COMMENT

Comments on a derivation and application of the "maximum entropy production" principle

G Grinstein and R Linsker

IBM Research Division, TJ Watson Research Center, PO Box 218, Yorktown
Heights, NY 10598, USA

E-mail: `ggrin@us.ibm.com`, `linsker@us.ibm.com`

Abstract. We show that: (1) an error invalidates the derivation (Dewar R C 2005
J. Phys. A: Math. Gen. **38** L371) of the "maximum entropy production" (MaxEP)
principle for systems far from equilibrium, for which the constitutive relations are
nonlinear; and (2) the claim (Dewar R 2003 *J. Phys. A: Math. Gen.* **36** 631) that
the phenomenon of "self-organized criticality" is a consequence of MaxEP for slowly
driven systems is unjustified.

In a pair of influential papers, Dewar [1,2] has claimed that a principle of "maximum entropy production" (MaxEP) follows from Jaynes' "maximum entropy" (MaxEnt) principle under certain conditions; that MaxEP holds even for systems far from equilibrium, for which the constitutive relations are nonlinear; and that the phenomenon of "self-organized criticality" (SOC) is a consequence of MaxEP for slowly driven systems.

In this note we first point out an error in the derivation [2, pp. L372-75] that invalidates the claimed proof of MaxEP for far-from-equilibrium systems. Second, we show that the claimed link [1, pp. 639-40] between MaxEP and SOC is unjustified.

1. Systems far from equilibrium

Using the same assumptions and notation as Dewar [2], we have verified his results through equation (13), the "fluctuation theorem" (FT): $p(\boldsymbol{f})/p(-\boldsymbol{f}) = \exp(2\sum_k \lambda_k f_k)$. Here $p(\boldsymbol{f})$ is the p.d.f. of the m-component vector $\boldsymbol{f} \equiv (f_1, ..., f_m)$, where the f_i's are functions whose known average values are $(F_1, ..., F_m)$, and λ_i is the Lagrange multiplier associated with F_i in the maximization of the information entropy in the MaxEnt approach. Dewar notes that "the generic FT [2, equation (12)] has important implications for the functional form of the relationship between $\boldsymbol{\lambda}$ and $\boldsymbol{F}$." He then states, correctly, that the quadratic approximation to $p(\boldsymbol{f})$ in the vicinity of $\boldsymbol{f} = \boldsymbol{F}$ is given by his equation (14):

$$p(\boldsymbol{f}) \propto \exp\{-\frac{1}{2}\sum_{j,k=1}^{m}(f_j - F_j)A_{jk}(f_k - F_k)\} \quad , \tag{1}$$

However, there is no reason to expect, as Dewar requires in the line following equation (14), that that equation be valid for all $\boldsymbol{f}$ including $\boldsymbol{f}$ far from $\boldsymbol{F}$; and therefore there is no basis for concluding (equation (15)) that $\lambda_k = \sum_j A_{jk}(\boldsymbol{F})F_j$. Indeed, if it were true, equation (15) would imply [using also equation (10) for the definition of $A_{nk}(\boldsymbol{F})$] that

$$A_{nk}(\boldsymbol{F}) \equiv \frac{\partial \lambda_k(\boldsymbol{F})}{\partial F_n} = \sum_j \frac{\partial A_{jk}(\boldsymbol{F})}{\partial F_n}F_j + A_{nk}(\boldsymbol{F}) \quad , \tag{2}$$

hence that $\partial A_{jk}(\boldsymbol{F})/\partial F_n = 0$ for arbitrary $\boldsymbol{F}$. But this is only true when A is independent of $\boldsymbol{F}$ (i.e., linear constitutive relations), contrary to Dewar's claim.

The point is that approximation (14) of ref. [2] (our equation (1)), is typically valid only for $\boldsymbol{f}$ near $\boldsymbol{F}$, while the FT, equation (13), involves p at both $\pm\boldsymbol{f}$. Since $\boldsymbol{f}$ and $-\boldsymbol{f}$ can *both* be near $\boldsymbol{F}$ only when $\boldsymbol{F}$ is near zero, however, equations (13) and (14) hold simultaneously only for small $\boldsymbol{F}$, i.e., in the linear ("near-equilibrium") regime where $A(\boldsymbol{F})$ is independent of $\boldsymbol{F}$.

Defining (as does Dewar) the mean dissipation $D = \langle d \rangle$ where $d \equiv 2\sum_k \lambda_k f_k$, and using the definition of $A_{nk}(\boldsymbol{F})$, we find

$$\frac{\partial D(\boldsymbol{F})}{\partial F_n} = 2\lambda_n + 2\sum_k A_{nk}(\boldsymbol{F})F_k \quad , \tag{3}$$

instead of Dewar's orthogonality condition $\partial D(\boldsymbol{F})/\partial F_n = 4\lambda_n$ [2, equation (18)]. (The dual equation (19) for $D(\boldsymbol{\lambda})$ is likewise found to be incorrect in general.) The orthogonality condition – which states that the vector $\boldsymbol{\lambda}(\boldsymbol{F})$ points along the direction of steepest ascent of $D(\boldsymbol{F})$ [2, p. L374] – is therefore not valid far from equilibrium. (By way of contrast, $\boldsymbol{\lambda}(\boldsymbol{F})$ indeed does point along the direction of steepest descent in the information entropy $S(\boldsymbol{F})$ [2, equation (9)] – a property of MaxEnt.)

The orthogonality condition for D is essential to the subsequent derivations in [2] of the principles of "maximum dissipation" and "minimum dissipation," and hence of the corresponding principles of MaxEP and minimum entropy production, respectively. Contrary to the claim in [2], therefore, the derivations using this orthogonality condition do require "the usual near-equilibrium assumption of linear constitutive relations" [2, p. L375]. Thus the question of the existence of possible extremal principles (and in particular, of MaxEP) that might apply to far-from-equilibrium regimes (having nonlinear constitutive relations) has not been settled by [1,2].

1.1. Illustrative example

To illustrate these points explicitly, we make use of an instructive model proposed by Bruers in a recent preprint [3] that also discusses Dewar's work. We consider a single-spin-chain simplification of that (two-spin-chain) model, make some small changes in definition (corresponding to differences in factors of τ, the number of spins in the chain), present some results that correspond to those in [3], and go on to show how and why Dewar's argument [2] fails. In the single-chain model, the distribution p_i is over all sequences of spins $i \equiv (\sigma_1, \sigma_2, \ldots, \sigma_\tau)$, where each $\sigma_t = \pm c$. We choose the quantity f to be the sum of spins $f(i) \equiv \sigma_1 + \ldots + \sigma_\tau$, and the constraint [2, equation (2)] to be $\langle f(i) \rangle \equiv \sum_i p_i f(i) = F$. [Dewar's vectors $\boldsymbol{\lambda}$, $\boldsymbol{f}$, and $\boldsymbol{F}$, and matrix A, are all scalars here since $m = 1$.] As Dewar requires (equation (11)), the possible sequences i can be grouped into pairs $(i+, i-)$ with respect to which f is antisymmetric, $f(i-) = -f(i+)$, by defining, for each $i+ = (\sigma_1, \ldots, \sigma_\tau)$, $i- \equiv (-\sigma_1, \ldots, -\sigma_\tau)$ [or, alternatively, $i- \equiv (-\sigma_\tau, \ldots, -\sigma_1)$].† In order that $F \to$ constant as τ is increased, c should go as $1/\tau$. We define $c = 1/\tau$ for convenience in what follows.

For this model we obtain, exactly (again as in [3], except for factors of τ, and for one spin chain rather than two): the partition function $Z(\lambda) = [2\cosh(\lambda/\tau)]^\tau$; $F(\lambda) = \partial \ln Z(\lambda)/\partial \lambda = \tanh(\lambda/\tau)$; $\lambda(F) = \tau \mathrm{arctanh} F$; and $A(F) \equiv \partial \lambda(F)/\partial F = \tau/(1 - F^2)$. The FT (Dewar's equation (12)), $p_{i+}/p_{i-} = \exp[2\lambda f(i+)]$, is also exact.

For finite τ, only a discrete set of f values can occur; therefore, to calculate the p.d.f. $p(f)$ of f, we pass to the large-τ limit. Using Stirling's approximation for the

---

† This model can be interpreted (as in [3], except for factors of τ) as one in which a constant macroscopic unit time interval is divided into τ microscopic steps, a microscopic quantity c of material or charge flows from left to right (if $\sigma_t = +c$) or right to left (if $\sigma_t = -c$) at each microscopic time step t, $f(i)$ is the total flux (net flow per macroscopic unit of time) for the microscopic path defined by i, and F is the observed value of f averaged over all microscopic paths.

factorial, and dropping terms of $O(\ln \tau / \tau, 1/\tau)$ on the RHS, we obtain:

$$\frac{1}{\tau} \ln p(f) = \frac{1-f}{2} \ln \frac{1-F}{1-f} + \frac{1+f}{2} \ln \frac{1+F}{1+f} \ . \tag{4}$$

Expanding this about $f = F$, we obtain the quadratic approximation $\frac{1}{\tau} \ln p(f) = -\frac{(f-F)^2}{2(1-F^2)} + O((f-F)^3)$. This corresponds to [2, equation (14)] for f near F, since $A(F) = \tau/(1-F^2)$. However, [2, equation (15)] is violated: the LHS is $\lambda(F) = \tau \mathrm{arctanh} F$, while the RHS is $A(F)F = \tau F/(1-F^2)$. These two quantities are only equal in the limit $F \to 0$, where both reduce to τF. In this limit the constitutive relation between λ and F becomes linear, and the quantity $A(F)$ becomes independent of F. Dewar emphasizes that the constitutive relation (equation (15)) is nonlinear owing to the dependence of A on $\boldsymbol{F}$; however, as we see, equation (15) is only satisfied in the linear limit in which A becomes independent of F.

Bruers [3, p.25] notes that the orthogonality condition for D is only valid for linear systems, contrary to Dewar's [2] claim. However, the error in [2] that invalidates Dewar's claim is not identified in [3], and [3, equation (40)] repeats Dewar's assertion that D is equal to $\langle 2 \sum_k \lambda_k f_k \rangle$ and also to $2 \sum_{j,k} A_{jk}(\boldsymbol{F}) F_j F_k$. As we have seen, the latter assertion is incorrect in the nonlinear (far-from-equilibrium) regime.

2. MaxEP and SOC

We turn now to Dewar's first article[1], in which it is argued that the occurrence of SOC in flux-driven systems follows generally from the principle of MaxEP, in the limit where the driving is infinitely slow with respect to the internal dynamics. This claim has great potential significance, since there is at present no clear, general understanding of the origins of SOC; nor have necessary and sufficient conditions for the occurrence of the phenomenon been delineated. We demonstrate, however, that the arguments about SOC in [1] are flawed, and that no conclusion about the occurrence of the phenomenon in the model considered there can be properly drawn.

To see this, consider the sandpile-like system considered in [1], using the same definitions and notation. In particular, F is the output flux of grains from the pile, F_{ext} the fixed external input of grains, and $p(F|F_{\mathrm{ext}})$ the output-flux probability distribution for given F_{ext}. Ref. [1] starts with the assumption (in equation (26)) that

$$p(F|F_{\mathrm{ext}}) = \frac{1}{Z(F_{\mathrm{ext}})} \exp\{H(F|F_{\mathrm{ext}})\} \ , \tag{5}$$

where $Z(F_{\mathrm{ext}}) \equiv \int dF \exp\{H(F|F_{\mathrm{ext}})\}$, and (equation (27))

$$H(F|F_{\mathrm{ext}}) = rF^2 + gF^4 \tag{6}$$

for small F, with the real constants r and g respectively positive and negative.

Dewar's claim that the model defined in this way exhibits SOC rests on the assertion that the quantity $\langle F^2 \rangle \equiv \int dF F^2 p(F|F_{\mathrm{ext}})$ diverges in the limit $F_{\mathrm{ext}} \to 0$. The only divergence in the integral defining $\langle F^2 \rangle$ can come from large F, since the integrand is well behaved near $F = 0$. In particular, for $\langle F^2 \rangle$ to be infinite at a given F_{ext},

$p(F|F_{\text{ext}})$ must not decrease more rapidly than F^{-3} as $F \to \infty$, aside from slowly varying corrections such as logarithms. Such an algebraically slow decrease of $p(F|F_{\text{ext}})$ with F means that the distribution, $D(n)$, of avalanche sizes n, also falls off algebraically slowly with n for large n, which is the defining characteristic of SOC[4].

The assertion that $\langle F^2 \rangle$ diverges is unjustified, however. In fact, since the behaviour of $H(F|F_{\text{ext}})$ at large F is unspecified in the model, no conclusion about the large-F behaviour of $p(F|F_{\text{ext}})$ and hence about the behaviour of $\langle F^2 \rangle$ can be drawn. If one assumes a particular large-F form for $p(F|F_{\text{ext}})$, moreover, that assumed form alone determines whether the model has SOC or not. If, for example, one assumes that $p(F|F_{\text{ext}}) \sim 1/F^\alpha$ for some power α as $F \to \infty$, then one has by *fiat* defined the model to exhibit SOC. (Any value of α suffices to ensure SOC, even though $\langle F^2 \rangle$ is finite for $\alpha > 3$.) If, on the other hand, one assumes that $p(F|F_{\text{ext}})$ decreases faster than a power of F – exponentially say – then by definition the model does not exhibit SOC.

It remains to understand, then, how ref. [1] concludes that $\langle F^2 \rangle$ diverges as $F_{\text{ext}} \to 0$. The point is that $\langle F^2 \rangle$ is computed in ref. [1] under a ("mean-field") approximation wherein $H(F|F_{\text{ext}})$ is expanded to quadratic order in $(F - \langle F \rangle)$ (with $\langle F \rangle = F_{\text{ext}} = \sqrt{-r/2g}$ in mean field approximation), the cubic and quartic terms being neglected. Implicitly, moreover, this quadratic form is assumed to hold not just for small F but for all F, so that $\langle F^2 \rangle$ can be computed.‡ Since the coefficient of the quadratic term vanishes like $|g|F_{\text{ext}}^2$ (i.e., like r) as $F_{\text{ext}} \to 0$, the integral defining $\langle F^2 \rangle$ diverges in this limit. Note, however, that if one keeps the higher order terms in the expansion of $H(F|F_{\text{ext}})$, and again assumes that the resulting form – now *quartic* in $(F - \langle F \rangle)$ – holds for all F to allow the computation of $\langle F^2 \rangle$, one readily sees that $\langle F^2 \rangle$ stays finite at $F_{\text{ext}} = 0$, where one finds $\langle F^2 \rangle \propto 1/\sqrt{|g|}$. (This result follows immediately from changing the integration variable from F to $y \equiv F|g|^{1/4}$ in the integrals defining Z and $\langle F^2 \rangle$, assuming that $|g| > 0$ at $F_{\text{ext}} = 0$.) Thus the apparent divergence of $\langle F^2 \rangle$ is an artefact of using the mean-field approximation and assuming that the resulting quadratic form remains valid for large F.

Though it has yet to be demonstrated, Dewar's conjecture in [1] that MaxEP or a similar principle underlies SOC remains an intuitively appealing notion.

[1] Dewar R 2003 *J. Phys. A: Math. Gen.* **36** 631
[2] Dewar R C 2005 *J. Phys. A: Math. Gen.* **38** L371
[3] Bruers S 2006 preprint (arXiv:cond-mat/0604482 v2, 5 Sep 2006)
[4] E.g., Jensen H J 1998 *Self-Organized Criticality* (Cambridge, UK: Cambridge University Press)

‡ The mean-field result $\langle F \rangle = \sqrt{-r/2g}$ is only self-consistent if the integrals over F in the computation of $\langle F^2 \rangle$ using the quadratic approximation for $H(F|F_{\text{ext}})$ are taken to run from $-\infty$ to ∞. Our conclusion remains valid if the integrals are taken to run from 0 to ∞, however.