum_i g(p_i) = 1.$$

The mean energy U and the A_ν are given by

$$U = \sum_i \epsilon_i \, g(p_i) \qquad A_\nu = \sum_i a_i^\nu \, g(p_i). \tag{7}$$

Assume now that the set $\mathcal{P}$ changes in such a way that

$$p_i \to p_i + dp_i,$$

with

$$\sum_i dp_i = 0 (\text{Cf.}(6)), \tag{8}$$

which in turn generates corresponding changes dS and dU in S, the A_ν, and U respectively.

We wish to extremize S subject to the constraint of fixing i) U and ii) the M values A_ν. This is achieved via Lagrange multipliers i) β and ii) M γ_ν. We need also a normalization Lagrange multiplier ξ:

$$\delta_{\{p_i\}}[S - \beta U - \sum_{\nu=1}^{M} \gamma_\nu A_\nu - \xi \sum_i p_i] = 0, \tag{9}$$

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

$$0 = \delta_{p_m} \sum_i p_i f(p_i) + \left(-\delta_{p_m} [\sum_i \beta g(p_i) (\sum_{\nu=1}^{M} \lambda_\nu \, a_i^\nu + \epsilon_i) - \xi] \right), \tag{10}$$

so that

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

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

$$0 = f(p_i) + p_i f'(p_i) + \left\{ -\beta [g'(p_i) (\sum_{\nu=1}^{M} \lambda_\nu \, a_i^\nu + \epsilon_i) - K] \right\}. \tag{11}$$

To see that this equation leads to the first law [27], we go back to the expression for the first law

$$dU - TdS + \sum_{\nu=1}^{M} dA_\nu \lambda_\nu = 0, \tag{12}$$

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

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

$$C_i^1 = [\sum_{\nu=1}^{M} \lambda_\nu \, a_i^\nu + \epsilon_i] \, g'(p_i)$$

$$C_i^2 = -kT[f(p_i) + p_i \, f'(p_i)], \tag{13}$$

where the primes indicate the 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 Eq. (13) to hold, we necessarily have $K_i = K_j$. But, since i and j have been arbitrarily chosen, a posteriori we find that $K_i = constant = K$

for all i. The value of K will be determined by the normalization condition on the probability distribution, to be determined by the relation:

$$K = \sum_i D_i^1 + D_i^2$$
$$D_i^1 = [\sum_{\nu=1}^{M} \lambda_\nu \, a_i^\nu + \epsilon_i] \, g'(p_i)$$
$$D_i^2 = -kT[f(p_i) \, + \, p_i \, f'(p_i)], \tag{14}$$

so that we can recast (14) in the form

$$T_i^1 = f(p_i) \, + \, p_i \, f'(p_i)$$
$$T_i^2 = -\beta[(\sum_{\nu=1}^{M} \lambda_\nu \, a_i^\nu + \epsilon_i) \, g'(p_i) - K]$$
$$\beta \equiv 1/kT, \;\; \text{that leads to :}$$
$$\sum_i T_i^1 + T_i^2 = 0. \tag{15}$$

Equation (15) comes from the first law while Eq. (11) comes from MaxEnt. Since it is apparent that the two equations are identical, our proof is complete.

3. Why is MaxEnt a useful teaching tool?

There are thousands of MaxEnt applications in the most fields of knowledge. Why is this useful for the theaching of Physics?

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

Of course, MaxEnt should be examined in more detail in teaching Thermodynamics and Statistical Mechanics. In additional MaxEnt 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 above discussion, we were concerned with discrete probabilities, while we now need 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

$$\frac{\partial f}{\partial t} + \nabla \cdot \mathbf{J} = 0, \tag{16}$$

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,

$$\frac{\partial f}{\partial t} - Q\frac{\partial^2 f}{\partial x^2} = 0, \tag{17}$$

where Q denotes the diffusion coefficient while the flux is given by

$$J = -Q\frac{\partial f}{\partial x}. \tag{18}$$

- ii) The general Liouville equation

$$\frac{\partial f}{\partial t} + \nabla \cdot (f\mathbf{w}) = 0, \tag{19}$$

with flux

$$\mathbf{J} = f\mathbf{w}. \tag{20}$$

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

$$\frac{d\mathbf{z}}{dt} = \mathbf{w}(\mathbf{z}), \tag{21}$$

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

- Hamiltonian ensemble dynamics, a particular instance of the Liouville equations (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

$$\nabla \cdot \mathbf{w} = 0, \tag{22}$$

and

ii) for this reason the Liouville equation simplifies to

$$\frac{\partial f}{\partial t} + \mathbf{w} \cdot \nabla f = 0 \tag{23}$$

equivalent to a relationship obeyed by the total time derivative

$$\frac{df}{dt} = 0, \tag{24}$$

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 to (23)-(24), so is any function $g[f(\mathbf{z}, t)]$.

4. MaxEnt ansatz for the continuity equation

A central point for our present discussion is that of considering an especially important ansatz for solving the equation of continuity (16), namely, the MaxEnt ansatz, that is expressed by

$$f_{ME} = \frac{1}{Z} \exp\left[-\sum_{i=1}^{M} \lambda_i A_i\right], \tag{25}$$

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

$$Z = \int \exp\left[-\sum_{i=1}^{M} \lambda_i A_i d^N z\right]. \tag{26}$$

The probability distribution (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):

$$S[f] = -\int f \ln f d^N z, \tag{27}$$

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

$$\langle A_i \rangle = \int A_i f d^N z. \tag{28}$$

The relevant mean values $\langle A_i \rangle$ and the associated Lagrange multipliers λ_i are related by the well-known Jaynes' relations

$$\lambda_i = \frac{\partial}{\partial \langle A_i \rangle} S, \tag{29}$$

and

$$\langle A_i \rangle = -\frac{\partial}{\partial \lambda_i} (\ln Z). \tag{30}$$

All the time dependence of the maximum entropy distribution (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 the evolving mean values values $\langle A_i \rangle$. Now, in general, the time derivatives of the aforementioned mean values are

$$\frac{d}{dt}\langle A_i \rangle = -\int A_i \nabla \cdot \mathbf{J} d^N z \qquad i = 1, \dots M. \tag{31}$$

Integrating by parts and making the usual assumption that $\mathbf{J} \to 0$ as quickly as $|z| \to \infty$, surface terms vanish (they do in 99.9% of physics problems!) and we finally obtain

$$\frac{d}{dt}\langle A_i \rangle = \int d^N z \, \mathbf{J} \cdot \nabla A_i, \qquad (i = 1, \dots, M). \tag{32}$$

The integrals appearing on the right hand sides of these equations generally involve, unfortunately, new mean values not included in the original set $\langle A_i \rangle (i = 1, \dots, M)$ (remember that the flux $\mathbf{J}$ depends on the distribution f). One way to implement the maximum entropy approach to solving the evolution Eq. (16) is to evaluate, at each instant of time, the right hand sides of (31) using the maximum entropy ansatz (25). In this way, the system of equations (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 (16) [28–32].

5. MaxEnt Solution to the Liouville Equation

According to Eq. (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

$$\begin{aligned} \frac{d\langle A_i \rangle}{dt} &= \int d^N z \, f \mathbf{w} \cdot \nabla A_i \\ &= \langle \mathbf{w} \cdot \nabla A_i \rangle \, (i = 1, \dots, M). \end{aligned} \tag{33}$$

Here we are going to assume that f is given by the ansatz (25)-(26). We can then regard the quantities Z, f, and λ_i's as functions of the set $\langle A_1 \rangle, \dots, \langle A_M \rangle$. Alternatively, it is also possible to regard all relevant quantities as functions of the λ_i's. The time derivative of the Lagrange multipliers reads

$$\begin{aligned} \frac{d\lambda_i}{dt} &= \sum_{j=1}^{M} \frac{\partial \lambda_i}{\partial \langle A_i \rangle} \frac{d\langle A_j \rangle}{dt} \\ \frac{d\lambda_i}{dt} &= \frac{\partial}{\partial \langle A_i \rangle} \left\{ \sum_{j=1}^{M} \lambda_j \frac{d\langle A_j \rangle}{dt} \right\} \\ &\quad - \sum_{j=1}^{M} \lambda_j \frac{\partial}{\partial \langle A_i \rangle} \frac{d\langle A_j \rangle}{dt}. \end{aligned} \tag{34}$$

Now, since

$$\sum_{j=1}^{M} \lambda_j \frac{d\langle A_j \rangle}{dt} = \sum_{j=1}^{M} \lambda_j \langle \mathbf{w} \cdot \nabla A_j \rangle = \langle \nabla \cdot \mathbf{w} \rangle, \tag{35}$$

the equation of motion for the Lagrange multipliers can be written

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^{M} \left[\lambda_j \int \frac{\partial f}{\partial \langle A_i \rangle} \mathbf{w} \cdot \nabla A_j d^N z - \frac{\partial \langle \nabla \cdot \mathbf{w} \rangle}{\partial \langle A_i \rangle} \right]. \tag{36}$$

Note that, for the important instance of a divergenceless flow, which implies that $\nabla \cdot \mathbf{w} = 0$, Eq. (36) specializes to

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^{M} \lambda_j \frac{\partial}{\partial \langle A_i \rangle} \frac{d\langle A_j \rangle}{dt}. \tag{37}$$

It is often the case that we deal with a set of relevant quantities $A_i,\ (i = 1, \ldots, M)$ entering (25)-(26) such that

$$\mathbf{w} \cdot \nabla A_i = \sum_j^M C_{ij}\, A_j,\ \ (i = 1, \ldots, M), \tag{38}$$

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

$$\frac{d\langle A_i\rangle}{dt} = \sum_j^M C_{ij}\, \langle A_j\rangle,\ \ (i = 1, \ldots, M). \tag{39}$$

Now, if $\nabla \cdot \mathbf{w} = 0$, we have, for the temporal evolution of the Lagrange multipliers in (25)-(26)

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^M \lambda_j \frac{\partial}{\partial\langle A_i\rangle} \frac{d\langle A_j\rangle}{dt}, \tag{40}$$

so that

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^M \lambda_j \frac{\partial}{\partial\langle A_i\rangle} \left[\sum_k C_{jk}\langle A_k\rangle\right], \tag{41}$$

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

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^M C_{ji}\lambda_j. \tag{42}$$

We can now study the time-evolution of

$$\sum_{i=1}^M \lambda_i A_i$$

using (38)-(42) and the fact that this dependence is entirely contained in the Lagrange multipliers. Thus,

$$\begin{aligned} \frac{d}{dt}\left(\sum_{i=1}^M \lambda_i A_i\right) &= \sum_i \frac{d\lambda_i}{dt} A_i \\ &= -\sum_i A_i \left[\sum_{j=1}^M C_{ji}\lambda_j\right], \end{aligned} \tag{43}$$

which, after interchanging sums over i and j, yields

$$\begin{aligned} \frac{d}{dt}\left(\sum_{i=1}^M \lambda_i A_i\right) &= -\sum_j \lambda_j \left[\sum_{i=1}^M C_{ji} A_i\right] \\ &= -\sum_j \lambda_j (\mathbf{w} \cdot \nabla A_j) \\ &= -\mathbf{w} \cdot \nabla \sum_j \lambda_j A_j, \end{aligned} \tag{44}$$

i.e.,

$$\frac{d}{dt}\left(\sum_{i=1}^M \lambda_i A_i\right) + \mathbf{w} \cdot \nabla \left(\sum_{i=1}^M \lambda_i A_i\right) = 0, \tag{45}$$

from which it follows that $\sum_{i=1}^M \lambda_i A_i$ is an *exact* solution of Liouville's equation, and so is (because of equation (24) any function of this quantity like the one that interests us here, this is, the MaxEnt anzatz (25)-(26).

5.1. Example: Application to the Harmonic Oscillator

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

$$H = \frac{p^2}{2m} + \frac{1}{2} m w^2(t) q^2, \tag{46}$$

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

$$\langle p\rangle\,, \langle q\rangle\,, \langle p^2\rangle\,, \langle q^2\rangle\,, \langle pq\rangle\,. \tag{47}$$

Making use of Hamilton's equations, we find

$$\begin{aligned} \frac{d\langle p\rangle}{dt} &= \langle \dot{p}\rangle = \left\langle -\frac{\partial H}{\partial q}\right\rangle = -m w^2(t)\,\langle q\rangle\,; \\ \frac{d\langle q\rangle}{dt} &= \langle \dot{q}\rangle = \left\langle \frac{\partial H}{\partial p}\right\rangle = \frac{\langle p\rangle}{m}, \end{aligned} \tag{48}$$

$$\frac{d\langle p^2\rangle}{dt} = \left\langle \frac{d}{dt} p^2\right\rangle = 2\,\langle p\dot{p}\rangle = -2m w^2(t)\,\langle pq\rangle\,, \tag{49}$$

$$\frac{d\langle q^2\rangle}{dt} = \left\langle \frac{d}{dt} q^2\right\rangle = 2\,\langle q\dot{q}\rangle = \frac{2\,\langle pq\rangle}{m}, \tag{50}$$

and

$$\begin{aligned} \frac{d\langle pq\rangle}{dt} &= \left\langle \frac{d}{dt} pq\right\rangle \\ &= \langle \dot{p}q\rangle + \langle p\dot{q}\rangle = -m w^2(t)\,\langle q\rangle^2 + \frac{\langle p\rangle^2}{m}. \end{aligned} \tag{51}$$

In the HO-case we have a divergenceless flow so that (37) applies, yielding

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^M \lambda_j \frac{\partial}{\partial\langle A_i\rangle} \frac{d\langle A_j\rangle}{dt}, \tag{52}$$

from which we find:

$$\frac{d\lambda_p}{dt} = -\frac{\lambda_q}{m}, \tag{53}$$

$$\frac{d\lambda_q}{dt} = \lambda_p m w^2(t), \tag{54}$$

$$\frac{d\lambda_{p^2}}{dt} = -\frac{\lambda_{pq}}{m}, \tag{55}$$

$$\frac{d\lambda_{q^2}}{dt} = \lambda_{pq} m w^2(t), \tag{56}$$

and

$$\frac{d\lambda_{pq}}{dt} = \lambda_{p^2} 2mw^2(t) - \frac{\lambda_{q^2} 2}{m}. \tag{57}$$

The system of linear differential equations (53-57) 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 MaxEnt ansatz (remember that all the time dependence of $f(q,p,t)$ is through the Lagrange multipliers λ_i)

$$f(q,p,t) = \frac{1}{Z} \exp(-\lambda_q q - \lambda_p p - \lambda_{q^2} q^2 - \lambda_{qp}\, qp - \lambda_{p^2} p^2), \tag{58}$$

with

$$Z = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp(-\lambda_q q - \lambda_p p - \lambda_{q^2} q^2 - \lambda_{qp} qp - \lambda_{p^2} p^2) dq dp, \tag{59}$$

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

6. Conclusions

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

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

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

3. MaxEnt 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 to several evolution equations, such as the Liouville, diffusion, and Fokker-Planck equations.

4. MaxEnt is today an indispensable tool in Physics, Chemistry, Engineering, etc., for confronting "real world" problems.

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

1. J. Willard Gibbs, *Elementary Principles in Statistical Mechanics* (New Haven, Yale University Press, 1902).
2. E.T. Jaynes, *Phys. Rev.* **106** (1957) 620; **108** 171; *Papers on probability, statistics and statistical physics*, edited by R.D. Rosenkrantz (Dordrecht, Reidel, 1987).
3. M. Tribus, *Thermostatics and Thermodynamics* (D. Van Nostrand Company, New York, 1961).
4. A. Katz, *Principles of Statistical Mechanics* (W.H. Freeman and Company, San Francisco and London, 1967).
5. G.A.P. Wyllie, *Elementary Statistical Mechanics* (Hutchinson, London, 1970).
6. R. Baierlein, *Atoms and Information Theory* (W.H. Freeman and Company, San Francisco, 1971).
7. L. Brillouin, *Science and Information Theory* (New York: Academic Press, New York, 1962).
8. Deepak Agrawal, J.K. Singh, and Akhilesh Kumar, *Biosystems Engineering* **90** (2005) 103.
9. A. Lukacs and E. Papp, *Journal of Photochemistry and Photobiology B: Biology* **77** (2004) 1.
10. A.M. Blokhin, R.S. Bushmanov, and V. Romano, *International Journal of Engineering Science* **42** (2004) 915.
11. Yihong Gong, Mei Han, Wei Hua, and Wei Xu, *Computer Vision and Image Understanding* **96** (2004) 181.
12. Ali A. Shams, Sergio Picozzi, and F. Bary Malik, *Physica B: Condensed Matter* **352** (2004) 269.
13. Kouichirou Amemiya, Shigenori Moriwaki, and Norikatsu Mio, *Physics Letters A* **315** (2003) 184.
14. S. Israel, R. Saravanan, N. Srinivasan, and R. K. Rajaram, *Journal of Physics and Chemistry of Solids* **64** (2003) 43.
15. Sun Kyoung Kim and Woo Il Lee, *International Journal of Heat and Mass Transfer* **45** (2002) 381.
16. Jonathan Clowsera and Costas Strouthosa, *Nuclear Physics B - Proceedings Supplements* **106-107** (2002) 489.
17. A. Elgarayhi, *Journal of Quantitative Spectroscopy and Radiative Transfer* **75** (2002) 1.
18. Soumya Raychaudhuri, Jeffrey T. Chang, Patrick D. Sutphin, and Russ B. Altman, *Genome Research* **12** (2002) 203.
19. S.A. El-Wakil, A. Elhanbaly, and M. A. Abdou, *Journal of Quantitative Spectroscopy and Radiative Transfer* **69** (2001) 41.
20. H.B. Callen, *Thermodynamics* (NY, J. Wiley, 1960); E.A. Desloge, *Thermal Physics* (NY, Holt, Rhinehart and Winston, 1968).

21. R.K. Pathria, *Statistical Mechanics* (Exeter, Pergamon Press, 1993).
22. F. Reif, *Statistical and Thermal Physics* (NY, McGraw-Hill, 1965).
23. J.J. Sakurai, *Modern Quantum Mechanics* (Menlo Park, Ca., Benjamin, 1985).
24. A. Katz, *Principles of Statistical Mechanics: The Information Theory Approach* (San Francisco, Freeman and Co., 1967).
25. D.J. Scalapino, in *Physics and Probability. Essays in Honor of Edwin T. Jaynes*, edited by W.T. Grandy Jr. and P.W. Milonni (NY, Cambridge University Press, 1993).
26. A.R. Plastino and A. Plastino, *Phys. Lett. A* **226** (1997) 257.
27. E. Curado and A. Plastino, *Phys. Rev. E* (2005) (in press).
28. A.R. Plastino, *Tsallis Theory, the Maximum Entropy Principle, and Evolutions Equations*, p. 157, in *NonExtensive Statistical Mechanics and Its Applications*, *Lecture Notes in Physics*, S. Abe and Y. Okamoto (Eds.) (Springer Verlag, Berlin, 2001).
29. E.D. Malaza, H.G. Miller, A.R. Plastino, and F. Solms, *Physica A* **265** (1999) 224.
30. A.R. Plastino and A. Plastino, *Physica A* **258** (1998) 429.
31. A.R. Plastino, H.G. Miller, and A. Plastino, *Phys. Rev. E* **56** (1997) 3927.
32. T.D. Franck, *Nonlinear Fokker-Planck Equations* (Springer-Verlag, Berlin, 2005).