Article

Maximum Entropy Production as an Inference Algorithm that Translates Physical Assumptions into Macroscopic Predictions: Don't Shoot the Messenger

Roderick C. Dewar

Division of Plant Sciences, Research School of Biology, R.N. Robertson Building, The Australian National University, Canberra ACT 0200, Australia; E-Mail: roderick.dewar@anu.edu.au

Abstract: Is Maximum Entropy Production (MEP) a physical principle? In this paper I tentatively suggest it is not, on the basis that MEP is equivalent to Jaynes' Maximum Entropy (MaxEnt) inference algorithm that passively translates physical assumptions into macroscopic predictions, as applied to non-equilibrium systems. MaxEnt itself has no physical content; disagreement between MaxEnt predictions and experiment falsifies the physical assumptions, not MaxEnt. While it remains to be shown rigorously that MEP is indeed equivalent to MaxEnt for systems arbitrarily far from equilibrium, work in progress tentatively supports this conclusion. In terms of its role within non-equilibrium statistical mechanics, MEP might then be better understood as Messenger of Essential Physics.

Keywords: maximum entropy; maximum entropy production; non-equilibrium; statistical inference; statistical mechanics

1. Introduction

What is the current status of Maximum Entropy Production (MEP)? Is it a working hypothesis about the macroscopic behaviour of non-equilibrium systems whose domain of application has yet to be firmly established? Or is it an accepted physical principle with a well-defined range of validity? Or has MEP already been shown to be fundamentally wrong? If it is a working hypothesis, what further critical tests would help to define its range of validity? If it is an established principle with a well-defined range of validity, what is its physical basis? And if it is just plain wrong, what if anything can be salvaged from the wreckage? Much of the discussion at the 2009 and previous MEP workshops,

and in the literature [1–7], has been framed around these fundamental questions, motivated by the need to underpin and guide the practical application of MEP to non-equilibrium phenomena across physics, chemistry and biology.

Behind all of these questions, however, lies the tacit assumption that MEP *is* a physical principle, *i.e.*, a statement about the behaviour of physical (or biological) systems that is open to experimental falsification, like Newton's inverse square law of gravitational attraction for example. However, if this assumption is wrong—as I tentatively suggest here—then much of the debate about the current status of MEP is misplaced, and worse, misleading. In that case, to avoid potential confusion, we need to clarify what MEP actually is, what we are doing when we apply it, and what it means when MEP predictions agree or disagree with experiment. This is the objective of the present paper.

To an extent, the view that MEP is a physical principle has been encouraged by the way in which thermodynamics is often presented in text books. The terms *thermodynamic force* and *thermodynamic potential* are suggestive analogues of gravitational force and gravitational potential (e.g., "concentration gradients are the driving force for molecular diffusion"). Thermodynamic entropy production (EP) is expressible as the product of various thermodynamic forces and fluxes. It is then tempting to think of MEP as physical fact, with EP being viewed as a sort of grand thermodynamic potential that governs the dynamics of non-equilibrium systems through the "principle" of MEP, driving the system up the "EP hill" in an (inverted) analogy with gravitational potential [1,7].

However, the statistical nature of thermodynamic laws (e.g., the relationship between molecular diffusion and concentration gradients) means that the analogy between thermodynamic and mechanical forces, *etc.* is just that. The prediction of thermodynamic behaviour by statistical mechanics has (as its name suggests) both statistical and physical ("mechanical") components, and potential confusion can arise if the distinction between the two is not respected, because it is only the latter that are open to experimental falsification [8]. For example, when the Boltzmann-Gibbs (B-G) statistical principle of maximum entropy is applied to equilibrium systems, disagreement between predictions and experiment signals a failure of the physical assumptions or other constraints under which the entropy was maximised, not a failure of the B-G principle itself. This is illustrated by the historical failure of classical radiation theory (which assumes a continuous energy spectrum) to reproduce, through the B-G principle, the observed energy spectrum of blackbody radiation, and Planck's spectacular success in doing so when he replaced the classical assumption by one of discrete energy levels [9]. The B-G statistical principle is not falsifiable because (and not everyone agrees here) it is not a physical principle.

This last statement is a consequence of Jaynes' information-based formulation of statistical mechanics [10–12] (which even today is not universally accepted). Jaynes' formulation interprets the B-G principle as an example of a more general algorithm of inference from partial information (the Maximum Entropy algorithm, or MaxEnt), whose scope extends beyond thermodynamics to other fields (e.g., image reconstruction, spectral analysis, inverse problems, economics) [13]. In its application to statistical mechanics, Jaynes' formulation clearly distinguishes between the physical information that is given or assumed (e.g., continuous *vs.* discrete energy levels, conservation laws, experimental constraints) and the inference algorithm used to make macroscopic predictions from that information (the B-G principle, *i.e.*, MaxEnt). The key point here (explained in Section 2.2) is that MaxEnt passively translates physical information into macroscopic predictions without introducing

any additional physical assumptions itself. Physically, MaxEnt is mute; it is an inference algorithm, not a physical principle—the messenger, not the message. Moreover, since MaxEnt can be extended to non-equilibrium behaviour (same algorithm, different input information), the same point applies there too.

Figure 1. Two contrasting interpretations of MEP. (a) MEP as an inference algorithm (MaxEnt) that converts physical assumptions into macroscopic predictions, with no physical content itself. Only the physical assumptions are falsifiable by observations (dotted arrows). (b) MEP as a physical principle, *i.e.*, with physical content over and above the physical assumptions in (a). Both the MaxEnt physical assumptions and MEP are open to experimental falsification.
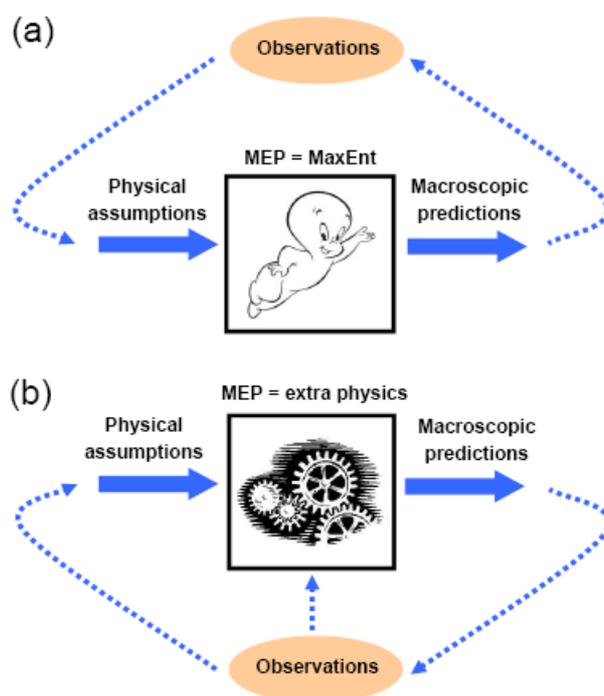
If it can be shown that MEP is equivalent to MaxEnt when applied to non-equilibrium systems (and there lies a challenge), it follows that MEP is not a physical principle after all but, rather, a passive algorithm for translating given or assumed physical information about non-equilibrium systems into macroscopic predictions (Figure 1a). In that case, when MEP predictions disagree with experiment, it is the physical information (assumptions, constraints) that is falsified, not MEP (hence the subtitle "Don't shoot the messenger"). In contrast, if MEP contains extra physical assumptions beyond those explicitly identified in the problem under investigation, then MEP itself is open to experimental falsification (Figure 1b).

The paper is structured as follows. Section 2 outlines the general programme of statistical mechanics, the aim of which is to identify the *essential physics* governing macroscopic phenomena; discusses the status of MaxEnt as an inference algorithm within this general programme, according to Jaynes' interpretation; and addresses the question of whether MaxEnt is restricted to the large-N limit, where N is the number of microscopic degrees of freedom. The discussion here provides the necessary

information-based perspective from which MEP will be subsequently viewed. Section 3 briefly summarises recent and on-going efforts to clarify the relationship between MaxEnt and MEP within non-equilibrium statistical mechanics, and re-examines some previously-raised issues concerning the practical application of MEP in the new light of its interpretation as an inference algorithm. Section 4 presents some tentative conclusions, and suggests that the acronym MEP might be re-interpreted as Messenger of Essential Physics.

2. MaxEnt as Messenger of Essential Physics

2.1. The problem of macroscopic prediction

The practical problem we are trying to address is how to predict the macroscopic behaviour of non-equilibrium physical systems (e.g., the large-scale structure of turbulent heat flow within the Earth's atmosphere and oceans, or the mean fluid velocity and temperature fields in a Rayleigh-Bénard cell). One approach would be to try to calculate the microscopic trajectory of the system directly, by (numerically) integrating the microscopic equations of motion forward in time from some initial microstate, typically requiring some approximate computational scheme involving discretization in space and time. Practical challenges here include the treatment of sub-grid scale dynamics, symmetry-breaking of the equations of motion by the computational scheme [14], and the rapid increase in computational costs with the number of microscopic degrees of freedom N. Therefore one might ask whether there is not an alternative approach.

Nature herself suggests that there is: at least for large N, macroscopic behaviour can be accurately reproduced experimentally through the control of a relatively small number of macroscopic degrees of freedom, implying that the initial microstate—which is not under experimental control—must be largely irrelevant to the experimental result. This remarkable degree of *microscopic redundancy* has a simple explanation—it must be that the overwhelming majority of possible initial microstates lead to microscopic trajectories that look essentially the same on macroscopic scales [10,11]. Hence, as far as macroscopic behaviour is concerned, it does not matter in which one of those initial microstates the system finds itself. Given this remarkable empirical fact, there ought to be a more efficient way to predict macroscopic behaviour directly from a tractably reduced set of physical assumptions and constraints—what one might call the *essential physics* governing the macroscopic phenomena under investigation. The essential physics should include not only the relevant experimental control parameters and other macroscopic constraints that are relatively straightforward to identify, but also certain general features of the microscopic equations of motion (e.g., conservation laws, the energy spectrum of initial states) which manifest themselves at the macroscopic scale, and which might be the subject of conjecture as in the historical example of blackbody radiation. The essential physics is not universal but will depend on the macroscopic phenomenon under investigation, including its spatial or temporal resolution. Two practical challenges, then, are how to identify the essential physics in any given problem, and how to make macroscopic predictions from it.

2.2. MaxEnt: an inference algorithm with no physical content

We may view Jaynes' information-based formulation of statistical mechanics as a response to these two challenges, within which the MaxEnt algorithm plays a central role. Broadly speaking, statistical mechanics takes an informed guess as to the essential physics and applies MaxEnt to make macroscopic predictions from that guess. Disagreement between prediction and experiment informs a new guess and so on until, by trial and error, acceptable agreement is reached.

Operationally, the MaxEnt algorithm is very simple, although its interpretation remains a subject of controversy. As a concrete and familiar example, consider the application of MaxEnt to a closed system in thermal equilibrium with a heat bath at temperature *T*. By closed we mean that no matter enters or leaves the system, but energy can be transferred between the system and the heat bath. In equilibrium, only the mean energy *U* of the system is fixed; the actual energy *E* of the system fluctuates about *U*. Let E_i denote the system energy in microstate *i*. Here the given physical information consists of *U*, plus knowledge of (or assumptions about) the spectrum of possible microstates *i* and their energies E_i. Given this information (or guess), MaxEnt makes macroscopic predictions by first assigning a probability p_i that the system will be in microstate *i* and then taking expectation values over p_i. In its simplest version [12], MaxEnt constructs p_i by maximising the Shannon information entropy:

$$H = -\sum_i p_i \ln p_i \tag{1}$$

with respect to the p_i, subject to the mean energy constraint:

$$\sum_i p_i E_i = U \tag{2}$$

and the normalisation constraint:

$$\sum_i p_i = 1 \tag{3}$$

The solution to this constrained optimisation problem is:

$$p_i = \frac{e^{-\lambda E_i}}{Z(\lambda)} \tag{4}$$

where $Z(\lambda) = \sum_i e^{-\lambda E_i}$ is a normalisation factor (partition function) and λ is a Lagrange multiplier associated with the mean energy constraint. The value of λ is obtained as a function of *U* by solving Equation 2, which is equivalent to:

$$U = -\frac{\partial \ln Z(\lambda)}{\partial \lambda} \tag{5}$$

The value of the maximised Shannon entropy is then:

$$H_{\max}(U) = \ln Z(\lambda) + \lambda U \tag{6}$$

which is a function of *U* alone since $\lambda = \lambda(U)$ *via* Equation 5. Contact with traditional thermodynamics is established through the identifications $kT = 1/\lambda$, $S = kH_{\max}$ and $A = -kT \ln Z$, where *k* is Boltzmann's constant, *T* is the temperature, *S* is the thermodynamic entropy, and *A* is the

Helmholtz free energy. Then Equation 4 is Gibbs' canonical distribution and Equation 6 is the standard thermodynamic relation $A = U - TS$. Once we have constructed the MaxEnt distribution p_i, the prediction of any observable x can be obtained from the expectation value $X = \sum_i p_i x_i$ (where x_i = value of x in state i). Actually we can construct the entire probability distribution $p(x)$, from $p(x) = \sum_i p_i \delta(x - x_i)$ (where $\delta(x)$ is the Dirac delta function), and therefore MaxEnt predicts statistical fluctuations about the mean, not just the mean value X itself.

What then is MaxEnt? What are we doing when we apply it? And what does it mean when MaxEnt predictions agree or disagree with experiment? In Jaynes' interpretation, MaxEnt is a passive inference algorithm that converts given information into predictions; it has no physical content itself (Figure 1a). This point may be understood as follows. In information terms, H represents the missing information (or uncertainty) about which microstate the system is in [12]. By maximising H subject to the given physical information, that information and no more is encoded into p_i. Any other distribution p_i' reflects more information (less uncertainty) than is actually given. Therefore, the statistical distribution $p(x)$ of any observable x, constructed as above from the MaxEnt microstate distribution p_i, faithfully reflects the given physical information—no more, no less. Then, when the predicted $p(x)$ differs from the experimental one, it means that the given or assumed physical information is erroneous or incomplete relative to the essential physics governing the statistical behaviour of x. Conversely, when MaxEnt predictions of $p(x)$ agree with experiment, it means that the given information captures the essential physics governing the statistical behaviour of x, all other information being irrelevant.

Experimental falsification of MaxEnt is thus a meaningless concept because, in Jaynes' interpretation, MaxEnt is not a physical principle. When MaxEnt predictions disagree with experiment, it is the assumed essential physics (the message) that is falsified, not MaxEnt (the messenger). Jaynes' interpretation also makes it clear that MaxEnt can be applied to any system, equilibrium or non-equilibrium, physical, biological or otherwise. In all cases, MaxEnt makes no claim as to the reality of its predictions; its role within the general programme of statistical mechanics is to ensure that the assumed essential physics (and no other physical assumptions) is faithfully represented in the predictions that are compared with experiment. MaxEnt thus plays a central role in the identification of the actual essential physics, which would otherwise be obscured by the implicit introduction of extra physical assumptions.

Nevertheless, one may still question whether MaxEnt is the appropriate algorithm to use in the first place. For example, maximum *relative* entropy (MaxREnt) is the appropriate generalisation of MaxEnt when the microstates are not *a priori* equiprobable (e.g., [15]), while Maximum Probability (e.g., [16]) is a combinatorial approach to statistical mechanics which differs significantly from MaxEnt (or MaxREnt) when N is small (Section 2.3). However, that question falls within the realm of statistical inference, not physics.

2.3. Does MaxEnt require N to be large?

This question arises at both a theoretical and a practical level. As an alternative to Jaynes' information-based formulation of statistical mechanics, the Maximum Probability (MaxProb) formulation (e.g., [16]) is a combinatorial approach which takes as its basic concept the number of ways W that a given macrostate can be realised microscopically. For example, the number of ways in

which N distinguishable entities can be assigned to M distinguishable states such that there are n_i entities in state i is given by the multinomial coefficient:

$$W = \frac{N!}{n_1! n_2! .. n_M!} \qquad (7)$$

The macrostate $\{n_i\}$ predicted by MaxProb is the one for which W is maximal, subject to given physical constraints. MaxProb is just Boltzmann's state-counting principle [17] extended to all values of N (not just large N); in contrast, MaxEnt (without its information-based interpretation) is the statistical principle adopted by Gibbs [18], as illustrated in Section 2.2 The mathematical connection between MaxProb and MaxEnt occurs in the limit of large N and large n_i, for then Stirling's approximation ($\ln N! \approx N \ln N$ and $\ln n_i! \approx n_i \ln n_i$) yields the asymptotic result:

$$\frac{1}{N} \ln W \to -\sum_{i=1}^{M} p_i \ln p_i = H \qquad (N, n_i \to \infty) \qquad (8)$$

where the state probability p_i is defined as $p_i = n_i / N$, *i.e.* the occupation frequency of state i. Within the MaxProb formulation of statistical mechanics, therefore, MaxEnt is an approximation to MaxProb that is only valid for large N (and large n_i), and one must apply corrections to MaxEnt when N is small (or when statistics other than multinomial are assumed) [19,20]. Conversely, if one accepts MaxEnt as the fundamental basis of statistical mechanics, it is the validity of MaxProb that is restricted to large N. MaxProb and MaxEnt also differ conceptually in their interpretation of p_i. In MaxProb, p_i is a physical frequency (a property of the real world) while in MaxEnt p_i is a Bayesian probability (describing our objective state of knowledge about the real world based on objectively-prescribed information) [12].

Therefore, if one adopts MaxProb as the fundamental basis of statistical mechanics, there is a theoretical issue regarding the validity of MaxEnt for small N. In this paper, however, I am considering MEP within the framework of Jaynes' information-based formulation of statistical mechanics, in which the fundamental concept is not W (Equation 7) but the Shannon information entropy H (Equation 1). Within Jaynes' formulation, there is nothing intrinsic to the MaxEnt algorithm which demands that N be large; as we have seen, MaxEnt has no physical content.

Nevertheless there remains the practical issue of whether MaxEnt still requires the assumption that N is large as part of the given physical information, if its predictions are ever to agree with experiment. Since MaxEnt predicts the probability distribution $p(x)$ of a physical observable x, not just its mean value $X = \sum_i p_i x_i$, it is pertinent here to consider how N affects the predicted and experimental $p(x)$. If N is known to be large then, as a general rule, MaxEnt predicts that $p(x)$ is sharply peaked about the mean X (the relative standard deviation declining typically as $1/\sqrt{N}$). An exception to this rule occurs at critical phase transitions (e.g., at the Curie temperature of a magnet). Even when N is of the order of Avogadro's number MaxEnt predicts that, at or near the transition point, the so-called macroscopic order parameter x (e.g., the magnetisation) will fluctuate widely about its mean value—$p(x)$ is very broad—and, indeed, the detailed nature of the predicted fluctuations (obtained from renormalisation group analyses of the MaxEnt solution [21]) has been confirmed experimentally (e.g., *via* the divergence of the magnetic susceptibility). The success of MaxEnt here reveals the essential physics governing the critical fluctuations (e.g., for a magnet at the Curie temperature, the thermal energy per

degree of freedom is of the same order as the interaction energy between individual magnetic moments, resulting in a highly sensitive balance between the disordering effect of thermal agitation and the ordering effect of the magnetic interaction).

If *N* is known to be small, then typically MaxEnt will predict large relative fluctuations in all observables—again, *p(x)* is broad. But this is also what one observes experimentally; as *N* decreases, the experimental behaviour becomes less reproducible under given experimental conditions as fluctuations about the mean behaviour become more important (e.g., [22]).

Figure 2. Different scenarios for MaxEnt-predicted (━━━━━) vs. experimental (▪▪▪▪▪▪▪▪▪▪▪) probability distributions *p(x)* of some macroscopic observable *x*. *N* = number of microscopic degrees of freedom. Disagreement between the predicted and observed *p(x)* signals either (a) erroneous or (b) incomplete physics. Agreement between predicted and observed *p(x)* reveals the essential physics governing the statistical behaviour of *x*, whether that behaviour (c) is sharply-peaked about the mean or (d) involves large fluctuations. MaxEnt remains practically useful whether *N* is large or small.
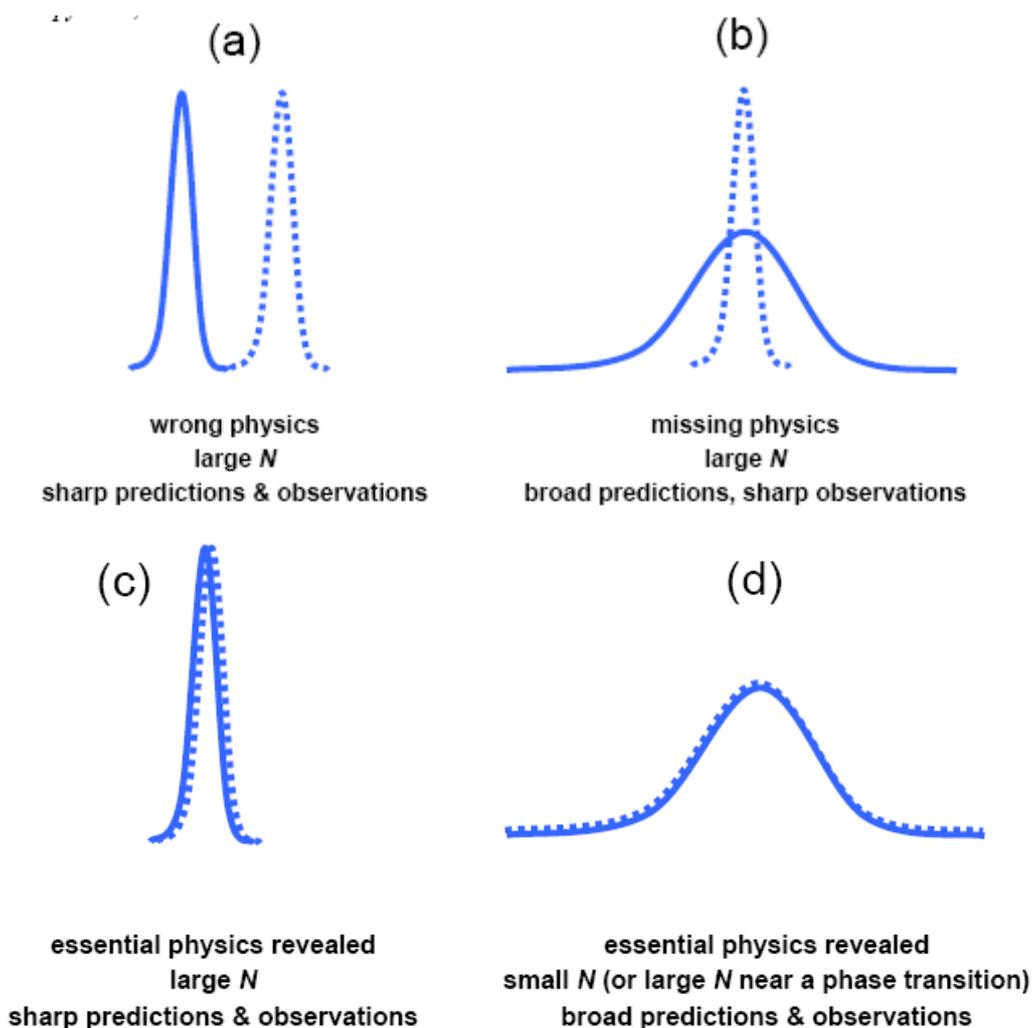
Thus, regardless of whether *N* is large or small, the predicted *p(x)* is the "best" possible one in the sense that it faithfully reflects the given physical information (which includes the size of *N*) and no

more. Figure 2 depicts various scenarios for the relation between the MaxEnt-predicted and experimental $p(x)$. When there is disagreement between the predicted and experimental $p(x)$—whether sharply-peaked or broad—it means that the given or assumed information is either erroneous (if the predicted and observed $p(x)$ are both sharply-peaked but mismatched, Figure 2a), or incomplete (if the predicted $p(x)$ is broad but the observed $p(x)$ is sharply-peaked, Figure 2b). One can also envisage a scenario (not shown) in which the predicted $p(x)$ is narrower than the experimental one, indicating an over-constrained set of physical assumptions. Conversely, agreement between the predicted and experimental $p(x)$—whether sharply-peaked or broad—means that the given or assumed information captures the essential physics governing the statistical behaviour of x (Figures 2c, 2d).

In summary, MaxEnt remains practically useful whether N is large or small, because MaxEnt predicts fluctuations as well as mean values, and fluctuations can be an important aspect of the macroscopic phenomena under investigation, especially for small N. However, MaxEnt is generally more efficient in exposing erroneous physical assumptions when N is large, for then generally both the predicted and observed $p(x)$ are sharply-peaked (Figure 2a).

3. Implications of MaxEnt for MEP

3.1. Is MEP equivalent to MaxEnt?

Here I give a brief status report. In two papers by this author [2,3], three different derivations of MEP from MaxEnt were attempted, all based on maximising the Shannon entropy of p_Γ, the probability distribution of microscopic trajectories Γ, subject to physical constraints. Subsequently, some technical limitations of these three derivations were noted [4–6]. Specifically, the first derivation [2] introduced an *ad hoc* assumption that the Shannon path entropy is an increasing function of EP, so that MaxEnt would imply maximum EP. However, Bruers [5] presented an example in which that assumption is false, implying a minimum—not a maximum—EP. Nevertheless, another well-established result—the so-called Fluctuation Theorem [23]—was also derived from the same formalism independently of the *ad hoc* assumption. In the second derivation [3], a Gaussian approximation for the MaxEnt macrostate probability distribution was used to derive an orthogonality condition for EP (first proposed by Ziegler [24]), from which max-EP was then derived. But as Bruers [5] and Grinstein and Linsker [6] correctly observed, the Gaussian approximation is only valid close to equilibrium. In a third derivation [3, Appendix], an upper bound constraint on EP was introduced, and it was then argued that the most probable macrostate was the one for which the upper bound was attained. However, the specific mathematical argument employed contains some unresolved technical issues [5] and is unconvincing. Nevertheless, the successful derivation of the Fluctuation Theorem and Ziegler's orthogonality relation suggests that the MaxEnt approach to understanding MEP is possibly on the right track.

More recently, Niven [25] has proposed a "conditional derivation" of MEP based on maximising the Shannon entropy defined in terms of macroscopic fluxes rather than microscopic trajectories. However, this derivation appears to be conditional on the key assumption that there is a net increase in EP during the dynamic transition from one macrostate to another, so that steady states are characterised by maximum EP. Unfortunately, until that assumption can be firmly justified, it would

seem that this application of MaxEnt falls short of a fundamental derivation of MEP. It also remains to be seen if and how this approach is related to that based on microscopic trajectories.

In work in progress [26], the above-mentioned technical difficulties of the MaxEnt path entropy approach to MEP may have been overcome. In brief, this work appears to establish the equivalence of MEP and MaxEnt for systems arbitrarily far from equilibrium, whilst retaining previous successes in deriving the Fluctuation Theorem and Ziegler's orthogonality property. The new feature of this work is the introduction of a general information-based measure of macroscopic irreversibility (I)—defined in terms of the probabilities of a microscopic trajectory Γ and the reverse trajectory Γ_R—with the central result that for non-equilibrium systems I attains its maximum value permitted by the applied constraints; I is identified with the thermodynamic EP. The only physical assumptions used to derive MEP from MaxEnt in this way appear to be the standard conservation laws (continuity equations) of mass, energy, momentum, electric charge, *etc*. Since these laws constitute a standard and unassailable component of the physical information already assumed as input to MaxEnt, we are in the situation depicted in Figure 1a. Details will be published elsewhere.

In summary, work in progress tentatively suggests that MEP is the expression of MaxEnt applied to physical systems arbitrarily far from equilibrium (for which the standard conservation laws remain valid). If this is indeed the case, then MEP is free of additional physical assumptions beyond those already used as input to MaxEnt, and is therefore an inference algorithm rather than a physical principle (Figure 1a).

3.2. What if the sea were made of vinegar?

This question arose during discussions at the 2009 MEP workshop in relation to the fact that MEP predictions of the large-scale structure of heat transport in Earth's climate obtained from Paltridge's 10-zone climate model do not depend on the physical nature of the transport medium [27]. A similar question was raised much earlier in relation to the model's predicted independence of planetary-scale heat flows on planetary rotation rate [28]. Would Earth's climate really be the same if the sea were made of vinegar or the Earth spun faster? Surely the answer is no (so it is claimed) and therefore MEP is fundamentally wrong. The implication would seem to be that, in this case, the broad agreement between MEP predictions and observations is a lucky accident.

But let us suppose that MEP indeed passively translates physical assumptions into macroscopic predictions (Figure 1a). In Paltridge's application of MEP to Earth's climate [27], the physical assumptions consist of steady-state energy balance, plus certain subsidiary model assumptions (e.g., regarding the relationship between atmospheric and surface temperatures) and parameterisations (e.g., mean surface albedos in each latitudinal zone). Then the rational conclusion to be drawn from the broad agreement between MEP predictions and observations is that the physical assumptions capture the essential physics governing the large-scale patterns of heat transport, and that all other physical details are irrelevant at that scale. The physical explanation would be that, given a sufficient number of microscopic degrees of freedom (especially assured at large scales), the detailed processes responsible for heat transport may well differ in different physical media or under different planetary rotation rates, but the same large-scale pattern of heat transport will result—a case of microscopic redundancy. Conversely, an example of MEP as a messenger of erroneous physics is provided by Paltridge's recent

application of MEP to climate change [29], in which unrealistic predictions were obtained when the water-vapour feedback derived from General Circulation Models formed part of the physical assumptions [see also Paltridge, this volume].

3.3. MEP and the Upper Bound Theory of fluid turbulence

The apparent success of MEP in revealing the essential physics governing planetary-scale heat flows [27,30] does not imply that the physics ignored there is not relevant in other problems. Clearly, the relevant physics will depend on the scale of the macroscopic phenomena under investigation, with phenomena at higher spatial resolution depending on more detailed physical information. Empirically, the fluid velocity fields in turbulent shear flow or thermal (Rayleigh-Bénard) convection, for example, do depend on the physical nature of the fluid (e.g., viscosity, thermal expansion coefficient) and its rotational forcing. But again, MEP may be used to reveal the essential physics at that scale. This is illustrated by the Upper Bound Theory (UBT) of fluid turbulence [31], a collection of variational principles identical or closely-related to MEP [32].

Several features of the empirical velocity fields in turbulent shear flow and thermal convection have been predicted—with varying degrees of success—by maximising various dissipation-like functionals of the flow under a restricted number of dynamical constraints derived from the Navier-Stokes equation (e.g., global energy balance, mean momentum balance) plus a small number of additional assumptions and constraints (e.g., incompressibility, boundary conditions). The search for the relevant constraints using UBT is exactly in the spirit of the search for the essential physics using MEP. Moreover, in UBT it remains an open question which dissipation-like functional is the most appropriate to maximise, analogous to the question of which expression for EP to use in applications of MEP. Maximum dissipation rate of the mean flow was derived as a statistical stability criterion for turbulent flow [33,34]. However, in UBT no single choice of functional to be maximised appears to work in all cases. For example, the dissipation rate of the mean flow appears to be the only candidate functional capable of predicting zero interior mean shear in plane Couette flow, but, conversely, this choice cannot be correct for plane Poiseuille flow where the interior shear does not vanish [31]. As Kerswell [31] concluded:

"The challenge therefore remains to find a [universal] functional whose optimization over a tractably reduced set of dynamical constraints leads to the emergence of realistic optimal velocity fields. Unfortunately, it remains unclear how to construct such a functional beyond intelligent guessing".

The close relationship between UBT and MEP suggests the possibility that such a universal functional might be Shannon information entropy, and that UBT, like MEP, may ultimately be an expression of MaxEnt. If this is indeed the case, then one might anticipate that previous [2–3,25] and on-going [26] efforts to derive MEP from MaxEnt will eventually lead to an unambiguous identification of the appropriate dissipation-like functional to be used in any given problem.

3.4. So what?

If MEP is indeed equivalent to MaxEnt and has no physical content itself (Figure 1a), then MEP can predict nothing about macroscopic phenomena beyond what is already predictable (in principal) from the microsopic equations of motions. So what has been gained? The answer is computational efficiency. Rather than simulate the microscopic physics in detail, only to discover subsequently that most of that detail is irrelevant to macroscopic behaviour, MEP seeks from the outset to identify a tractably reduced set of physical assumptions that gives the same result. Identification of the essential physics in this way might only be achieved through trial and error. Nevertheless, when microscopic redundancy is effective enough that the essential physics *is* tractably reduced, the trial and error process is very efficient indeed—as the first 150 years of statistical mechanics have taught us.

4. Conclusions

In this paper I have suggested that MEP is not a physical principle but, rather, an inference algorithm (MaxEnt) that passively translates physical assumptions into macroscopic predictions (Figure 1a). In that case, falsification of MEP is a meaningless concept. Only the physical assumptions are falsifiable. This suggestion redefines the status of MEP within the general programme of non-equilibrium statistical mechanics, the aim of which is to identify the essential physics governing non-equilibrium macroscopic phenomena. We should pick the MEP state not because the real system is driven there physically, but because only the MEP state faithfully reflects the physical assumptions we have adopted as our current guess for the essential physics. Picking a non-MEP state is unwarranted because it is tantamount to introducing extra physical assumptions beyond that guess, thus obscuring the identification of the actual essential physics when we make comparisons with observations.

If this suggestion is correct, then previous statements by various authors (including this one) regarding "the evidence or otherwise for MEP" are misleading, and ought to be refocused on "the evidence or otherwise that the physical assumptions capture the essential physics governing the macroscopic phenomena in question." It might then be more informative to reinterpret the acronym MEP as Messenger of Essential Physics or Messenger of Erroneous Physics, according to whether its predictions agree or disagree with observations.

However, it remains to be seen whether indeed MEP is a faithful expression of MaxEnt, particularly for systems that are far from equilibrium, although work in progress tentatively supports that view [26]. If MEP is not a faithful expression of MaxEnt, the implication is that MEP itself introduces some additional physical assumptions beyond those explicitly identified in the problem (Figure 1b); then, only when those additional assumptions are valid (or invalid but irrelevant) would MEP be practically useful as a Messenger of Essential Physics. Ultimately, however, it is to MaxEnt that we should look in providing the fundamental inference algorithm for non-equilibrium statistical mechanics.

Acknowledgements

References and Notes

1. Ozawa, H.; Ohmura, A.; Lorenz, R.D.; Pujol, T. The second law of thermodynamics and the global climate system: A review of the maximum entropy production principle. *Rev. Geophys.* **2003**, *41*, 1018-1042.

2. Dewar, R.C. Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *J. Phys. A: Math. Gen.* **2003**, *36*, 631-641.

3. Dewar, R.C. Maximum entropy production and the fluctuation theorem. *J. Phys. A: Math. Gen.* **2005**, *38*, L371-L381.

4. Martyushev, L.M.; Seleznev, V.D. Maximum entropy production principle in physics, chemistry and biology. *Phys. Rep.* **2006**, *426*, 1-45.

5. Bruers, S. A discussion on maximum entropy production and information theory. *J. Phys. A: Math. Theor.* **2007**, *40*, 7441-7450.

6. Grinstein, G.; Linsker, R. Comments on a derivation and application of the "maximum entropy production" principle. *J. Phys. A: Math. Theor.* **2007**, *40*, 9717-9720.

7. Kleidon, A. Nonequilibrium thermodynamics and maximum entropy production in the Earth system. *Naturwissenschaften* **2009**, *96*, 653-677.

8. Caticha [*From Inference to Physics*, arXiv:0808.1260, **2008**] has suggested that Newtonian mechanics itself is open to a statistical mechanical interpretation. This suggestion highlights the fact that in statistical mechanics what one considers to be a *macroscopic* phenomenon depends on the scale of the problem. Only by historical accident was the "mechanics" in statistical mechanics first taken to be Newtonian; the statistical interpretation of Newtonian mechanics itself as a macroscopic phenomenon posits a "sub-Newtonian mechanics" at even smaller scales. But none of this alters the point (to be made in the following paragraph) that, in any given problem, statistical mechanics distinguishes between the physical information that is given or assumed (both the macroscopic constraints and the assumed microscopic mechanics) and the statistical algorithm (MaxEnt) used to make macroscopic predictions from it.

9. Longair, M.S. *Theoretical Concepts in Physics*; Cambridge University Press: Cambridge, UK, 1984; pp 188-225.

10. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620-630.

11. Jaynes, E.T. Information theory and statistical mechanics. II. *Phys. Rev.* **1957**, *108*, 171-190.

12. Jaynes, E.T. *Probability Theory: The Logic of Science*; Bretthorst, G.L., Ed.; Cambridge University Press: Cambridge, UK, 2003.

13. Wu, N. *The Maximum Entropy Method*; Springer: Berlin, Germany, 1997.

14. Essex, C.; Ilie, S.; Corless, R.M. Broken symmetry and long-term forecasting. *J. Geophys. Res.* **2007**, *112*, 1-9.

15. Dewar, R.C.; Porté, A. Statistical mechanics unifies different ecological patterns. *J. Theor. Biol.* **2008**, *251*, 389-403.

16. Niven, R.K. Origins of the combinatorial basis of entropy. In *MaxEnt07*; Knuth, K.H., Caticha, A., Center, J.L., Giffon, A., Rodriguez, C.C., Eds.; AIP: Melville, NY, USA, 2007; pp 133-142.

17. Boltzmann, L. *Lectures on Gas Theory*, Dover: New York, NY, USA, 1995.

18. Gibbs, J.W. (1902) *Elementary Principles of Statistical Mechanics*. Reprinted by Ox Bow Press: Woodridge, CT, USA, 1981.

19. Niven, R.K. Exact Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics. *Phys. Lett. A* **2005**, *342*, 286-293.

20. Niven, R.K.; Grendar, M. Generalized classical, quantum and intermediate statistics and the Polya urn model. *Phys. Lett.* **2009**, *373*, 621-626.

21. Amit, D.J.; Martin-Mayer, V. *Field Theory, The Renormalization Group, and Critical Phenomena*, 3rd ed.; World Scientific Publishing: Singapore, 2005.

22. Wang, G.M.; Sevick, E.M.; Mittag, E.; Searles, D.J.; Evans, D.J. Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales. *Phys. Rev. Lett.* **2002**, *89*, 050601.

23. Evans, D.J.; Searles, D.J. The fluctuation theorem. *Adv. Phys.* **2002**, *51*, 1529-1585.

24. Ziegler, H. *An Introduction to Thermomechanics*; North-Holland: Amsterdam, The Netherlands, 1983.

25. Niven, R.K. Steady state of a dissipative flow-controlled system and the maximum entropy production principle. *Phys. Rev. E* **2009**, *80*, 021113:1-021113:15.

26. Dewar, R.C.; Maritan, A. Work in progress, 2009.

27. Paltridge, G.W. The steady-state format of global climate. *Quart. J. R. Met. Soc.* **1978**, *104*, 927-945.

28. Rodgers, C.D. Comments on Paltridge's "Minimum entropy exchange principle". *Quart. J. R. Met. Soc.* **1976**, *102*, 455-457.

29. Paltridge, G.W.; Farquhar, G.D.; Cuntz, M. Maximum entropy production, cloud feedback, and climate change. *Geophys. Res. Lett.* **2007** *34*, L14708.

30. Lorenz, R.D.; Lunine, J.I.; Withers, P.G.; McKay, C.P. Titan, Mars and Earth: Entropy production by latitudinal heat transport. *Geophys. Res. Lett.* **2001**, *28*, 415-418.

31. Kerswell, R.R. Upper bounds on general dissipation functionals in turbulent shear flows: revisiting the "efficiency" functional. *J. Fluid. Mech.* **2002**, *461*, 239-275.

32. Ozawa, H.; Shimokawa, S.; Sakuma, H. Thermodynamics of fluid turbulence: A unified approach to the maximum transport properties. *Phys. Rev. E.* **2001**, *64*, 026303.

33. Malkus, W.V.R. Statistical stability criteria for turbulent flow. *Phys. Fluids* **1996**, *8*, 1582-1587.

34. Malkus, W.V.R. Borders of disorder: in turbulent channel flow. *J. Fluid Mech.* **2003**, *489*, 185-198.