ELSEVIER

Mathematical Biosciences 153 (1998) 151–161

Mathematical Biosciences
an international journal

Demography and the canonical ensemble

Jonathan D.H. Smith [*]

Department of Mathematics, Iowa State University, Ames, IA 50011-2064, 400 Carver Hall, USA

Received 23 September 1997; received in revised form 14 July 1998

Abstract

The Gibbs canonical ensemble of statistical mechanics is used to describe the probability distribution of the age classes of mothers of new-borns in an age-structured population. The Malthusian parameter emerges as a Lagrange multiplier corresponding to a generation time constraint, while a new perturbation parameter appears as the Lagrange multiplier corresponding to a maternity constraint. Classical Lotka stability reduces to the unperturbed case of the more general canonical ensemble model. The model is used in a case study of the female (peninsular) Malaysian population of 1970. The Malthusian parameter and perturbation are calculated easily by linear regression. Use of the model identifies an anomaly in the population due to the effects of World War II. © 1998 Elsevier Science Inc. All rights reserved.

1. Introduction

The Gibbs canonical ensemble model from statistical mechanics assigns a probability to each of a set of states of a system, subject precisely to knowledge of the expected value of one or more numerical parameters associated to each state ([1] and p. 151 in [2]). In equilibrium thermodynamics, the parameter is the energy of a state. The canonical ensemble describes the probability distribution of the states when the system is in equilibrium with its surroundings at a given temperature. The (inverse) temperature appears in the model as the Lagrange multiplier associated with the constraint of known expected energy, imposed as a condition on the maximization of the entropy of the distribution.

[*] Tel.: +1-515 294 8172; fax: +1-515 294 5454; e-mail: jdhsmith@pollux.math.iastate.edu

PII: S0025-5564(98)10035-4

In a previous paper [3], the canonical ensemble model was applied to the dynamics of a non-equilibrium situation in biology, namely Eigen's phenomenological rate equations describing the evolution of competing species in the presence of fixed resources. Here each species represented a state, and the numerical parameter associated with each state was the natural growth rate of the species in the presence of unlimited resources. In this application, the Lagrange multiplier corresponding to the constraint turned out to be the intrinsic age of the system.

The current paper applies the canonical ensemble model to the study of an age-structured population. The states are the various age classes of the mothers of new-borns. Two numerical parameters are associated with each state: its age, and the negated logarithm of the corresponding value of the net maternity function. Since two constraints are placed on the maximization of the entropy, there are two Lagrange multipliers appearing. The first is identified as the Malthusian parameter, while the second is identified as a perturbation from Lotka stability. In fact, Lotka stability (or Lotka equilibrium) appears in context as a special case of the canonical ensemble model, namely the case of equal age differences between successive age classes and vanishing of the perturbation parameter.

Details of the classical projection-matrix approach to the modelling of age-structured populations are summarized in Section 2. Constancy of the projection matrix elements is assumed. The notation has been adapted to the context of this paper (e.g. preparing for the subsequent focus on the reproductive age classes). Successive age classes are not assumed to have constant age differences. Although such an assumption underlies the classical eigenvalue method, it is not needed for the canonical ensemble model. One minor difference from many treatments is to work with row vectors, matrices then multiplying on the right. Besides the obvious typographic advantages of row vectors over column vectors, this also enables one to identify the projection matrix directly as the incidence matrix of the weighted directed graph (Fig. 1) describing the development of the population. (Transposition of matrices and reversal of the order of their multiplication recovers the opposite convention.) Section 2 concludes with specialization to the Lotka-stable case, recalling the logarithm of the dominant eigenvalue of the projection matrix as the Lotka growth rate r_1 (Eq. (7)).

Section 3 applies the canonical ensemble model, assigning a probability (Eq. (14)) for the mother of a randomly chosen new-born to belong to a given reproductive age class. The assignment assumes that one's knowledge about the population amounts precisely to knowledge of the generation time (Eq. (9)) and (negated) logarithmic maternity (Eq. (10)), nothing else being known. The Malthusian parameter r is the Lagrange multiplier corresponding to the generation time constraint, while the perturbation s is the Lagrange multiplier corresponding to the maternity constraint.

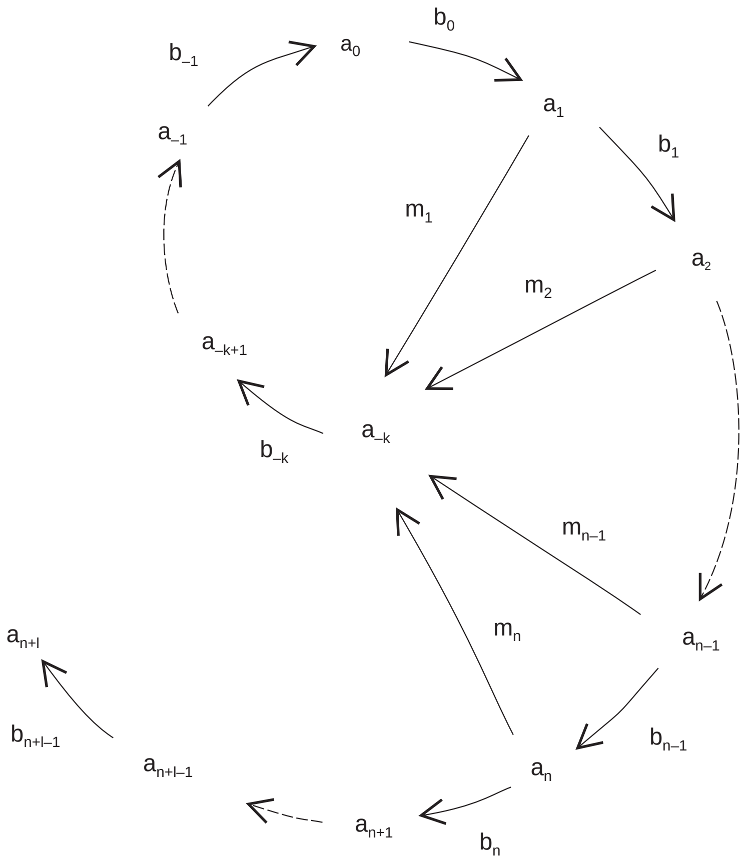

Fig. 1. The clock of ages.

A comparison with the case of Lotka stability is made in Section 4. It is shown that the canonical ensemble model reduces to Lotka stability on setting its Malthusian parameter r to equal the Lotka growth rate r_1, and on letting its perturbation s vanish. Of course, for classical Lotka stability, one also needs to assume a constant difference between successive demographically active age classes.

Section 5 presents a canonical ensemble analysis of a specific age-structured population, namely (peninsular) Malaysian females in 1970. It uses data from [4], summarized in Table 1. For this population, one obtains $r = 0.03402$ and $s = -0.02591$, contrasting with the Lotka growth rate $r_1 = 0.02697$. As a predictor of growth, the Malthusian parameter r here does about as well as the Lotka growth rate r_1. However, the Malthusian parameter r is obtained easily by linear regression, while the Lotka growth rate has to be extracted as a root

Table 1
Malaysia 1970 ([4], pp. 380–381)

Age class	Age group	Female pop. 1970	Female pop. 1975	Births by mother's age	Estimated births	Net maternity function
a_1	10–14	611 164	672 984	270	280	0.0010
a_2	15–19	513 404	594 032	27 778	25 600	0.1229
a_3	20–24	392 904	493 447	88 673	86 200	0.5092
a_4	25–29	286 722	386 734	76 088	84 400	0.5935
a_5	30–34	275 611	284 690	60 386	58 500	0.4847
a_6	35–39	219 704	272 234	30 679	31 400	0.3048
a_7	40–44	190 613	212 909	10 746	10 800	0.1209
a_8	45–49	161 671	186 893	2075	2090	0.0268
a_9	50–54	141 504	152 740	470	457	0.0067
Total		2 793 297	3 256 663	297 165		

of a high-order algebraic equation. Comparing the probability distribution for the age structure of the mothers of new-borns given by the canonical ensemble model with that actually arising from the data, a notable discrepancy arises amongst the cohort of mothers born in the early 1940s. This may shed light on disruptive effects of the occupation of (then) Malaya during World War II, effects that were not built in to the canonical ensemble model. A similar canonical ensemble analysis of other populations may turn out to be an extremely useful tool for the detection of such anomalies.

2. Projection matrices and Lotka stability

Consider a population with age classes $a_{-k}, \dots, a_0, a_1, \dots, a_n, a_{n+1}, \dots, a_{n+l}$, arranged in order of increasing ages. An individual of age a_i has a constant positive probability b_i of surviving to the next age a_{i+1}. No individual of age class a_{n+l} survives to a later age class. The only reproductive age classes are $a_1, \dots, a_n$. During their sojourn in each such reproductive age class a_i, individuals give birth to a positive average number m_i of new-borns falling into age class a_{-k}. In breakdowns of female human populations such as [4], one might have $k = 1, n = 9, l = 7$, with a_{-1} the 0–4 year age class, $\dots, a_1$ the 10–14 year olds, $\dots, a_{15}$ the 80–84 year olds, and a_{16} those aged 85 or more.

The process is represented by the directed weighted graph or directed network of Fig. 1, the 'clock of ages'. Demographic significance resides in the subnetwork obtained by deleting the post-reproductive vertices $a_{n+1}, \dots, a_{n+l}$ and the edges incident with them. The adjacency matrix of this subnetwork is the (transposed) Leslie or *projection matrix*

$$A = \left[\begin{array}{ccccc|ccccc}
0 & b_{-k} & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & & 0 & 0 & 0 & 0 & & 0 & 0 \\
 & & & & & & & & & \\
0 & 0 & & 0 & b_{-1} & 0 & 0 & & 0 & 0 \\
0 & 0 & & 0 & 0 & b_0 & 0 & & 0 & 0 \\
\hline
m_1 & 0 & & 0 & 0 & 0 & b_1 & & 0 & 0 \\
m_2 & 0 & & 0 & 0 & 0 & 0 & & 0 & 0 \\
 & & & & & & & & & \\
m_{n-1} & 0 & & 0 & 0 & 0 & 0 & & 0 & b_{n-1} \\
m_n & 0 & & 0 & 0 & 0 & 0 & & 0 & 0
\end{array}\right]. \tag{1}$$

For $n > 1$, the clock of ages has cycles $a_{-k}a_{-k+1}\dots a_1a_{-k}, \dots, a_{-k}a_{-k+1}\dots a_na_{-k}$ of coprime lengths, so that the matrix A is primitive [5–7]. The Perron–Frobenius Theorem ([8], Theorem XVI.5; [9], Section 13.2) then implies that the non-negative matrix A has a real, positive, dominant simple eigenvalue λ with a positive (row) eigenvector u. Simplicity of λ means that its eigenspace just consists of scalar multiples of u. Dominance of λ means that λ exceeds the absolute value of any other eigenvalue of A.

For an explicit determination of λ and u, it is convenient to introduce some notation. For $i \leqslant n$, define

$$l_i = \prod_{j<i} b_i, \tag{2}$$

the probability that a new-born survives to age a_j. The characteristic equation of A may then be written in the form

$$\lambda^{k+1} = \sum_{i=1}^{n} \lambda^{-i} l_i m_i \tag{3}$$

([10], Section 3), and the components of the eigenvector $u = [u_{-k}\dots u_{-1}\ u_0\ u_1\dots u_n]$ may be taken as

$$u_i = \lambda^{-i} l_i \tag{4}$$

for $-k \leqslant i \leqslant n$ (section 6 in [10]).

Now suppose that there is a constant time difference z between the successive non-post-reproductive age classes $a_{-k}, \dots, a_n$. If the components of the row vector $x^t = [x^t_{-k}\dots x^t_0\ x^t_1\dots x^t_n]$ record the number of individuals in each such age class at time t, the projection matrix (Eq. (1)) may be used to describe the development of the population over time via

$$x^{t+z} = x^t A. \tag{5}$$

The population is said to exhibit *Lotka stability* at time t if x^t lies in the eigenspace of λ. One then has $x^{t+z}_i = \lambda x^t_i$ for $-k \leqslant i \leqslant n$, so that

$$\log x_i^{t+1} = z^{-1} \log \lambda + \log x_i^t. \tag{6}$$

The logarithmic increase

$$r_1 = z^{-1} \log \lambda. \tag{7}$$

over unit time is called the *Lotka growth rate*.

3. The canonical ensemble

The canonical ensemble model starts with an age-structured population as described in the previous section, but without the assumptions of Lotka stability or constant difference z between the non-post-reproductive age classes. Consider the experiment of choosing a new-born at random, and determining the age class of her mother. Suppose that the mother lies in age class a_i with probability q_i, for $1 \leqslant i \leqslant n$. One thus has

$$\sum_{i=1}^{n} q_i = 1. \tag{8}$$

Define the *generation time*

$$T = \sum_{i=1}^{n} q_i a_i \tag{9}$$

as the expected age of the mother. Define the (*logarithmic*) *maternity*

$$M = -\sum_{i=1}^{n} q_i \log l_i m_i \tag{10}$$

as the expected value of the negated logarithm of the *net maternity function* $l_i m_i$. Note that the quotient M/T is the *reproductive potential* ([11], ref. 9). Suppose that the numerical values of the generation time and maternity (or generation time and reproductive potential) are known, but that one has no further information about the probability distribution $[q_1\ q_2 \ldots q_n]$. As discussed in [3], the appropriate probability distribution for this model is the one that maximizes the *entropy*

$$H = -\sum_{i=1}^{n} q_i \log q_i \tag{11}$$

subject to the constraints (8)–(10). (Briefly, the validity of a different distribution would amount to different knowledge of details of the experiment from that summarized by specification of the generation time and logarithmic maternity.)

Determination of the probability distribution is achieved using the Lagrangean function

$$\begin{aligned} L(q_i,\alpha,r,s) = {} & -\sum_{i=1}^{n} q_i \log q_i + \alpha\left(1-\sum_{i=1}^{n} q_i\right) \\ & + r\left(T-\sum_{i=1}^{n} q_i a_i\right) + (1+s)\left(M+\sum_{i=1}^{n} q_i \log l_i m_i\right). \end{aligned} \tag{12}$$

The stationarity conditions $\partial L/\partial q_i = 0$ for $1 \leqslant i \leqslant n$ lead to

$$\log q_i = -(1+\alpha) - r a_i + (1+s)\log l_i m_i \tag{13}$$

or $q_i = \mathrm{e}^{-ra_i}(l_i m_i)^{1+s}/\exp(1+\alpha)$. Substituting into Eq. (8), noting that α is independent of i, one obtains

$$q_i = Z(r,s)^{-1}\mathrm{e}^{-ra_i}(l_i m_i)^{1+s} \tag{14}$$

with the *partition function* or *Zustandsumme*

$$Z(r,s) = \sum_{i=1}^{n} \mathrm{e}^{-ra_i}(l_i m_i)^{1+s}. \tag{15}$$

The Lagrange multiplier r corresponding to the generation time constraint (9) is called the *Malthusian parameter*. The (shifted) Lagrange multiplier s corresponding to the maternity constraint (10) is called the *perturbation*.

4. Comparison with Lotka-stable populations

In order to understand the probability distribution (14) given by the canonical ensemble, and the interpretation of the Lagrange multipliers, it is useful to consider the case of a Lotka-stable population with a constant time difference z between successive non-post-reproductive age classes. In this case, Eq. (4) yields the probability

$$q_i = \lambda^{-i} l_i m_i \Big/ \sum_{j=1}^{n} \lambda^{-j} l_j m_j \tag{16}$$

for the mother of a randomly chosen new-born to lie in a reproductive age class a_i, namely $1 \leqslant i \leqslant n$. Note that

$$a_i = a_0 + iz \tag{17}$$

for $1 \leqslant i \leqslant n$. By Eq. (7), one then has

$$\lambda^{-i} = \mathrm{e}^{-r_1 iz} = \mathrm{e}^{r_1 a_0}\mathrm{e}^{-r_1 a_i}. \tag{18}$$

The Lotka-stable probability Eq. (16) may thus be written in the form

$$q_i = Z(r_1,0)^{-1}\mathrm{e}^{-r_1 a_i} l_i m_i \tag{19}$$

with

$$Z(r_1, 0) = \sum_{i=1}^{n} \mathrm{e}^{-r_1 a_i} l_i m_i. \tag{20}$$

Comparing Eqs. (19) and (20) with Eqs. (14) and (15), it emerges that the Lotka-stable distribution is just the special case of the canonical ensemble distribution with zero perturbation, the Lotka growth rate being the Malthusian parameter of the canonical ensemble model. One may thus characterize Lotka stability within the canonical ensemble model as the freedom from perturbation.

5. A case study: Malaysia 1970

To illustrate the application of the canonical ensemble model, a case study of the 1970 female population of (peninsular) Malaysia will be undertaken, based on [4], pp. 380 ff. This choice is motivated by the availability of good data, and by the relative stability (in the non-technical sense) of the population at that time. The immediate post-independence ‘insurgency’ had abated, while the subsequent rapid industrialization had not yet taken hold. The key data are summarized in Table 1, except for the Lotka growth rate $r_1 = 0.02697$. (Note that the ‘estimated births’ column of Table 1 does not come from [4]: its origin will be clarified below.)

In the absence of information about the variation of the male/female birth ratio with the age of the mother, this ratio will be assumed to be constant. Dividing the entries of column 5 by the total at the bottom thus gives a list of nine relative frequencies yielding an approximation $[p_1\ p_2 \ldots p_9]$ to the probability distribution $[q_1\ q_2 \ldots q_9]$. Negating Eq. (13) and using these relative frequencies p_i in place of q_i yields a system of linear equations

$$1.(1+\alpha) + a_i r - (\log\, l_i m_i)(1+s) = -\log\, p_i, \tag{21}$$

$1 \leqslant i \leqslant 9$. The ages a_i are taken to be central in each corresponding class, i.e. $a_1 = 12, a_2 = 17, a_3 = 22, \ldots, a_9 = 52$.

For a number of reasons, not the least being that the canonical ensemble model is only a model, the system (21) is inconsistent. Its least squares solution yields

$$\log Z = -0.1681 \tag{22}$$

for the ‘thermodynamic potential’ $(1+\alpha)$ ([3], (3.10)),

$$r = 0.03402 \tag{23}$$

for the Malthusian parameter, and

$$s = -0.02591 \tag{24}$$

for the perturbation. Reinstating these values into the left-hand side of Eq. (21), and normalizing to the total number of births from the bottom of the 'birth's by mother's age' column of Table 1, one obtains the 'estimated births' listed in the adjacent column of Table 1. One may view these as the best estimate of the age distribution of this total number of births, assuming that the canonical ensemble model were exact. The most striking discrepancy is the model's 10% overestimate of the births to the 25–29 year old mothers. Note that these mothers were born during the occupation of World War II. Comparing the relative sizes of the age classes for 1970 and 1975, it is apparent that the cohort aged 25–29 in 1970 was anomalously small. The canonical ensemble model only incorporated Eqs. (9) and (10), not additional knowledge about World War II.

The value $r = 0.03402$ of Eq. (23) for the Malthusian parameter may be contrasted with the value $r_1 = 0.02697$ for the Lotka growth rate. Comparison of the total 10–54 year old female populations in 1970 and 1975, from the bottom of the columns of Table 1, suggests a value of $\frac{1}{5}\log(3\,256\,663/2\,793\,297) = 0.03070$ for the true growth rate. The Lotka growth rate underestimates this by 12%, while the Malthusian parameter of the canonical ensemble model overestimates by 11%. It should be noted that calculation of the Malthusian parameter by the least squares solution to the linear system (21) is mathematically simpler than extraction of the Lotka growth rate as a logarithm of a root of the high-degree algebraic Eq. (3).

6. Discussion

There have been previous applications of the canonical ensemble model to demography, most notably by Demetrius et al., [11–15]. Demetrius' approach is quite different from that presented above. Except for what physicists describe as 'toy models' in the much harder density-dependent case (Section 5 in [11]), his approach focusses on the density-independent case, with a constant projection matrix (Eq. (1)). Taking a phase space Ω consisting of all infinite paths through the clock of ages, and a shift operation τ corresponding to the passage of one time unit, it considers the invariant measure μ on Ω given by the Lotka-stable probability Eq. (16). It then uses results from ergodic theory concerning the dynamical system (Ω, μ, τ), for example ([12,(6.3)] [13,(2.16)] [14,(5.6)] [11,[10]]) obtaining the relationship

$$r_1 = (H - M)/T \tag{25}$$

between Eq. (7) and Eqs. (9) and (10) "by invoking the thermodynamic formalism described in [16]" ([11], p. 3494).

The mathematical basis for the current paper – merely linear algebra and Lagrange multipliers – is far simpler than the deep ergodic theory used in

Demetrius' approach. For example, taking the negative of the expected value of each side of Eq. (13) in the Lotka-stable case $r = r_1, s = 0$ of Section 4 yields

$$H = \log Z(r_1, 0) + r_1 T + M. \tag{26}$$

However, with $a_0 = (k + 1)z$, the characteristic Eqs. (3) and (20) show that the partition function $Z(r_1, 0)$ reduces to unity in this case. Thus Eq. (26) yields Demetrius' relationship Eq. (25) in completely elementary fashion, without reliance on the elaborate machinery of ergodic theory. Incidentally, it should be noted that Eq. (25) may still give good approximations to the Lotka growth rate in populations that do not exhibit Lotka stability. In the 1970 Malaysia population of Section 5, Eq. (25) yields $r_1 \simeq 0.02722$, a 1% overestimate.

Charlesworth ([17], p. 180) offered the following critique of Demetrius' use of entropy in [13]:

> Demetrius has argued that a measure of entropy of the life-history is a useful predictor of gene-frequency change. While this may be true under some special circumstances, the result is derived from the properties of the intrinsic rate of increase, and therefore offers no new insights into the dynamics of selection.

Demetrius [18] pointed out, however, that a uniform rescaling of the net maternity function values $l_i m_i$ in the Lotka stable case will change the Lotka growth rate, while the entropy remains invariant. The general case of non-negative perturbation s presented in this paper gives further evidence that the entropy has significance independently of the Lotka growth rate. Indeed, in the present context, the entropy Eq. (11) and the partition function Eq. (15) determine each other as mutual Legendre transforms [19,20], while the partition function acts as the moment generating function for the distribution q_i. In this way the entropy function may be said to determine all the relevant macroscopic quantities.

Acknowledgements

The author is grateful for correspondence with L. Demetrius, and the comments of anonymous referees, that helped in the preparation of this paper.

References

[1] J.W. Gibbs, in: Scientific Papers of J. Willard Gibbs, H.A. Bumstead and R.G. van Name (Eds.), Longmans, London, 1906.

[2] W. Pauli, in: A. Smekal (Ed.), Die allgemeinen Prinzipien der Wellenmechanik, in Handbuch der Physik, Bd. 24/1 Quantentheorie, Springer, Berlin, 1933, p. 83.

[3] J.D.H. Smith, Competition and the canonical ensemble, Math. Biosci. 133 (1996) 69.

[4] N. Keyfitz, W. Flieger, World Population Growth and Aging, University of Chicago, Chicago, IL, 1990.

[5] L. Demetrius, Primitivity conditions for growth matrices, Math. Biosci. 12 (1971) 53.

[6] B. Parlett, Ergodic properties of population I: the one sex model, Theor. Pop. Biol. 1 (1970) 191.

[7] D. Rosenblatt, On the graphs and asymptotic forms of finite Boolean relation matrices and stochastic matrices, Naval Res. Logist. Quart. 4 (1957) 151.

[8] G. Birkhoff, Lattice Theory, American Mathematical Society, Providence, RI, 1967.

[9] F.R. Gantmacher, Matrizentheorie, Deutscher Verlag der Wissenschaften, Berlin, 1986.

[10] E.G. Lewis, On the generation and growth of a population, Sankhyā 6 (1942) 93.

[11] L. Demetrius, Directionality principles in thermodynamics and evolution, Proc. Nat. Acad. Sci. 94 (1997) 3491.

[12] L. Demetrius, Statistical mechanics and population biology, J. Stat. Phys. 30 (1983) 709.

[13] L. Demetrius, Growth rate, population entropy, and perturbation theory, Math. Biosci. 93 (1989) 159.

[14] L. Demetrius, The thermodynamics of evolution, Physica A 189 (1992) 417.

[15] L. Arnold, V.M. Gundlach, L. Demetrius, Evolutionary formalism for products of positive random matrices, Ann. Appl. Prob. 4 (1994) 859.

[16] D. Ruelle, Thermodynamic Formalism, Addison-Wesley, Reading, MA, 1978.

[17] B. Charlesworth, Evolution in Age-Structured Populations, 2nd Ed., Cambridge University, Cambridge, 1992.

[18] L. Demetrius, Evolutionary entropy, punctuated equilibrium and phyletic gradualism, preprint, 1998.

[19] W.T. Grandy Jr., Foundations of Statistical Mechanics, Vol. I, Reidel, Dordrecht, 1987.

[20] J.D.H. Smith, Barycentric algebras, canonical distributions and Legendre transforms, Iowa State University Mathematics Report Number M98-01, 1998.