Validity of the Boltzmann distribution under non-linear constraints

1988 J. Phys. A: Math. Gen. 21 1467

(http://iopscience.iop.org/0305-4470/21/6/021)

J. Phys. A: Math. Gen. 21 (1988) 1467-1469. Printed in the UK

COMMENT

Validity of the Boltzmann distribution under non-linear constraints

A J M Garrett†

Sir Frank Packer Theoretical Department, School of Physics, University of Sydney, NSW 2006, Australia

Received 13 October 1987

Abstract. The Boltzmann distribution $\exp(-\lambda E_i)$ for the occupancy p_i of a fixed energy level E_i derives from a maximum entropy procedure under the constraint of energy conservation $\Sigma_i E_i p_i = U_0$. This constraint is linear in the p_i. It is shown that the form of the distribution still holds when the energy levels E_i depend on the occupancies p_i, provided that $\mathrm{d}U \equiv \Sigma_i E_i \,\mathrm{d}p_i$ is a perfect differential so that energy is conserved. The energy constraint then becomes non-linear in the p_i. Specific applications of this result are covered, in particular the Poisson-Boltzmann equation for the electrostatic potential in plasma in thermal equilibrium.

The Boltzmann, or canonical, distribution for the probability of occupancy p_i of a fixed energy level E_i is

$$p_i = Z(\lambda)^{-1} \exp(-\lambda E_i) \tag{1}$$

(Boltzmann 1871), where the partition function $Z(\lambda)$ is given by insisting that (1) be normalised, and λ is in principle expressible in terms of the expectation value of the energy, U_0, by substituting (1) into the equation of energy conservation. The distribution is derived by maximising the information entropy relative to uniform measure on the i,

$$S \equiv -\sum p_i \ln p_i \tag{2}$$

subject to the constraints of normalisation and energy conservation alone; λ is the Lagrange multiplier associated with the latter. To work out the form of this constraint, calculate the energy increment $\mathrm{d}U$ by building up the occupancy bit by bit:

$$\mathrm{d}U = \sum_i E_i \,\mathrm{d}p_i \tag{3}$$

whence, on integrating and choosing the zero of energy appropriately,

$$U = \sum_i E_i p_i. \tag{4}$$

Energy conservation states simply that $U = U_0$. It is assumed that the expectation value coincides with the actual value.

† Present address: 234 Dialstone Lane, Stockport, Cheshire SK2 7LE, UK.

0305-4470/88/061467+03\$02.50

(A technical point: during the build-up process the probability distribution appears not to be normalised; initially all the p_i are zero. In fact (3) is a shorthand form in which p_i is postulated to represent a normalised physical quantity.)

The variational procedure used is the principle of maximum entropy, an enormously powerful tool introduced into physics and probability theory in fullest generality by Jaynes (1983). *Ab initio* derivation of the Boltzmann distribution have traditionally derived the principle combinatorially (see, for example, French 1958); but it also arises in many other diverse ways, as stressed by Jaynes (1986).

Now suppose that the energy levels are brought to depend dynamically on their occupancies, so that E_i is not only a function of i but also of the p_i. It will be shown that (1) still holds, albeit as a transcendental relation between the p_i, provided that $\partial E_i/\partial p_j = \partial E_j/\partial p_i$; (3) is then always a perfect differential, and $E_i = \partial U/\partial p_i$. Extension to the functional case, with continuous suffixes, is immediate. Clearly if (3) is imperfect, the system is non-conservative and it makes no sense even to refer to energy conservation.

As before, we maximise (2) subject to normalisation $N \equiv \Sigma_i p_i = 1$ (with Lagrange multiplier $(\ln Z - 1)$, for convenience) and energy conservation. Now

$$\mathrm{d}[S - (\ln Z - 1)N - \lambda U] = \sum_i [-\ln(Zp_i) - \lambda E_i(\boldsymbol{p})]\,\mathrm{d}p_i \qquad (5)$$

$$= 0 \qquad \forall \mathrm{d}p_i \qquad (6)$$

at an extremum, whence $p_i = Z^{-1}\exp[-\lambda E_i(\boldsymbol{p})]$ immediately. It is no longer trivial to isolate $Z(\lambda)$, and the solutions for p_i may be non-unique or complex. Complex solutions imply the absence of any physically meaningful maximum; the concept of ordering numbers according to their magnitude, so as to find the maximum, is necessarily restricted to the reals. In linear-constraint theory, by contrast, the existence of a unique absolute maximum is guaranteed (Jaynes 1963).

The significance of this result is in applying the principle of maximum entropy to a class of problems with constraints non-linear in the p_i, and in establishing conditions for the validity of the Boltzmann distribution. Necessary conditions are that energy be conserved and that no other constraint (apart from normalisation) operate. In addition, the solution for p_i must be real; no general test for this condition exists.

Specific non-linear-constraint problems have arisen before. Joyce and Montgomery (1973) derived the Boltzmann distribution (and the principle of maximum entropy, using combinatorics) for a quadratic form of $U(\boldsymbol{p})$ arising from charge–charge interactions in two dimensions. This derivation is reproduced by Ting *et al* (1987). E_i becomes the potential $\phi(\boldsymbol{r})$, and its dynamical relation to $p_i = p(\boldsymbol{r})$ is Poisson's equation. The 'self-consistent' equation for $\phi(\boldsymbol{r})$ found by eliminating $p(\boldsymbol{r})$ between Poisson's equation and the Boltzmann distribution is called, logically, the Poisson–Boltzmann equation; it is more usually—but incorrectly—written down without regard to non-linear constraints. Its solution for given λ is unique (Garrett and Poladian 1988). Another quadratic form of $U(\boldsymbol{p})$ has been considered by van Kampen (1964) in the context of imperfect gas theory; van Kampen states that van der Waal's equation was derived in similar fashion long before (Ornstein 1908). Non-linear constraints in maximum entropy analysis are nothing new! Evans (1979) has asserted that $1/f$ noise can be treated by means of a modified non-linear constraint analysis, and also that in some circumstances the use of maximum entropy with non-linear constraints is inconsistent, a claim denied here. Finally, a mathematical analysis identical to the present is used in Bayesian image processing, in which the noise is represented by a non-linear statistic (almost always χ^2) and is constrained to take a particular value; see Gull and Skilling (1984). These authors confirm uniqueness in the most commonly encountered case, but it can fail for phaseless Fourier data: Gull and Daniell (1978) quote the existence of a counterexample due to Skilling.

The author acknowledges receipt of a Professor Harry Messel Research Fellowship from the School of Physics of the University of Sydney.

References

Boltzmann L 1871 *Wien. Ber.* **63** 397, 679, 712
Evans R B 1979 *The Maximum Entropy Formalism* ed R D Levine and M Tribus (Cambridge, MA: MIT) p 169
French A P 1958 *Principles of Modern Physics* (New York: Wiley)
Garrett A J M and Poladian L 1988 *Phys. Rep.* submitted
Gull S F and Daniell G J 1978 *Nature* **272** 686
Gull S F and Skilling J 1984 *Proc. IEE* **131F** 646
Jaynes E T 1963 *Statistical Physics* (*1962 Brandeis Lectures*) ed K W Ford (New York: Benjamin) ch 4 (reprinted in Jaynes (1983) ch 4)
—— 1983 *Papers on Probability, Statistics and Statistical Physics* (*Synthese series* **158**) ed R D Rosencrantz (Dordrecht: Reidel)
—— 1986 *Maximum Entropy and Bayesian Methods in Applied Statistics* ed J M Justice (Cambridge: Cambridge University Press) p 1
Joyce G and Montgomery D 1973 *J. Plasma Phys.* **10** 107
Ornstein L S 1908 *Thesis* Leiden
Ting A C, Chen H H and Lee Y C 1987 *Physica* **26D** 37
van Kampen N G 1964 *Phys. Rev.* **135** A362