Statistical Mechanical Approach to Modelling Ecosystems

Dissertation submitted to Jawaharlal Nehru University in partial fulfilment for the award of the degree of

MASTER OF PHILOSOPHY

RAKESH KUMAR RANJAN

School of Environmental Sciences
Jawaharlal Nehru University
New Delhi-110067
INDIA

2004

जवाहरलाल नेहरू विश्वविद्यालय
Jawaharlal Nehru University
SCHOOL OF ENVIRONMENTAL SCIENCES
New Delhi – 110 067

CERTIFICATE

The research work embodied in this dissertation entitled "*Statistical Mechanical Approach to Modelling Ecosystems*" has been carried out in the School of Environmental Science, Jawaharlal Nehru University, New Delhi. This work is original and has not been submitted in part or full for any other degree or diploma in any university/institution.

G. P. Malik
Prof. G. P. Malik
(Supervisor)

Ranjan
Rakesh Kumar Ranjan
(Candidate)

Prof. J. Behari
(Dean, SES)

GRAM : JAYENU TEL. : 26107676, 26167557 EXT. : 2302, 03, 04 TELEX : 031-73167 JNU IN
FAX : 91-011-26165886, 26172438

Dedicated to my dear brother

Mintu..

ACKNOWLEDGMENT

With deep sense of gratitude, I would like to thank my teacher and supervisor Prof. G. P. Malik for his invaluable insights and precious guidance. I also acknowledge having benefited immensely from his lectures on the subject of 'Mathematical Ecology' and 'Mathematical Modelling'.

I am thankful to Prof. J. Behari, Dean, School of Environmental Sciences, for providing adequate facilities in carrying out this work. Special thanks are due for Prof. J. Subba Rao and Prof. V. K. Jain for their suggestions and discussions, which proved vital during the study.

I take this opportunity to thank my parents Shri Mohan Mahto and Mrs. Sonamati Mahto who have been a pillar of strength for me in all ups and downs.

I would like to thank Mr. Akhilananda and Mr. Amit Prakash for their valuable suggestions and help at various stages of the work.

I deeply acknowledge the cooperation of my friends Avi, Banti, Babu, Chandan, Dilip, Gangeya, Manoj, Shambhu and Vikas who helped me through each and every moment whenever I got into a problem and Who were always there for me to boost up.

I am indebted to library staff of Jawaharlal Nehru University, Delhi University, National Medical library, AIIMS and Indian Agricultural Research Institute For their cooperation.

RAKESH KUMAR RANJAN

CONTENTS

Chapter 1 *Page no.*

An overview of Ecology and introduction to modelling ecosystem via Statistical Mechanics

Chapter 2

Kerner's statistical mechanical approach to modelling ecosystem

Continued.

Chapter 3

Statistical mechanics of the Gompertz model via Kerner's method (with and without the assumption of the anti-symmetry of the community matrix)

Chapter 4

Experimental evidence for log-normal distribution for interacting multi-species ecosystems

Chapter 1

Chapter 1

An overview of Ecology and introduction to modelling ecosystem via Statistical Mechanics

1.1 Introduction
1.2 An overview of different approaches to study an ecosystem
1.3 Evolution of mathematical models of populations
1.4 The essence of Statistical Mechanics
1.5 An example of statistical mechanics at work: A 2-level system
1.6 Motivation for use of statistical Mechanics to study the ecosystem
1.7 Scope of the dissertation

1.1 Introduction

The word "ecology", which denotes a branch of biology that deals with the relations of organisms and their environment, including their relations with other organisms, is derived from the word "eco", which means "house", or a "dwelling place". Ecology is an interdisciplinary field, cutting across the life and geophysical sciences. Ecological investigations look into two directions: (1) The nature of environments and the demands which these environments make upon the organisms that inhabit them; and (2) the characteristics of organisms (plant or animal), species, and groups that permit or promote their tolerance of specific environmental conditions. In recent years, particular emphasis has been placed on the studies of groups rather than single species and this has given rise to the term *ecosystem.* Odum [1] has defined an *ecosystem* as "any entity or natural unit that includes living and nonliving parts interacting to produce a stable system in which the exchange of materials between the living and nonliving parts follows circular paths."

Our concern in this dissertation is with the statistical mechanical approach to modeling of ecosystems. A pioneering work in the field is a model given by Kerner [2]; before we deal with it, and its variants, it seems desirable that we present a broad overview of the different approaches that may be used to study an ecosystem. This is done in the next section, where definitions of some basic ecological terms are also given. This section ends by highlighting the role of mathematical models in an overall context. Section 3 of this chapter is devoted to tracing the evolution of mathematical models of populations, ending with equations that describe a community of a large number of species. In Sec. 4, we try to convey the essence of the framework of statistical mechanics in physics. In Sec. 5, we show this framework at work for a particularly simple system. Section 6 is devoted to motivating the use of statistical mechanics to study populations in an ecosystem. Section 7 delineates the scope of the dissertation and the plan for the following chapters.

1.2 An overview of different approaches to study an ecosystem

(i) Environmental or Habitat approach:

Characteristics of an environment fall into three major categories: (a) physical; (b) chemical; and (c) biotic. In connection with any of these factors, if the presence or absence of a given condition is necessary for sustenance of whatever group, species, etc. that is being considered, such condition is referred to as a limiting factor. Physical factors include light, temperature, wind, fire, soil structure, etc. Chemical factors include the composition of the water (pH, dissolved gases and solids, etc.); the composition of the air (pollutants, water vapor, etc.); the composition of the soil (alkalinity, acidity, presence of various elements and compounds) etc.; Biotic factors are related to food supply and presence and behavior of neighboring organisms (predators, parasites, etc.).

Environments are classified into four major types: (a) *freshwater*; (b) *marine*; (c) *terrestrial*: and (d) *symbiotic*. There are several subdivisions of each. For example, terrestrial environments, called *biomes*, are divided into deserts, tundra, grasslands, savannas, different types of forests, etc.

In the symbiotic environment, the factors of concern are other organisms, rather than nonorganic factors. There are numerous examples where two or more species may inhabit a given area of close proximity in the absence of predators. Usually, each species present will contribute in some fashion to the well being of others.

(ii) Communities approach:

The term *population* is used to describe a group of individuals composed of a single species or of several closely associated species that occupy a definite environmental area. Where all of the populations that occupy a given geographic area are studied, the term *biotic community* is used. The portion of the earth on or in which life exists is termed as the *biosphere*. *Habitat* is defined as that particular environment in which a population lives. For example, a catfish that likes slow-moving streams and lakes is described as having this particular habitat. Obviously, the variation among habitats considering the tens of thousands of life forms is tremendous, although generalizations can be made. *Niche* is a term used to describe the role played by a population within its community and ecosystem. Numerous factors determine niche – eating habits, predaroriness, etc.

The first branch of ecology to develop beyond the stage of life-history study was the description of vegetation. Its basic method is to study the detailed distribution of vascular plants in terms of communities of various types, the pattern and complexity of which depend largely upon the climate and soil. The major contribution of this school is the idea of *seral succession* toward a stable climax. According to this idea, if a new environment is created for terrestrial plants, or an old one drastically changed, the vegetation on it does not remain unchanged for eternity, but rather alters the environment so that it becomes more suitable for some new and different kind of vegetation. Ultimately, a climax vegetation develops which is stable under the prevailing climatic conditions. A closely related concept is that of *climax community*.

Probably the greatest body of coherent ecological theory has been created via studies of populations. The study of field populations has uncovered many interesting phenomena, such as the periodic oscillation of arctic small vertebrate populations and the seasonal changes in abundance of planktonic organisms. These observations have stimulated both laboratory experiments and the development of deductive mathematical theory.

1.3 Evolution of mathematical models of populations

Malthusian model:

It is a useful to consider first the simplest idealized ecosystem and build a quantitative basis for them. The most idealized system is one with a single species in an environment in which the resources are unlimited. It may be assumed that in such a system the growth rate per individual is same for all individuals and is constant with respect to time. It is the one of the simplest ecosystem models and it is assumed in this model that the species are immortal, rate of reproduction is constant with time and the species are uniformly distributed over the space.

The basic equation for this model is

$$\frac{dN}{dt} = \alpha N, \qquad (1)$$

and the solution of the equation is

$$N(t) = N(0)\exp(\alpha\, t), \qquad (2)$$

where

N (t) = population of the species at time t.

α = intrinsic rate constant.

If death as well as birth is also taken into account then the constant α is replaced by r (=b-d). Where 'd' stands for death rate and 'b' stands for birth rate. If $r > 0$, i.e. birth is more than death, it is called biotic potential growth. If $r < 0$, i.e. death is more than birth, then the species becomes extinct in course of time.

Verhulst-pearl model:

The environment is in reality not an unlimited one. The food available to the population is sooner or later going to become limited because of rising population. Hence the growth rate 'α' has to be such that it shows a decrease as the population rises. The simplest possibility that we could consider is if 'α' is replaced by 'α-βN'. This is the logistic approach and reason behind this approach is that at higher density the rate of increase of population decreases because the space and food are limited in any ecosystem. Here the factor 'β' reflects the degree to which the density decreases the rate of increase of the population and significant due to the competition among the individuals of species. When 'N' is small 'α-βN' tends to 'α' and when 'N' is large then,

$$\frac{d(\alpha - \beta N)}{dN}$$

is negative. The basic equation for this approach is

$$\frac{dN}{dt} = (\alpha - \beta N)N \qquad (3)$$

This is the well-known Pearl-Verhulst logistic equation of population growth. And the solution of this is

$$\ln\left(\frac{|N - \alpha / \beta|}{N}\right) = -\alpha \,(t + c) \qquad (4)$$

Two cases arise

1. $N-\alpha/\beta > 0$ i.e. $N>\alpha/\beta$, and
2. $N-\alpha/\beta <0$ I.e. $N<\alpha/\beta$.

In case of $N>\alpha/\beta$ solution becomes

$$N = \frac{\alpha / \beta}{1 - \exp(-\alpha(t + c))} \qquad (5)$$

If t tends to infinity then N tends to α/β.

In case of $N<\alpha/\beta$ the solution is

$$N = \frac{\alpha / \beta}{1 + \exp(-\alpha\,(t + c))} \qquad (6)$$

In this case again N tends to α/β, as time goes to infinity.

According to this approach, the population rises initially as in the Malthusian model, but then the growth rate begins to slow down and the population turns towards its asymptotic value, which is (α/β). This value is the maximum that the population can reach and is therefore called the carrying capacity of the given environment.

Lotka- Voltera model:

We now consider a situation where we have two interacting population in a given ecosystem, say a prey and a predator. The simplest pray predator model given independently by Lotka (1925) and Voltera (1926). Two very general assumptions used in such Pray predator model are;

1) the two populations inhabit the same area, so density is directly proportional to numbers.
2) there is no time lag in the response of either population to changes due to one other.

If 'N_1' represents the population of pray while 'N_2' represents that of predator then fundamental equations for this model are

$$\frac{dN_1}{dt} = \alpha\ N_1 - \beta\ N_1 N_2 \qquad (7)$$

and

$$\frac{dN_2}{dt} = \gamma\ N_2 - \delta\ N_1 N_2, \qquad (8)$$

where α, β, γ and δ are positive quantities. Here α and γ are average rate of growth and decay per individual in the absence of the other species and β and δ are interaction parameters. If the niche is not overlapping the interaction is reduced and the values of β and δ are very small. The sign in the interaction terms in the above equation are based on the expectation that the interaction will generate predators at the cost of pray. These equations are coupled equation so solved by assuming equilibrium point.

At equilibrium point

$$\frac{dN}{dt}=0$$

With this consideration one can show the trajectory in the phase space is given by

$$\frac{(N_1-\gamma/\delta)^2}{A_1{}^2}+\frac{(N_2-\alpha/\beta)^2}{A_2{}^2}=1 \qquad (9)$$

where

$$A_2=\sqrt{\frac{\alpha}{\gamma}.\frac{\delta}{\beta}}.A_1$$

This model shows that both prey and predator populations under go oscillations with constant amplitudes. The behavior is generally referred to as one of the neutral stability. The point (N_1^*, N_2^*) corresponds to the point of neutral stability, where $N_1^*=\gamma/\delta$ and $N_2^*=\alpha/\beta$

Leslie Gower model:

An alternative formulation of prey-predator equations was suggested by Leslie and Gower [3]. The rate equation for Prey (N_1) and Predator (N_2) are given by,

$$\frac{dN_1}{dt}=(\alpha-\gamma\,N_2)\,N_1 \qquad (10)$$

and

$$\frac{dN_2}{dt}=(\delta\ -\eta\,N_2/N_1)\,N_2, \qquad (11)$$

where α, γ, δ and η are positive. We generally use isocline's method to study the system. The 'N_1- isocline' and 'N_2-isocline' are the curve in 'N_1-N_2' phase space on which the time rate of change of prey and predator respectively are zero. Thus on 'N_1-isocline', $dN_1/dt = 0$ and on 'N_2- isocline', $dN_2/dt = 0$. From equation (10) and (11) we can see that

N_1 isocline: $N_1 = \alpha/\gamma$

N_2 isocline: $N_2 = (\delta/\eta)N_1$

Thus both isoclines are straight lines, which intersect at point (N_1^*, N_2^*), where $N_1^* = \alpha/\gamma$ and $N_2^* = \delta/\eta$

At point (N_1^*, N_2^*) both $dN_1/dt = 0$ and $dN_2/dt = 0$. This means that at this point population of prey and predator remains constant with time; hence this point is the equilibrium point. This model takes into account the likely effect of predators per capita growth rate of the relative sizes of interacting population. Thus larger the ratio, N_1/N_2, the smaller the number of prey per predator and consequently, the less rapid the growth of predator population.

Leslie's approach is different from Voltera's approach in the following way. In case of Voltera, predator number depends only upon the density of prey whereas in the case of Leslie it depends upon the number of prey per predator. Voltera relates the rate of increase of predators to the rate at which the prey are being eaten whereas in Lesile's approach there is no relationship between the rate at which predator eats and the rate at which it reproduces.

Holling -Tanner model:

The Holling-Tanner model [3] is slightly more elaborate than that of Leslie and Gower model.

$$\frac{dN_1}{dt} = \alpha\, N_1\left(1 - \frac{N_1}{k}\right) - \frac{\gamma\, N_2}{D + N_1} N_1 \tag{12}$$

$$\frac{dN_2}{dt} = \left(\delta\ - \eta\, N_2 / N_1\right) N_2, \tag{13}$$

where α is intrinsic growth rate of prey, k is the maximum number of prey allowed by the resources of the system, γ is maximum predation rate, D is a constant determining how fast the functional response curve increases at low prey densities. The value of δ and η are same as discussed in Leslie Gower model. In this model the additional term $(\frac{\gamma N_2}{D+N_1} N_1)$ has introduced in place of $\gamma N_2 N_1$ the addition of D+ N_1 corresponds that even N_1 is very large we have still limitations of attack rate of predator result of this population oscillates in closed curve and independent of initial conditioned.

Competing species model (Gauss's model):

If more than one species are present in the ecosystem then interaction takes place not only among the individuals of the same species but also among the individuals of different species. Both of these interactions are generally negative due to the competition between them and predation between the two species, again because of limited food supply and space. In this approach the basic equations are

$$\frac{dN_1}{dt} = \alpha\, N_1 - a_{11} N_1 N_1 - a_{12} N_1 N_2 \qquad (14)$$

and

$$\frac{dN_2}{dt} = \beta\, N_2 - a_{21} N_2 N_1 - a_{22} N_2 N_2 \qquad (15)$$

These are nonlinear, coupled equation and cannot be directly solved.

In case of three species the basic equations can be written down as

$$\frac{dN_1}{dt} = \alpha\, N_1 - a_{11} N_1 N_1 - a_{12} N_1 N_2 - a_{13} N_1 N_3 \qquad (16)$$

$$\frac{dN_2}{dt} = \beta\, N_2 - a_{21} N_2 N_1 - a_{22} N_2 N_2 - a_{23} N_2 N_3 \qquad (17)$$

and

$$\frac{dN_3}{dt} = \gamma\, N_3 - a_{31} N_3 N_1 - a_{32} N_3 N_2 - a_{33} N_3 N_3 \qquad (18)$$

Generalizing to "k" species yields

$$\frac{dN_1}{dt} = r_1 N_1 - a_{11}N_1N_1 - a_{12}N_1N_2 ......a_{1k}N_1 N_k \tag{19}$$

$$\frac{dN_2}{dt} = r_2 N_2 - a_{21}N_2N_1 - a_{22}N_2N_2 ...... a_{2k}N_2 N_k \tag{20}$$

....

$$\frac{dN_k}{dt} = r_k N_k - a_{k1}N_kN_1 - a_{k2}N_kN_2 .....a_{kk}N_k N_k \tag{21}$$

This can be compactly written as

$$\frac{dN_i}{dt} = r_i N_i - \sum_{j=1}^{k} a_{ij}N_iN_j \quad (i = 1,2,...k) \tag{22}$$

The first term on right side shows how particular species grow if it left to itself in a given ecosystem and no other species interact with it. 'r_i' (birth minus death rate) may be positive or negative which result exponential rise or fall respectively. The remaining term express the interaction of species 'i' with all other species 'j' stating that increase or decrease of N_i per unit time is effectively proportional to the number of encounters per second between 'i' and 'j', taken to be measured by the product N_i Nj.

1.4 The essence of Statistical Mechanics

One major problem of physics involves the prediction of the macroscopic properties of matter in terms of the molecules of which it is composed. According to classical physics, this could have been accomplished by a determination of the detailed motion of each molecule and by a subsequent superposition or summation of their effects. The Heisenberg uncertainty principle now indicates that this process is impossible, since we cannot acquire sufficient information about the initial state of the molecules. Even if this were not so, the problem would be practically insoluble because of the extremely large number ($\approx 10^{23}$) of molecules involved in nearly all observations. Many successful predictions can be made, however, by considering only the average, or most probable, behavior of the molecules, rather than the behavior of individuals. This is the method used in statistical mechanics, which is based on the following postulates:

Postulate 1: The existence of microstates
Postulate 2: The principle of equal a priori probabilities of individual microstates
Postulate 3: The law of conservation of energy.
Postulate4:The values of the macroscopic variables when the system is in thermodynamic equilibrium are given by the ensemble averages of the corresponding quantities in the statistical mechanical approach

Let us first remark that the average of a physical quantity in statistical mechanics can be determined in two ways. The system could be followed over a long period of time during which it would sample a large number of microstates, independent of the initial state of the system (this is known as ergodicity). The average of the physical quantity over the long period would give the time-averaged value of the physical quantity. Alternately, one could consider a large collection (ensemble) of identical closed systems prepared under identical conditions and average the physical quantity over all these systems at one instant of time to determine the ensemble average. In practice when estimating an ensemble average, it is convenient to restrict the ensemble to systems containing one and only one copy of each of the microstates accessible to the closed system. The ergodic hypothesis is invoked to assure the equality of ensemble and time averages of physical quantities in equilibrium thermodynamics.

Consider now a small system in thermal equilibrium with a heat reservoir at absolute temperature T. The probability that the system is in the i^{th} microstate is proportional to the Boltzmann factor, i.e.

$$p_i = A \exp(-\beta\varepsilon_i), \qquad (1)$$

where $\beta = 1/kT$, k is the Boltzmann constant and ε_i is the energy of the i^{th} microstate. Since the small system must be in one of its possible microstates, one has

$$\sum_i p_i = 1, \qquad (2)$$

this implies

$$A = \frac{1}{\sum_i \exp(-\beta\varepsilon_i)} \qquad (3)$$

Thus

$$p_i = \frac{\exp(-\beta\varepsilon_i)}{Z}, \qquad (4)$$

where

$$Z = \sum_i \exp(-\beta\varepsilon_i) \qquad (5)$$

is the *partition function.* The sum in Eq. (5) is over all the microstates of the system.

Postulate 2 now implies that the average of any variable α, when the system is in thermal equilibrium with the heat reservoir, is

$$\bar{\alpha} = \frac{\sum_i \alpha_i \exp(-\beta\varepsilon_i)}{Z}, \qquad (6)$$

where α_i is the value of the variable α when the system is in its i^{th} microstate. The partition function of a system is the most important construct of statistical mechanics, because its knowledge enables one to deduce various thermodynamic properties of the system such as the internal energy, pressure, volume, entropy, Helmholtz free energy, etc.

1.5 An example of statistical mechanics at work: A 2-level system

Let a system have accessible to it only two non-degenerate microstates, which have energies ε_0 and ε_1. Let the system be in equilibrium with a heat reservoir at temperature, T. The partition function of the system is then given by Eq. (5) as

$$Z = \exp(-\beta\varepsilon_0) + \exp(-\beta\varepsilon_1). \qquad (7)$$

Applying Eq. (4), we have

$$p_0 = \frac{\exp(-\beta\varepsilon_0)}{\exp(-\beta\varepsilon_0) + \exp(-\beta\varepsilon_1)}$$

$$= \frac{1}{1+\exp(-\theta/T)} \qquad (8)$$

$$p_1 = \frac{\exp(-\beta\varepsilon_1)}{\exp(-\beta\varepsilon_0) + \exp(-\beta\varepsilon_1)} \exp(-\beta\varepsilon_0)$$

$$= \frac{\exp(-\theta/T)}{1+\exp(-\theta/T)} \qquad (9)$$

$$\theta = (\varepsilon_1 - \varepsilon_0)/k. \qquad (10)$$

Clearly, both p_0 and p_1 depend upon the temperature of the reservoir. For $T/\theta = 0.01$, 0.1 and 0.2, p_1 is equal to $3.72\text{x}10^{-44}$, $4.54\text{x}10^{-5}$ and $6.69\text{x}10^{-5}$, respectively; for $T/\theta=10$, p_1 equals 0.475. According to Eq. (6), the ensemble average of the system is

$$\bar{\varepsilon} = \sum_i \varepsilon_i p_i$$

$$= \frac{\varepsilon_0 + \varepsilon_1 \exp(-\theta/T)}{1+\exp(-\theta/T)} \qquad (11)$$

For $T=\theta$, for instance, Eq. (11) gives

$$\bar{\varepsilon} = 0.731\varepsilon_0 + 0.269\varepsilon_1. \qquad (12)$$

1.6 Motivation for use of statistical Mechanics to study the ecosystem

An ecosystem is a complex system consisting of the biotic and abiotic components of the environment. Different organisms in it are continuously interacting among themselves and with the other organisms. Further, the nature of these interactions also depends on the abiotic component of the environment. For example, there are hundreds of species of butterflies and moths in a forest and thousands of species of other insects; and there are as many different environments within one forest as there are kind of insects. One could, in principle, set up differential equations of the kind noted above and try to solve them in order to determine how different species evolve in time. But the set of the differential equations is very large in number (as number of species is very large); furthermore, these equations are coupled, nonlinear equations, and then there is the intrinsic uncertainty in fixing the values of the different interaction parameters. Therefore prognosis of the system via this approach is a hopeless task.

The above situation is reminiscent of the statistical mechanical approach, which deals with systems comprising enormous numbers of particles in the Physics. Just as we can derive results of the kinetic theory of gases/thermodynamics by the application of statistical mechanics, we could attempt to unravel the broad features of the dynamics of an ecosystem via a suitable adaptation of the methods of statistical mechanics.

1.7 Scope of the dissertation

In chapter 2, we discuss the Kerner's statistical mechanical approach to modeling an ecosystem. Essentially, this involves casting the Volterra's equations for k-species into the Newtonian form and showing that the system satisfies the Liouville theorem. One then shows that the Kerner approach admits of a conserved quantity, which is identified as the biomass. We also consider the micro canonical ensemble, which yields a result equivalent to the equipartition of energy, and the canonical ensemble which yields the result that different components of the ecosystem are characterized by the same "temperature".

In chapter 3, The Kerner method is applied to Gompertz model. This is the study carried out by Gunasekran and Pande [4]. The motivation for this work was to obtain theoretically a lognormal distribution for the abundance of a species in a multi-species ecosystem, empirical evidence for which was noted by Preston [5]. Notable features of this derivation were the use of the anti-symmetry property of the community matrix, and the result that the different components of the system are characterized by the same "temperature". This chapter also deals with the work of Sitaram and Varma [6], who also used the property of anti-symmetry of the community matrix and obtained a log-normal distribution, but showed that the system is characterized by not one, but n/2 "temperatures". This chapter finally deals with a study carried out by Singh and Pande [7], in which they obtain the log-normal distribution without imposing the restriction that the community matrix be anti-symmetric,

In chapter 4 we try to examine the experimental evidence for and against the lognormal distribution for the abundance of a species in an ecosystem based on the field studies carried out by Saunders [8], Dirks [9], and Williams [10].

Chapter 2

Chapter 2

Kerner's Statistical Mechanical Approach to Modeling an Ecosystem

2.1 Introduction

In any ecosystem the number of species is very large. The first systematic attempt to study the structural patterns for species in an ecosystem was made by Corbet [10], Fisher [10] and Williams [10], and later by Preston [5]. On the basis of empirical data, they found mathematical relations between the number of species and the number of individuals. Corbet, Fisher and Williams [10] in 1943 were led to the Maxwell-Boltzmann distribution, while Preston obtained a lognormal distribution. In the wake of these findings, it became of interest to investigate if theoretical models could yield these results. One possibility for this purpose was to invoke the equations embodying interactions of different species. Since equations describing the interactions of a multi-species system are coupled, nonlinear differential equations, and are very large in number, solving them is a very difficult and impractical proposition. If we confine ourselves to the study of an ecosystem in the equilibrium state, we may try to tackle the problem through techniques of statistical mechanics. This technique enables us to calculate some conserved quantities of physical interest for the equilibrium state without actually solving the equations of an ecosystem model.

The first attempt to develop a statistical mechanics of biological association, taking Voltera's model as model for the system, was developed by Kerner [2]. S. Takeanka

[11] had already observed that the Volterra model [12] can be cast in the form of Hamilton's equations, and hence satisfies Liouville's theorem. In his work Kerner [2] sought to determine constants of motion for the ecosystem analogous to energy of a physical system, and tried to relate them to different parameters of the system.

2.2 Casting Volterra's ecosystem equations into the Newtonian form

Starting with Voltera's equation of model of the ecosystem, Kerner [2] derived an equation of motion for the ecosystem.

According to Voltera,

$$\frac{dN_r}{dt} = \varepsilon_r N_r + \frac{1}{\beta_r} \sum_s \alpha_{sr} N_s N_r \qquad (1)$$

The term $\varepsilon_r N_r$ tells about how each species propagates if it left to itself in a given environment and no other species interacts with it. This is nothing but Malthusian assumption that provides exponential growth of the population depending upon birth & death rate. The 2nd term of right hand side express the interaction of species 'r' with all other species 's' stating that increase or decrease of 'N_r' per unit time is effectively proportional to the number of encounters per unit time between 'r' and any 's', taken to be measured by the product '$N_s N_r$'.

Kerner [2] assumed that the interaction parameters α_{ij} be antisymmetric with respect to their indices 'i' & 'j', i.e. $\alpha_{ij}= -\alpha_{ji}$. This simply means for i=j, $\alpha_{ii}=\alpha_{jj}=0$. This assumption of Kerner was based on binary-prey-predator system model in which, if 'r' gains because of encounter then's' must lose. Kerner was interested in stationary state and took an equilibrium value of 'N_r' is 'q_r'.

At equilibrium sate,

$$\frac{dN_r}{dt} = 0$$

Hence applying equilibrium condition in equation (1) we have

$$\varepsilon_r q_r + \frac{1}{\beta_r}\sum_s \alpha_{sr} q_r q_s = 0$$

$$\Rightarrow \varepsilon_r \beta_r + \sum_s \alpha_{sr} q_s = 0 \qquad (2)$$

On rewriting the equation (1) we get

$$\frac{\beta_r}{N_r}\frac{dN_r}{dt} = \varepsilon_r \beta_r + \sum_s \alpha_{sr} N_s \qquad (3)$$

Putting the value of $\varepsilon_r\beta_r = -\Sigma\alpha_{sr}q_s$ from (2) we got,

$$\frac{\beta_r}{N_r}\frac{dN_r}{dt} = \sum_s \alpha_{sr}(N_s - q_s) \qquad (4)$$

Kerner introduced new dependent variable,

$$v_r = \ln\frac{N_r}{q_r} \qquad (5)$$

$$\Rightarrow \frac{dv_r}{dt} = \frac{q_r}{N_r}.\frac{1}{q_r}.\frac{dN_r}{dt} = \frac{1}{N_r}.\frac{dN_r}{dt}$$

From equation (4) we got,

$$\beta_r \dot{v}_r = \sum_s \alpha_{sr} q_s (e^{v_s} - 1) \qquad (6)$$

Where $\dot{v}_r = \frac{dv_r}{dt}$

This equation is analogous to Newton's law of motion.

2.3 Statistical mechanics of the ecosystem: The Liouville theorem

Kenner introduced the new variable ($v_r = \ln N_r/q_r$) to serve the Lioville's theorem. Considering a large number of copies, a Gibbs ensemble of biological association each of the same character and each controlled by the same differential equation (6). The initial values of all v_r are different for all species. Kerner has taken to be sufficiently numerous points in the Cartesian space of the v_r (phase space) so considered that the point constitute a fluid of, say, density $\rho(v_1,v_2,\ldots v_n)$ at a point $(v_1,v_2,\ldots,v_n)$ and velocity, $V(\dot{v}_1,\dot{v}_2,\ldots\dot{v}_n.)$, at this point. Therefore from the hydro dynamical equation of continuity we have,

$$\frac{\partial\rho}{\partial t}+div\rho\, V=0 \qquad (7)$$

$$\Rightarrow\frac{\partial\rho}{\partial t}+\Sigma\frac{\partial\left(\rho\,\dot{v}_r\right)}{\partial v_r}=0$$

$$\Rightarrow\frac{\partial\rho}{\partial t}+\Sigma\dot{v}_r\frac{\partial\rho}{\partial v_r}+\Sigma\rho\frac{\partial\dot{v}_r}{\partial v_r}=0$$

As $\dot{v}_r$ is independent of v_r [from equation (6)] we got,

$$\Sigma\rho\frac{\partial\dot{v}_r}{\partial v_r}=0$$

$$\therefore\frac{\partial\rho}{\partial t}+\Sigma\dot{v}_r\frac{\partial\rho}{\partial v_r}=0 \qquad (8)$$

This is the Liouville's theorem [18] of the conservation of density in phase. This simply means that as one goes along with the motion of one system point, the density in neighborhood remains invariable. This shows that the Voltera model follows Liouville's theorem.

2.4 The existence of a conserved quantity in the Kerner model: The biomass

Now Kerner tried to find out a conserved quantity, called the constant of motion. By multiplying $q_r[\exp(v_r)-1]$ to equation (6) and sum over all r, we have

$$\sum_r q_r\left(e^{v_r}-1\right)\beta_r \dot{v}_r = \sum_r\sum_s \alpha_{sr} q_s q_r \left(e^{v_r}-1\right)\left(e^{v_s}-1\right) \quad (9)$$

And, as interaction parameter 'α_{sr}' is antisymetric (i.e. $\alpha_{sr} = -\alpha_{rs}$), he got

$$\sum_r \beta_r q_r \dot{v}_r \left(e^{v_r}-1\right) = 0 \qquad (10)$$

$$\Rightarrow \sum_r \int \left[\beta_r q_r \frac{dv_r}{dt}\left(e^{v_r}-1\right)\right] dt = \text{Constant}$$

$$\Rightarrow \sum_r \int \left[\beta_r q_r \frac{d\left(e^{v_r}\right)}{dt} - \beta_r q_r \frac{d\left(v_r\right)}{dt}\right] dt = \text{Constant}$$

$$\Rightarrow \sum_r \int \left[\frac{d\left(\beta_r q_r e^{v_r}\right)}{dt} - \frac{d\left(\beta_r q_r v_r\right)}{dt}\right] dt = \text{Constant}$$

$$\Rightarrow \sum \left(\beta_r q_r e^{v_r} - \beta_r q_r v_r\right) = \text{Constant}$$

Kerner said this constant 'G'

$$\sum_r \beta_r q_r \left(e^{v_r} - v_r\right) = G \qquad (11)$$

As 'G' is a constant therefore dG/dt = 0

Thus Kerner got a constant 'G' which is a conserved quantity and analogous to the total energy of an isolated physical system. This 'G' is a universal single valued constant of motion.

2.5 Micro canonical ensemble: The equivalent of the "equipartition" theorem for an ecosystem

Adopting the proposition from Gibbs ensemble theory [13] for physical system, Kerner defined the micro canonical ensemble in phase space for the system under consideration in terms of density function,

$$\rho = \rho_0 \delta (G - G_0) \qquad (12)$$

where 'δ' is the standard Dirac delta function and 'ρ_o' is a numerical constant. Hence the volume of the ensemble phase space can be written as,

$$\tau = \int \rho \, d\tau$$

where '$d\tau$' is the infinitesimal volume in the Gibbs space, and given by

$$d\tau = dv_1, dv_2 \ldots . dv_n$$

Here in this case,

$$d\tau = ds.dn$$

where,
ds = element of area on a surface of constant 'G'
dn = increment of length normal to the surface.

As in statistical mechanics the equilibrium value is just the average value of quantity over the micro canonical ensemble. Hence micro canonical average of quantity 'f' is given by,

$$\langle f \rangle = \frac{\int f \rho_0 \delta (G - G_0) \, d\tau}{\int \rho_0 \delta (G - G_0) \, d\tau} \qquad (13)$$

$$\Rightarrow \langle f \rangle = \frac{\int f \rho_0 \delta (G - G_0) \, ds \, dn}{\int \rho_0 \delta (G - G_0) \, ds \, dn}$$

Now the gradient of 'G' is given by,

$$\nabla G = \frac{dG}{dn}\hat{n} \qquad (14)$$

where '$\hat{n}$' is a unit vector along '$\vec{dn}$'

Now taking scalar product of '∇G' with '$\vec{dn}$', we have,

$$\nabla G.\vec{dn} = \frac{dG}{dn}\hat{n}.\vec{dn}$$

$$\Rightarrow |\nabla G|\, dn = dG$$

$$\Rightarrow dn = \frac{dG}{|\nabla G|} \qquad (15)$$

Here dG = difference in 'G values' of two neighboring constant 'G surface'. Putting the above substitution we have,

$$\langle f\rangle = \frac{\int f\rho_0\delta\left(G-G_0\right)ds.\frac{dG}{|\nabla G|}}{\int \rho_0\delta\left(G-G_0\right)ds.\frac{dG}{|\nabla G|}}$$

$$\Rightarrow \langle f\rangle = \frac{\int_{G_0} f\frac{ds}{|\nabla G|}}{\int_{G_0}\frac{ds}{|\nabla G|}} \qquad (16)$$

Now Kerner took the function 'f' to be

$$T_r = v_r\frac{\partial G}{\partial v_r}$$

$$\Rightarrow T_r = \beta_r q_r\left(\exp(v_r)-1\right)v_r$$

$$\Rightarrow T_r = \beta_r q_r \left(\frac{N_r}{q_r} - 1 \right) \ln \frac{N_r}{q_r}$$

As,

$$\nabla G = \Sigma \frac{\partial G}{\partial v_r} \hat{v}_r$$

And also,

$$\nabla G = |\nabla G| \hat{n}$$

$$\therefore \hat{n} . \hat{v}_r = \frac{\partial G}{\partial v_r} / |\nabla G|$$

$$\Rightarrow \frac{\partial G}{\partial v_r} . \frac{ds}{|\nabla G|} . v_r = v_r \left(\hat{n} . \hat{v}_r \right) ds$$

$$v_r \frac{\partial G}{\partial v_r} . \frac{ds}{|\nabla G|} = \hat{n} . \vec{v}_r \, ds \qquad (17)$$

From equation (16) we have,

$$\langle T_r \rangle = \frac{\int v_r \frac{\partial G}{\partial v_r} . \frac{ds}{|\nabla G|}}{\int \frac{ds}{|\nabla G|}}$$

From equation (17),

$$\langle T_r \rangle = \frac{1}{\Delta_0} \int \hat{n} . \vec{v}_r \, ds \qquad (18)$$

where

$$\int \frac{ds}{\nabla G} = \Delta_0$$

From Gauss' divergence theorem;

$$\langle T_r \rangle = \frac{1}{\Delta_0} \int div \, \vec{v_r} \, d\tau \qquad (19)$$

$$\Rightarrow \langle T_r \rangle = \frac{\tau_0}{\Delta_0} \qquad (20)$$

where 'τ_0' $= \int div \, \vec{v_r} \, d\tau$, i.e. volume enclosed by 'surface G_0'

This result is analogous to the equipartition theorem of physics. It means the total 'T' of the biological association in the mean is equally distributed amongst all species as in the case of statistical mechanics. The quantity 'T' will be recognized to be a loose analog of kinetic energy

If Kerner took the function, f, to be $D_r = \partial G/\partial v_r$ ($= \beta_r q_r \left({N_r}/{q_r} - 1 \right)$) then he got $\langle D_r \rangle = 0$. This means that the mean of 'N_r' is 'q_r'. Now integrating equation (4) over the limit '0' to 't' we have;

$$\frac{\beta_r}{t} \log \frac{N_r(t)}{N_r(0)} = \varepsilon_r \beta_r + \Sigma \alpha_{sr} \left(\frac{1}{t} \int_0^t N_0(t) dt \right)$$

$$\Rightarrow \frac{\beta_r}{t} \log \frac{N_r(t)}{N_r(0)} = \varepsilon_r \beta_r + \Sigma \alpha_{sr} q_s \qquad (21)$$

This equation is analogous to Volterra mechanics as shown in equation (2). So Kerner suggested that the system under study might be Ergodic. He finally concluded that all species, not just same as proven by Voltera, must in general exhibits undamped oscillation; for the only alternative to continual oscillation is a tendency of the 'N' towards finite limits, and such tendency in any 'N' clearly will not permit the system point to return near a given starting point.

2.6 Canonical ensemble: Different components of an ecosystem have the same "temperature"; the Kerner distribution

In order to find out a temperature like quantity, θ, corresponding to the conserved quantity, G, Kerner defined canonical ensemble for the system via density function with the help of Gibbs theory. And the density function was given by,

$$\rho = \exp\left(\frac{\psi - G}{\theta}\right),$$

where 'Ψ' is the free energy of a system in the thermodynamic equilibrium. This distribution can also be written as,

$$\rho = K\exp\left(\frac{-G}{\theta}\right), \qquad (22)$$

where 'K' is a numerical constant. Now the distribution of phase points is characterized by the constant 'θ' rather than by 'G'. This 'θ' is analogous to the temperature function as in the case of statistical mechanics.

Again Kerner took the function '$D_r = \partial G/\partial v_r$' and calculated the canonical average for this in same manner as in the case of micro canonical ensemble and got zero. Which means that ensemble average of 'N_r' is again 'q_r'. Similarly he got the canonical average for the quantity 'T_r' was 'θ'.

$$\langle T_r \rangle = \frac{\int v_r \dfrac{\partial G}{\partial v_r}.\exp\left(-G/_\theta\right) d\tau}{\int \exp\left(-G/_\theta\right) d\tau} \qquad (23)$$

$$= \theta$$

This gives the earlier result of equpartition of 'T' as Kerner got in case of micro canonical ensemble. This 'T' also gave an insight into the meaning of 'temperature', θ, of biological association.

Now Kerner took the function as 'D_r^2', computed its canonical average and got a more perspicuous view of 'θ'.

$$\langle D_r^2 \rangle = \frac{\int \left(\frac{\partial G}{\partial v_r} \right)^2 \exp\left(-G/_\theta\right) d\tau}{\int \exp\left(-G/_\theta\right) d\tau} \tag{24}$$

$$= \theta \beta_r q_r$$

$$\Rightarrow \beta_r^2 q_r^2 \overline{\left(\frac{N_r}{q_r} - 1 \right)^2} = \theta \beta_r q_r$$

$$\Rightarrow \theta = \frac{\beta_r}{q_r} \overline{(N_r - q_r)^2} \tag{25}$$

Therefore Kerner concluded that the temperature, θ, is the measure of the mean square deviation of the population from their stationary state value, and vice-versa. A zero temperature means a completely 'quite' stationary state of biological association which is a kind of indicator of the level of extinction from stationary state.

The ensemble density in equation (22) gives the probability density that a given system is found in a state ($v_1, v_2, \ldots, v_n$). It enabled Kerner to calculate the probability that a particular species 'r' will have its n_r where $n_r = N_r / q_r$, in the range n_r and $n_r + dn_r$. This probability is given by,

$$P(n_r) dn_r = \frac{n_r^{x_r - 1} e^{-x_r n_r}}{x_r^{-x_r} \sqrt{x_r}} \tag{26}$$

This equation is analogous to the Maxwell-Boltzmann distribution law. This Kerner distribution is exactly the χ^2 – distribution for large value of 'θ', which is exactly the distribution of the intrinsic abundance of a species assumed by Cobert, Fisher, and Williams [10].

Thus Kerner statistical approach is based on the assumption of antisymetric relationship of so called community matrix. This assumption is quite restrictive which is only applicable to even-numbered species ecosystem and the distribution for the relative abundance of species is Maxwell-Boltzmann distribution not the Preston lognormal distribution.

Chapter 3

Chapter 3

Statistical mechanics of the Gompertz model via the Kerner method (with and without the assumption of the anti-symmetry of the community matrix)

3.1 Introduction

3.2 Statistical mechanics of the Gompertz model: Gunasekran and Pande's approach

3.3 Comments on Gunasekran and Pande's work

3.4 Statistical mechanics of the Gompertz model: Sitaram and Varma's approach

3.5 Statistical mechanics of the Gompertz model: Singh and Pande's approach

3.1 Introduction

By using the same approach as used by Kerner [2], Gunasekaran and Pande [4], and later Sitaram and Varma [6] derived Preston's Log-normal distribution pattern for the relative abundance of species. They took Gompertz model for the ecosystem instead of the Voltera model considered by Kerner [2].

According to Gompertz model,

$$\frac{dN_i}{dt} = \varepsilon_i N_i + \sum_j \alpha_{ij} N_i \log N_j \,,$$

where 'N_i' is the population of i^{th} species and 'ε_i' and 'α_{ij}' are real constant similar to those used in Voltera mechanics. The first term of RHS of the equation (1) represents the natural growth or decay of population i.e. Malthusian growth and the second term is due to interaction among the individuals.

Substitution of '$\log N_j$' in Voltera model in the place of 'N_j' makes the Gompertz model in exactly solvable mode. It may be emphasized that the new form (Gompertz model) continues to be nonlinear, as the Voltera model is. The advantage accruing from this new form is that a transformation to new variable, $x_i = \log N_i$, puts the model into a linear and solvable form.

We note that Gunasekaran and Pande [4], and later Sitaram and Varma [6] derived Preston's Log-normal distribution pattern for the relative abundance of species by assuming the anti-symmetry property of the community matrix. The studies of these authors are dealt with in sections 3.2-3.3 and 3.4, respectively. In the final section of this chapter, we report the work of Singh and Pande [7], where the anti-symmetry property of the community matrix is done away with, and yet the distribution obtained is a log-normal one. Before we discuss these models, it seems desirable to collect below the rudiments of the log-normal distribution.

Log-normal distribution:

Normal distribution or Gaussian distribution [14] is one of the examples of continuous distribution. It is defined by the equation,

$$Y = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(X-\mu)^2}{\sigma^2}\right),$$

where 'μ' = mean and 'σ' = standard deviation.

When the variable 'X' is expressed in terms of standard units z = (X - μ)/σ, the equation for normal distribution becomes

$$Y = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)$$

This is called the standard form of normal distribution and the 'z' is normally distributed with mean zero and variance one.

Now 'X' is said to have a log-normal distribution if 'ln X' is normally distributed. i.e. if 'X' is replaced by 'ln X' then the distribution obtained is called log-normal distribution.

The probability distribution is now therefore is given by

$$\ln Y = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{\left(\ln X - \ln\mu_g\right)^2}{\sigma_g^2}\right)$$

where 'μ_g' is called geometric mean and 'σ_g' is geometric standard deviation, and they are given by

$$\ln \mu_g = \frac{\Sigma n_i \ln X_i}{\Sigma n_i}$$

and,

$$\ln \sigma_g = \left[\frac{\Sigma n_i \left(\ln X_i - \ln \mu_g \right)^2}{\Sigma n_i - 1} \right]^{1/2}$$

3.2 Statistical mechanics of the Gompertz model: Gunasekran and Pande's approach

Gunasekaran and Pande [4] attempted to construct a statistical mechanics of the ecosystem in which the interactions among the individuals are of the form of Gompertz model [15]. From Gompertz model:

$$\frac{dN_i}{dt} = \varepsilon_i N_i + \sum_j \alpha_{ij} N_i \log N_j \qquad (1)$$

Gunasekaran and Pande [4] introduced the transformation,

$$x_i = \log N_i \qquad (2)$$

On differentiation of equation (2) we get,

$$\frac{dx_i}{dt} = \frac{1}{N_i} \frac{dN_i}{dt}$$

In order to construct a statistical mechanics of the ecosystem, Gomportz model can be written as by making the use of equation (1)

$$\frac{dx_i}{dt} = \varepsilon_i + \sum_j \alpha_{ij} x_j \qquad (3)$$

Now they put the constraint, $\alpha_{ij} = -\alpha_{ji}$, i.e. constraint of antisymetry as put by Kerner in his work.

At equilibrium state,

$$\frac{dx_i}{dt} = 0$$

$$\therefore \varepsilon_i + \sum_j \alpha_{ij} x_j^* = 0 \tag{4}$$

From equation (3) and equation (4) we have,

$$\frac{dx_i}{dt} = \sum_j \alpha_{ij}(x_j - x_j^*) \tag{5}$$

Multiplying both side by $(x_i - x_i^*)$ and summing over 'j' we get,

$$\sum_i x_i (x_i - x_i^*) = \sum_{ij} \alpha_{ij}(x_i - x_i^*)(x_j - x_j^*)$$

And applying antysymetry condition, RHS of above equation vanishes and the equation (5) gives a conserved quantity, G.

$$G = \sum_i \left(\frac{1}{2} x_i^2 - x_i x_i^* \right) \tag{6}$$

This 'G' is the constant of motion of the ecosystem, which was established by Gunasekaran and Pande [4].

They also defined the micro canonical ensemble as defined by the Kerner [2],

$$\rho = \rho_0 \delta (G - G_0),$$

where 'ρ_o' is a constant 'δ' is the Dirac delta function. The ensemble average of any function, f, is given by,

$$\langle f \rangle = \frac{\int f \rho_0 \delta (G - G_0)\, d\tau}{\int \rho_0 \delta (G - G_0)\, d\tau}$$

Micro canonical ensemble is some times more useful instead of the canonical ensemble, which can be defined easily for a smaller part of the system as well. Gunasekaran and Pande [4] got the same result for function 'T_r' and 'D_r' as obtained by Kerner. For more realistic situation we now go over to canonical ensemble. In canonical ensemble fluctuation/variation in 'G' is allowed and hence a temperature,θ, (called ecotemperature) is defined corresponding to the value of 'G'.

The density function for such system was defined by Gunasekaran and Pande [4], which was given by,

$$\rho = \exp\left(-G/\theta\right)$$

Hence the canonical average of any function, f, is given by

$$\langle f \rangle = \frac{\int f \exp\left(-G/\theta\right) d\tau}{\int \exp\left(-G/\theta\right) d\tau}$$

In order to give a meaning full interpretation to 'θ' they took the function, f, as same as Kerner has taken for the same.

$$f = \left(\frac{\partial G}{\partial x_i}\right)^2$$

$$\therefore f = \left(x_i - x_i^*\right)^2$$

Now ensemble average for canonical ensemble can be written as,

$$\langle f \rangle = \frac{\int \left(\frac{\partial G}{\partial x_i}\right)^2 \exp\left(-G/\theta\right) d\tau}{\int \exp\left(-G/\theta\right) d\tau}$$

$$= \theta$$

$$\therefore \theta = \overline{\left(x_i - x_i^*\right)^2}$$

i.e. the 'θ ' is the measure of mean square deviation of 'x' from its equilibrium value 'x_i^*'. This is analog to what has obtained by Kerner [2] for the same. In order to find out the probability, ($P(x_i)dx_i$), for any species 'i' to have its 'x_i' in the range 'x_i' and 'x_i+dx_i, Gunasekaran and Pande [4] considered a phase integral, Z, which is given by,

$$Z = \int \exp\left(-G/\theta\right) dx_1, dx_2 \ldots\ldots dx_n$$

As $G = \Sigma G_i$,

$$\therefore Z = \int \exp\left(-\frac{G_1 + G_2 + \ldots + G_n}{\theta}\right) dx_1 . dx_2 \ldots\ldots dx_n$$

$$= \prod_i \int \exp\left(-G_i/\theta\right) dx_i = \prod_i Z_i$$

Now we can write,

$$Z_i = \int \exp\left(-G_i/\theta\right) dx_i$$

$$Z_i = \int \exp\left(-\frac{(x_i^2 - 2x_i x_i^*)}{2\theta}\right) dx_i \text{ [From equation (6)]}$$

$$= \sqrt{2\pi\theta}\ \exp\left[\frac{\left(x_i^*\right)^2}{2\theta}\right]$$

where i = 1, 2…n.

Hence probability distribution for any species 'i' to have its 'x_i' in the interval between 'x_i' and '$x_i + dx_i$' is given by,

$$P(x_i)\, dx_i = \frac{\exp\left(-G_i/\theta\right)}{Z_i}\, dx_i$$

$$\Rightarrow P(x_i)\,dx_i = \frac{\exp\left(\dfrac{-(x_i^2 - 2x_i x_i^*)}{2\,\theta}\right)}{\sqrt{2\pi\theta}\,\exp\left(\dfrac{(x_i^*)^2}{2\,\theta}\right)}\,dx_i$$

$$\Rightarrow P(x_i)\,dx_i = \frac{1}{\sqrt{2\pi\theta}}\exp\left(\frac{-(x_i - x_i^*)^2}{2\,\theta}\right)dx_i$$

Now putting the value of 'x_i' as we assumed i.e. $x_i = \log N_i$, we get

$$\Rightarrow P(\log N_i)\,dN_i = \frac{1}{\sqrt{2\pi\theta}}\exp\left(-(\log N_i - \log N_i^*)^2/\ 2\theta\right)dN_i$$

If we put $\theta = \sigma^2$, varience as $\theta = \overline{\left(x_i - x_i^*\right)^2}$, we get,

$$\Rightarrow P(\log N_i)\,dN_i = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(\log N_i - \log N_i^*)^2\right)dN_i$$

This is the standard log-normal distribution for the species N_i. Thus the result for the probability distribution of the species N_i is similar to the Preston's result [5]; only difference is that the logarithm here are to the base 'e' instead of to the base '2'.

3.3 Comments on Gunasekran and Pande's work

One of the important things of Gunasekran and Pande's approach is that the stable variances unlike in most other approaches, which generate log-normal distributions with even increasing variance with time. They used the same approach as used by Kerner [2] and made the use of Gompertz model for the ecosystem and got the log-normal distribution for intrinsic abundance of species. The idea of equipartition prevails in their approach as well, which is evidenced by the properties of $\langle T_r \rangle$ that enabled them to define canonical ensemble for the ecosystem. They also tried to show that that the ecosystem admits of the ergodic hypothesis as well. The ergodic hypothesis ' first advanced by Boltzmann, states that time average of a macroscopic quantity under equilibrium condition is the same as an ensemble average. Because they also got $\langle D_r \rangle = 0$ that shows ensemble average of 'x_i' is equal to the equilibrium value of 'x_i' i.e. 'x_i^*'

3.4 Statistical mechanics of the Gompertz model: Sitaram and Varma's approach

The result obtained by Gunasekaran and Pande [4] is most encouraging. They computed one constant of motion with antisymetric consideration. In 1984 Sitaram and Varma [6] in their work told that the statistical mechanics of the ecosystem is not straight forward as Gunasekaran and Pande [4] envisage in their work in 1982. Sitaram and Varma [6] set up the statistical mechanics of the system and tried to find out the 'n/2' independent temperatures instead of one constant temperature obtained by Kerner [2] and later by Gunasekaran and Pande [4]

Sitaram and Varma [6] again took the Gompertz model for the n-interacting species in the ecosystem. And the Gomportz model is,

$$\frac{dN_i}{dt} = \varepsilon_i N_i + \sum_j \alpha_{ij} N_i \log N_j \qquad (1)$$

They substituted,

$$x_i = \ln\left(\frac{N_i}{N_i^*}\right) \qquad (2)$$

This substitution is similar to what was done by Kerner in his work. Here N_i^* corresponds to the equilibrium value.

At equilibrium,

$$\frac{dN_i}{dt} = 0$$

$$\therefore \varepsilon_i + \sum_j \alpha_{ij} \log N_i^* = 0$$

$$\Rightarrow \varepsilon_i = -\sum_j \alpha_{ij} \log N_i^* \qquad (3)$$

Here it is assumed that N_i^*, equilibrium population of any species 'i' is positive and real. Now,

$$\frac{dx_i}{dt} = \frac{1}{N_i}\frac{dN_i}{dt} = \frac{1}{N_i}\left(\varepsilon_i N_i + \sum_j \alpha_{ij} N_i \log N_j^*\right)$$

$$\Rightarrow \frac{dx_i}{dt} = \varepsilon_i + \sum_j \alpha_{ij} x_j$$

With the help of equations (1), (3) and (4) we got,

$$\Rightarrow \frac{dx_i}{dt} = \sum_j \alpha_{ij} x_j$$

If 'Ω' is the matrix which diagonalizes 'α' then in terms of variable, $Z = \Omega^{-1}x$.
The solution of this equation is,

$$Z_i(t) = Z_i(0)\exp(\lambda_i t), \qquad (5)$$

where 'λ_i' are eigenvalues of 'α' which were supposed by Sitaram and Varma [6] to be distinct and hence the system possesses (n-1) constants of motion which is given by,

$$c_{j-1} = Z_1^{\lambda_j} Z_j^{-\lambda_1}, \qquad (6)$$

where j = 2, 3…n

Again they took the antisymmetric interactions of the ecosystem that was considered by Kerner [2] and then Gunasekaran and Pande [4]. Under this antisymmetric condition the all eigenvalues will be imaginary and arranged so that

$$\lambda_{2j-1} = iv_j, \quad and \quad \lambda_{2j} = -iv_j; \text{j =1, 2…k (n = 2k)}$$

i.e. $\lambda_1 = iv_1$, $\lambda_2 = -iv_1$, $\lambda_3 = iv_2$, $\lambda_4 = -iv_2$, and so on.

They now took $Z_j = |Z_j| \exp(i\theta_j)$ and wrote equation (6) as,

$$|Z_{2j-1}|^2 = |Z_{2j}|^2 = A_j \text{; Where j = 1, 2...n} \tag{7}$$

$$\Rightarrow \exp\left[i\left(\theta_j - \theta_1\delta_j\right)\right] = B_j\left(A_1^{\delta_j} / A_j\right)^{1/2} \text{; j = 2, 3...n} \tag{8}$$

here A_j are real and B_j are complex constants with $\delta_j = v_j/v_1$. One of the important things is that the value of 'n' is '2k' because of antysemetric condition which is applicable only to the even-numbered species.

With the help of equation (7) and (8) they found out the volume in the phase space occupied by the ensemble which was given by

$$\tau\left[\left(A_j\right)\right] = \int\left[\sum_{j=1}^{n} dz_i dz_i^*\right]^{1/2}$$

$$= \Delta\sigma$$

where 'Δ' is constant and 'σ' is given by,

$$\sigma = \left[\sum_{j=1}^{k} v_j^2 A_j\right]^{1/2}$$

And then they defined entropy of the system that is given by

S = ln τ = ln Δ + ln σ,

and the 'k-temperatures', 'σ_j' are defined by,

$$\frac{1}{\sigma_j} = \frac{\partial S}{\partial A_j} = \frac{v_j^2}{2\sigma^2}$$

The density of phase space in the canonical ensemble is then

$$\rho \propto \exp\left(-\sum_{j=1}^{k} A_j / \sigma_j\right)$$

And with this consideration Sitaram and Varma [6] got that the intrinsic abundance for the species is log-normal provided that the ratio of all the eigenvalues of 'α' are rational.

3.5 Statistical mechanics of the Gompertz model: Singh and Pande's approach

The result obtained by Kerner [2], Gunasekran and Pande [4], and then Sitaram and Varma [6] are quite fascinating and elegant. But all of them put the same restriction of antisymetric interactions among the individuals, which was needed to ensure the existence of conserved quantities for the ecosystem. The presumption of antisymetric interaction condition put a very strong restriction on a particular community. And for an odd species community the antisymetric matrix is necessarily singular and one of its eigenvalue is zero, leading to unwanted complications.

In 1991, Singh and Pande [7] tried to develop the n-species Gompertz Model that omitted the anti-symmetry condition of the community matrix, which was followed by Kerner [2], Gunasekran and Pande [4], and Sitaram and Varma [6]. They developed micro canonical and canonical ensembles for the ecosystem which are applicable to all kinds of the ecosystem irrespective of weather they contain an even or an odd number of species and the probability distribution of any kind of communities are again Log-normal that was also obtained by Gunasekran and Pande [4], and then Sitaram and Varma [6].

Once again Gompertz model was taken for the ecosystem by Singh and Pande [7] in order to calculate the constant of motion, entropy and other conserved quantities. From the Gompertz model we have,

$$\frac{dN_i}{dt} = \left(\varepsilon_i + \sum_j \alpha_{ij} \log N_j \right) N_i \qquad (1)$$

and they made the similar transformation as done by Gunasekaran and Pande [4],

$$x_i = \log N_i$$

$$\Rightarrow \frac{dx_i}{dt} = \varepsilon_i + \sum_j \alpha_{ij} x_j \qquad (2)$$

At equilibrium state,

$$\frac{dx_i}{dt} = \varepsilon_i + \sum_j \alpha_{ij} x_j^* = 0 \qquad (3)$$

$$\Rightarrow \frac{dx_i}{dt} = \sum_j \alpha_{ij} \left(x_j - x_j^*\right) \qquad (4)$$

This set of equation is consistent if the matrix 'α' is nonsingular. Hence Singh and Pande [7] assumed that the equation is solvable for 'x_j^*' and has finite positive solution. They again introduced,

$$Y_i = x_i - x_j^*$$

Therefore equation (4) took the form of

$$\frac{dY_i}{dt} = \sum_j \alpha_{ij} Y_j \qquad (5)$$

This equation has been be solved by them by assuming a square matrix 'Ω' of order 'n' such that

$$\Omega^{-1} \alpha\ \Omega = D, \qquad (6)$$

where 'D' is diagonal matrix with elements λ_I = 1,2,3…,n. 'α' is the square matrix with element 'α_{in}' of order 'n' and 'λ_I' being eigenvalues of 'α'.

Therefore equation (5) can be written in matrix form as,

$$\dot{Y} = \alpha\ Y, \qquad (7)$$

where 'Y' is a column matrix with the elements 'Y_i', i = 1, 2…n. Now equation (7) takes the form,

$$\Omega^{-1}\dot{Y} = \Omega^{-1}\alpha\ \Omega\Omega^{-1}Y \qquad (8)$$

They took $\dot{Z} = \Omega^{-1}\dot{Y}$ and $Z = \Omega^{-1}Y$. So the equation (8) can be written as

$$\dot{Z} = DZ \qquad (9)$$

The solution of this equation has been found out by Bhatt and Pande [14] and it was,

$$Z_i = A_i \exp(\lambda_i\, t), \qquad (10)$$

where 'A_i' is time independent. While we solving equation (7), some times we get 'A_i' as a polynomial in time 't'. That is not compatible with the equilibrium approach. So Singh and Pande [7] didn't consider that solution, as their work was based on equilibrium approach. Further they used this form of solution and calculated the micro canonical ensemble volume 'τ' and then found the expression for the entropy of the system as

S=logτ.

The micro canonical ensemble volume is nothing other than length of the system trajectory in the phase space. In order to calculate this length Singh and Pande considered the n-dimensional space of the variable 'Z_i'. So the state of the system at any time't' is represented by the point (Z_1, Z_2,..,Z_n). Hence in time, dt, the point moves a distance, ds, is given by,

$$ds = \left[\sum_{i=1}^{n} dZ_i dZ_i^*\right]^{1/2} \qquad (11)$$

$$\Rightarrow \frac{ds}{dt}dt = \left[\sum_{i=1}^{n} \frac{dZ_i}{dt}\frac{dZ_i^*}{dt}\right]^{1/2} dt \qquad (12)$$

Hence the length of the system trajectory or the total volume covered in the entire time is given by,

$$\tau = \int_0^\infty ds = \int_0^\infty \left(\frac{ds}{dt} \right) dt \qquad (13)$$

$$\tau = \int_0^\infty \left[\sum_{i=1}^{n} \frac{dZ_i}{dt} \frac{dZ_i^*}{dt} \right]^{1/2} dt \qquad (14)$$

By the knowledge of the equation (9) and its solution Singh and Pande [7] tried to find out the value of this ensemble volume. The value of this integral depends upon the eigenvalues 'λ_I'. So the integral cannot be evaluated exactly in the most general case. They there fore considered the three possibilities and calculated the corresponding values of 'τ' (volume) and 'S' (entropy).

Case 1: when all 'λ_I' are pure imaginary then,

$$\tau = \int_o^\infty \left[2 \sum_{i=1}^{k} v_i^2 M_i \right]^{1/2} dt$$

$$\therefore \quad S = \log \tau = \left[2 \sum_{i=1}^{k} v_i^2 M_i \right]^{1/2} + K_1 \qquad (15)$$

Case 2: when all 'λ_i' are pure imaginary except one, which is real and negative i.e. for i = n +1, λ_i is real and negative. This is the case of (n+1) dimensional case. Then,

$$S = \log \tau = \left[2 \sum_{i=1}^{k} v_i^2 M_i \right]^{1/2} - \log 2r + K_2 \qquad (16)$$

where '–r' = r^th real negative eigenvalues.

Case 3: Some of the eigenvalues are pure imaginary, some are complex with real negative parts and some are pure real and negative. In this case it is not possible to calculate the integral exactly. Therefore Singh and Pande [7] made an assumption that the term corresponding to eigenvalues with nonzero real parts is small hence neglected, because these term carry negative real exponents so that they would normally damped out sufficiently well in short term duration and in the long run they have not significant contribution. Hence they got the same result as in the case of (1) except the constant.

$$S = \log\tau = \left[2\sum_{i=1}^{k} v_i^2 M_i\right]^{1/2} + K_3, \qquad (17)$$

where K_1, K_2, K_3 are constant those are independent of M_i. Now they have chosen to arrange the system in such a way that the pure imaginary eigenvalues appear first in the diagonal matrix 'D' in pairs and another eigenvalues later. So they got,

$\lambda_1 = iv_1$, $\lambda_2 = -iv_1$, $\lambda_3 = iv_2$, $\lambda_4 = -iv_2$, and so on.

And similarly they got,

$A_1A_1^* = M_1 = A_2A_2^*$ and $A_3A_3^* = M_2 = A_4A_4^*$ and so on.

Singh and Pande [7] didn't consider the case of eigenvalues having positive real parts, which ultimately leads to population explosion that disturbs the equilibrium configuration of the ecosystem.

The constant of integral appeared in the equation (15), (16) and (17) makes the entropy infinite as time, t, tends to infinity. Therefore they remarked that however it is always possible to separate out the relevant finite parts, which exhibit the appropriate dependence on the variables of the system, from the remainder, which is basically a constant whether finite or infinite. The remainder will not contribute in manipulations, which are to follow, as the latter involves only the derivatives of entropy, not the entropy itself. They further remarked that all classical physical theories never be use absolute magnitude of the entropy. So renormalizing the entropy function and throwing out the constant parts in the same sprit irrespective of weather they are finite or infinite.

In real situation physical system is not in complete isolation but rather in contact with other system. In fact they are often parts of the larger system which they maintain equilibrium. The description of such system is then more appropriately given within the framework of canonical ensemble rather than the micro canonical ensemble. And the canonical ensemble volume,τ, is then given by the integral over the entire phase space of density function, ρ;

$$\tau = \int \ldots\ldots\ldots\ldots \int dy_1 dy_2 dy_3 \ldots\ldots\ldots dy_n \, \rho \, (y_1 y_2 \ldots\ldots y_n) \qquad (18)$$

Now corresponding to each conserved quantities of interest for the system Singh and Pande [7] established a general temperature with property that in equilibrium all these temperatures must be correspondingly is equal to those of the surroundings separately. This is the idea behind the canonical ensemble. In order to define the canonical ensemble density and to find the probability distribution they once again took the equation (9) in the form of

$$Z_i = \lambda_i Z_i \qquad (20)$$

$$\Rightarrow \frac{1}{Z_i}\frac{dZ_i}{dt} = \lambda_i \qquad (21)$$

$$\Rightarrow \frac{\lambda_j}{Z_i}\frac{dZ_i}{dt} = \lambda_i \lambda_j$$

$$and, \frac{\lambda_i}{Z_i}\frac{dZ_i}{dt} = \lambda_j \lambda_i$$

$$hence, \lambda_i \frac{1}{Z_j}\frac{dZ_j}{dt} - \lambda_j \frac{1}{Z_j}\frac{dZ_i}{dt} = 0$$

$$\Rightarrow \frac{d}{dt}\left(\lambda_i \log Z_j - \lambda_j \log Z_i\right) = 0$$

$$\Rightarrow \frac{d}{dt}\left[\log\left\{(Z_j)^{\lambda_i}(Z_i)^{\lambda_j}\right\}\right] = 0 \qquad (22)$$

This equation possesses 'n^2' conserved quantities or the system has 'n^2' constants of motion, which are given by,

$$C_{ij} = \left(Z_j\right)^{\lambda_i}\left(Z_i\right)^{-\lambda_j} \qquad (23)$$

But all these 'n^2' constants are not independent. Out of these 'n^2' constants only (n-1) numbers of constants are independent. While the Kerner's approach to apply statistical mechanics to the ecosystem had not able to get these (n-1) number of constants. He got only one constant of motion and that was 'G'. However Sitaram and Varma [6] were able to establish these (n-1) numbers of constants of motion in their work.

Now Singh and Pande [7] took only 'k' number of constants out of '(n-1)' independent constants of motion and considered that remaining '(n-1-k)' will not play any role. The 'k' independent quantities, which were given by Singh and Pande, are $C_{12}, C_{34}, C_{56}, \ldots.. C_{2k-1,2k}$ or simply $M_1, M_2 \ldots M_k$.

Then they considered a simple conventional situation in which the system is specified by its energy, which is a conserved quantity, and then tried to generalize the canonical ensemble density function. This kind of generalization where the system possesses more than one conserved quantities, was suggested first by Grad in 1952. If the system has 'n' number of conserved quantities 'G_i' there are 'n' number of generalized temperature 'θ_i' in the system and the density function 'ρ' generalizes to

$$\rho = \exp\left(- \sum_{i=1}^{n} \frac{G_i}{\theta_i} \right) \qquad (24)$$

Here in the work of Singh and Pande [7], they have taken only 'k' number of appropriately chosen conserved quantities 'M_i' and correspondingly the same number of generalized temperature 'σ_i'. Hence they defined canonical ensemble density function as,

$$\rho = \exp\left(- \sum_{i=1}^{k} \frac{M_i}{\sigma_i} \right) \quad \ldots \qquad (25)$$

Before doing so, they defined,

$$\frac{1}{\sigma_i} = \frac{\partial S_i}{\partial M_i}, \ i = 1, 2, \ldots, k. \qquad (26)$$

This is analogous to the conventional situation, i.e.

$$\frac{1}{\theta} = \frac{dS}{dE},$$

where 'E' = energy that is conserved, 'θ' is temperature and 'S' is the entropy. In this situation 'θ' was replaced by 'σ_i' and 'E' was replaced by 'M_i'

The canonical ensemble volume was given in equation (18) and hence canonical ensemble of any quantity 'f' is then can be written as,

$$\langle f \rangle = \frac{\int \ldots\ldots\ldots \int dy_1 \ldots\ldots dy_n f \exp\left[-\sum_{i=1}^{k} \frac{M_i}{\sigma_i} \right]}{\int \ldots\ldots\ldots \int dy_1 \ldots\ldots dy_n \exp\left[-\sum_{i=1}^{k} \frac{M_i}{\sigma_i} \right]} \qquad (27)$$

Hence the probability distribution function P (y_1) for any variable 'y_i' may be written as,

$$p(y_1) = \frac{\int \ldots\ldots \int dy_2\, dy_3 \ldots\ldots dy_n f \exp\left[-\sum_{i=1}^{k} \frac{M_i}{\sigma_i} \right]}{\int \ldots\ldots \int dy_1\, dy_2 \ldots\ldots dy_n \exp\left[-\sum_{i=1}^{k} \frac{M_i}{\sigma_i} \right]} \qquad (28)$$

where 'M_i' is the 'k' conserved quantities and is given by,

$$M_i = A_i A_i^{*} = Z_i Z_i^{*} \qquad (29)$$

By using the transformation

$$Z_i = \sum_j \left(\Omega^{-1} \right)_{ij} Y_j \qquad (30)$$

$$i.e.\ Z_i = \sum_{j=1}^{n} a_{ij} Y_j$$

and $Z_i^* = \sum_{j=1}^{n} a_{ij}^* Y_j$

They calculated the probability distribution, P (y_i), of any of the 'y_i' and got a normal distribution. Hence the distribution of relative abundance of species follows the log-normal distribution, and the probability distribution was given by

$$NP\left(\log N_\xi\right) = \frac{N}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\left(\log N_\xi - \log N_\xi^*\right)^2}{2\sigma^2}\right], \qquad (31)$$

where 'ξ' is some mean index and 'N' is the total number of species in the system. This result this result is similar to what was obtained by Preston [5] in his work. Thus the work of Singh and Pande [7] explored the possibility of constructing a statistical mechanics which is free from the odd-even dichotomy: it is applicable to a system that may have an odd or even number of species. It was also found that there was situation when the distribution becomes ill defined. This happens when the system does not have enough oscillatory solutions.

Chapter 4

Chapter 4

Experimental evidence for log-normal distribution for interacting multi-species ecosystems

4.1 Introduction

Some of the species in nature are extraordinarily abundant at a particular time and place, whereas some of the others seem rare in comparison. This observation has been closely verified by Saunders [5] for the samples collected for the birds of Quaker Run valley, western New York, by Williams [10] for moths collected in light traps at Rothamsted, by Dirks [9] for moths in light traps at Orono, Maine, by King [5] for moths in a light traps at Saskatoon, Preston, and by Norris [5] for birds of the Frith. Under these circumstances (commonness and rareness) it would be logical merely as a matter of convenience to try plotting abundance of a logarithmic base. But there is a more cogent basis for so doing, as dealt with below.

4.2 Preston's octave approach

For the sake of simplicity, Preston [5] in his work argued that the commonness is a relative matter and one would say that a certain species was so many times as common as another. This led Preston to group the specimen into what he calls "OCTAVES". The intervals 1-2, 2-4, 4-8....etc. are presented by different octaves and they are represented by A, B, C.... octave respectively. The actual details of Preston's scheme for this grouping is as follows,

Number of the octave	Arithmetic grouping of the number of individuals/ species	Corresponding logarithmic grouping	Number of the species belonging to that octave (frequency)		
			Half of +	All	+ Half of
A	1 – 2	0 – 1	1	-	2
B	2 – 4	1 – 2	2	3	4
C	4 – 8	2 – 3	4	5,6,7	8
D	8 – 16	3 – 4	8	9 – 15	16
E	16 – 32	4 – 5	16	17 – 31	32
F	32 – 64	5 – 6	32	33 – 63	64
G	64 – 128	6 – 7	64	65 – 127	128
H	128 – 256	7 – 8	128	129 – 255	256
I	256 – 512	8 – 9	256	257 – 511	512
J	512 – 1024	9 – 10	512	513 – 1023	1024
K	1024 – 2048	10 – 11	1024	1026 – 2047	2048
L	2048 – 4096	11 – 12	2048	2049 – 4095	4096
M	4096 – 8192	12 – 13	4096	4097 – 8191	8192
N	8192 - 16384	13 – 14	8192	8193 - 16383	16384
....	....	....	....	....	....

A species falling on as octave may be thought of as having roughly the same degree of commonness in comparison to those falling in another octave. A species that falls on a group boundary is treated as contributing half a species to the left octave and half to the right octave. If some species are represented by 5, 6, 7 specimen it clearly falls in an octave ‘C’ and if it is represented by 33, 34, 35...63 then it falls in octave ‘F’. If a species is represented by ‘32’ specimen then octave ‘E’ is credited with half a species and octave ‘F’ is credited with the other half.

The mid point of each octave is double that of the preceding one. Thus Preston in his work took the abscissa in the species abundance curve becomes simply a scale of octaves which is equivalent to taking logarithm of the number of individual per species, to the base 2 i.e. he Plotted S_2 against $\log_2 N_i$.

Now in next section we discuss the experimental evidence to the theoretical calculation. The sample collected by Saunders [8], Dirks [9] and Williams [10] has been taken for this purpose.

Before doing so it is necessary to discuss that how a sample will represent a real picture of the ecosystem. If we consider the sample theory we know that the sample will be a sufficiently accurate replica of the universe provided (a) it is a perfectly "random sample and (b) no species is represented in the sample by less than 20-30 individual. But in most of the ecological work condition 'b' will never be obtained and condition 'a' will be poorly fulfilled.

In case of Ecological research the universe changes rapidly. We are dealing with a fleeting and fluctuating assemblage, a 'universe' continually expanding, contracting and changing in composition. Thus this important to recognize at the outset for the purpose of our present investigation, the 'universe' from which the sample is drawn is that universe declared to us by the sample itself and not our preconceived notion of what the universe ought to be.

4.3 Evidence for lognormal distribution from field studies

(a) The birds of Quaker Run Valley, Western New York collected by Saunders [8]

Saunders collected the birds on a tract of 16967 acres of land at Quaker Run Valley situated in western New York. His main reliance was upon a count of singing males aided by a few auxiliary methods. He collected about 80 species with varying individuals from 1 to 1670. The Preston form of sample data is shown in table 1 and corresponding plot is in fig 1.

(b) Moths caught in light trap at Orono, Maine, Uk by Dirks [9]

Dirks collected moths in light trap over period of four years (1931-34). The specimen is classified in table 2 and corresponding graph can be seen in fig 2.

(c) Female moths caught in light trap at Orono, Maine, UK by Dirks [9]

In the same period (1931-34), Dirks categories the data for female moths. The condensed form of data is grouped in the form of octave in table 3 and corresponding plot is reproduced in fig 3.

(d) Moths caught in light trap at Rothamasted, England by Williams [10]

The octave form of this sampled data collected by Williams is given in table 4 and corresponding graph is shown in fig 4.

Examination of above sampled data

Now plotting the above grouped data with abscissa as the scale of octave and the number of species belonging to a particular octave as the ordinate. We observed that each curve (1-4) exhibits a maximum in some octave to the right of the first one and the observed octave frequencies first increase and then decrease. The octave which corresponds to the maximum in the graph containing maximum number of species is known as the modal octave. The main feature, which emerges from this examination, is that in all cases the data is well fitted by a symmetrical Gaussian curve truncated on the left. The general equation of Gaussian curve is

$$n(R) = n_0 \exp\left[-(aR)^2\right], \qquad (1)$$

where 'n_0' is the number of species in the modal octave. n (R) is the number of species in an octave distance 'R' octave from the modal octave and,

$$a^2 = \frac{1}{2\sigma^2}. \qquad (2)$$

where 'σ' is the measure of the mean square deviation of the population from its mean value. We thus have

$$R = \left(\log_2 i - \log_2 i^*\right) \qquad (3)$$

where i is the number of individuals and i^* is the particular value of i corresponding to the peak of the curve which is plotted $\log_2 i$ at x-axis against the species per octave. Hence i^* is the mid value for the modal octave. Note that

$$\overline{R^2} = \overline{\left(\log_2 i - \log_2 i^*\right)^2}, \qquad (4)$$

and given by

$$\overline{R^2} = \frac{\int_{-\infty}^{+\infty} n(R)\,R^2\,dR}{\int_{-\infty}^{+\infty} n(R)\,dR} \qquad (5)$$

$$= \frac{n_0\sqrt{\pi}\big/2a^3}{n_0\sqrt{\pi}\big/a}$$

$$\Rightarrow \overline{R^2} = \frac{1}{2a^2} = \sigma^2 \qquad (6)$$

Now the total number of species, N, for the system is given by

$$\int_{-\infty}^{+\infty} n(R)\,dR = \frac{n_0\sqrt{\pi}}{a} = N \qquad (7)$$

This, N, is the total number of species that are theoretically observed.

For data (a) the experimentally calculated value of 'a' is 0.194. This particular data is well fitted in the curve

$$n(R) = 10\ \exp\left[-(0.194R)^2\right] \qquad (8)$$

Substituting this in equation (7), we get N = 91.

From the empirical data of Saunders [8], we see that there are 141 species of birds that are known to occur at the site of sampling and its vicinity. Out of these, 27 are migratory species; 13 breed out side the site and 11 species are rare. These amounts to

51. Of the remaining 90 species, 11 are irregular in their occurrence and are not found every year; the rest (79) are regular breeders. Thus we see that there is good agreement between the experimentally observed and theoretically calculated values for the total number of species in the environment under consideration.

For data (b), collected by Dirks [9], the experimentally calculated value for 'a' is '0.207' and it is found that the best fit to data is as given by

$$n(R) = 48 \exp\left[-(0.207R)^2\right], \qquad (9)$$

and the theoretically calculated value of total number of species (N) is 410. The number of species, which observed in the sample, was 349 that are about 85% of the estimated one. This is again a very good approximation, which was obtained by Gaussian curve. And $R = \left(\log_2 i - \log_2 i^*\right)$, hence species abundance follows log-normal distribution.

For data (c), collected by Dirks [9], the best fit curve is represented by the equation

$$n(R) = 42 \exp\left[-(0.205R)^2\right] \qquad (10)$$

where 'a = 0.205' and the estimated value for the total number of the species, N, is 363.

While the sample collected by Dirks [9] for female moth has 226 species that is about 62% of the estimated one.

For the data (d), collected by Williams [10] experimentally calculated value of 'a' is '0.227' and the data is well fitted by the curve

$$n(R) = 35 \exp\left[-(0.227R)^2\right] \qquad (11)$$

For this data, the estimated value for 'N' is 273 while the observed number of species in the sample was 240. Again it is a good approximation of the log-normal distribution.

TABLE 1

Octave	No. of individuals Per species (i)	$Log_2 i$	No. of species per octave
-	<1	-	>1
A	1 – 2	0 - 1	1.5
B	2 - 4	1 - 2	6.5
C	4 – 8	2 – 3	8
D	8 – 16	3 – 4	9
E	16 – 32	4 – 5	9
F	32 – 64	5 – 6	12
G	64 – 128	6 – 7	6
H	128 – 256	7 – 8	9
I	256 – 512	8 – 9	11
J	512 – 1024	9 – 10	4
J	1024 – 2048	10 – 11	3

(a) The birds of Quaker Run Valley, Western New York collected by Saunders

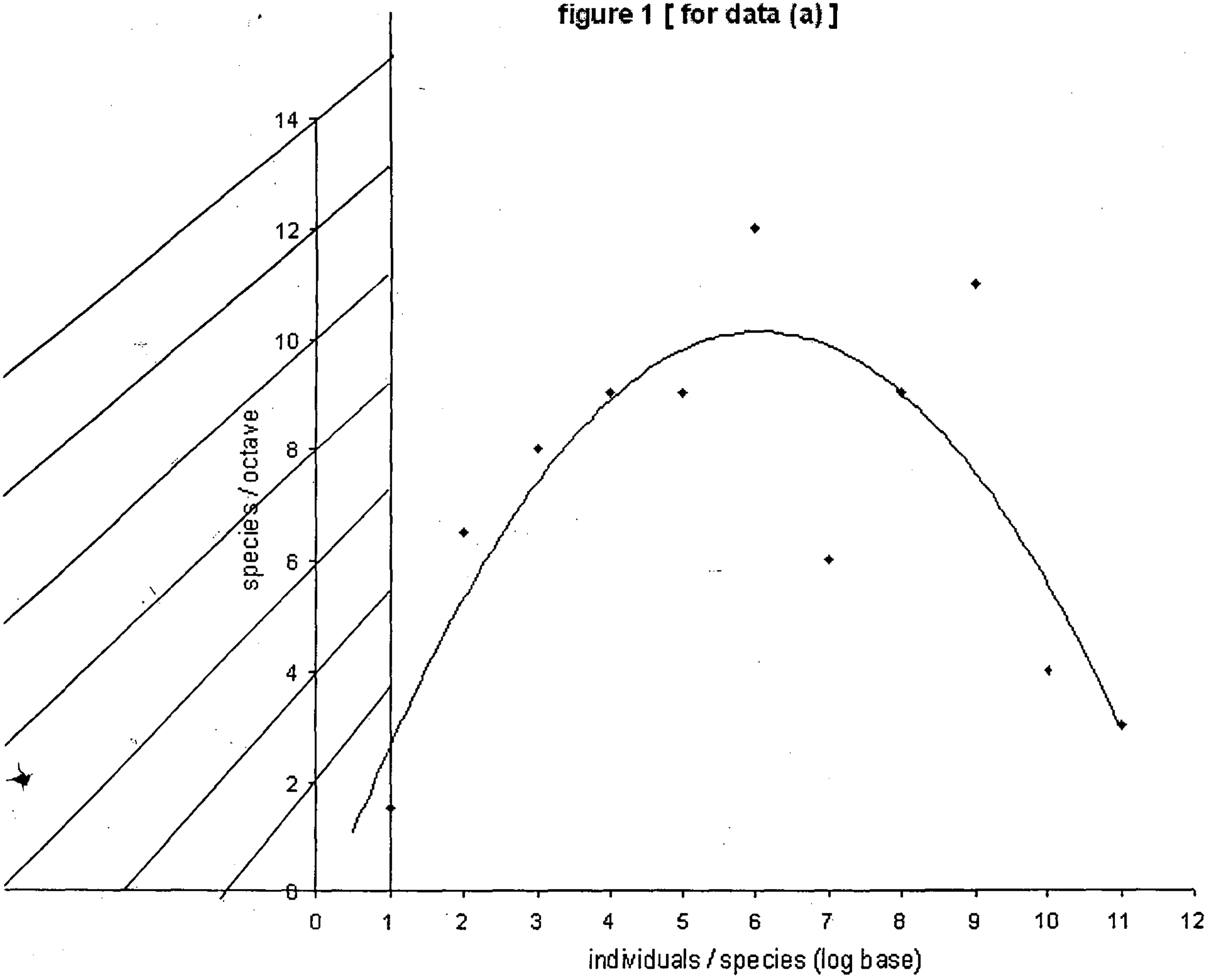

Fig (1): The birds of Quaker Run Valley, Western New York collected by Saunders

TABLE 2

Octave	No. of individuals Per species (i)	Log_2i	No. of species per octave
-	<1	-	>19
A	1 - 2	0 - 1	37
B	2 - 4	1 - 2	42
C	4 – 8	2 – 3	49.5
D	8 – 16	3 – 4	45.5
E	16 – 32	4 – 5	42
F	32 – 64	5 – 6	28.5
G	64 – 128	6 – 7	26.5
H	128 – 256	7 – 8	30
I	256 – 512	8 – 9	14
J	512 – 1024	9 – 10	9
K	1024 – 2048	10 – 11	2
L	2048 – 4096	11 – 12	0
M	4096 – 8192	12 – 13	2
N	8192 - 16384	13 – 14	2

(b) Moths caught in light trap at Orono, Maine, Uk by Dirks

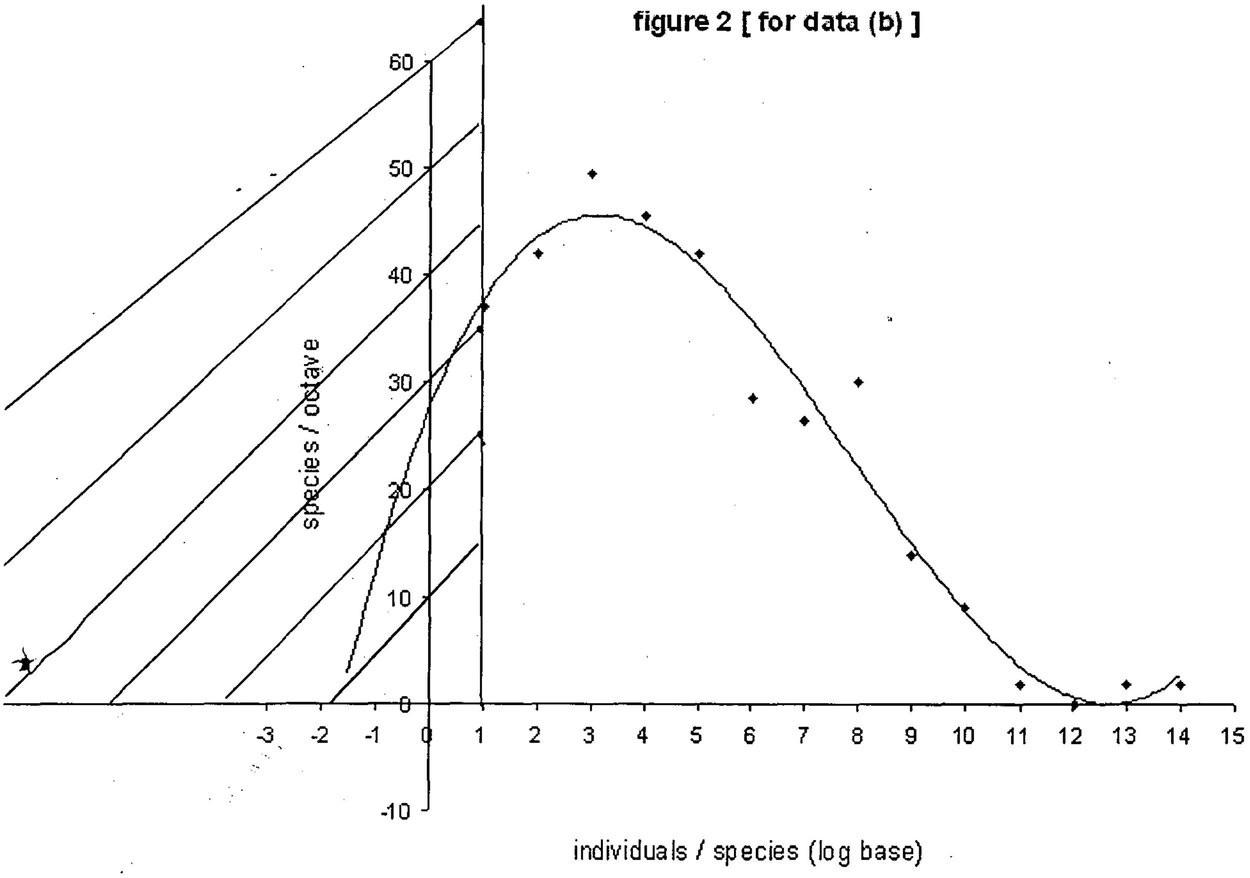

Fig (2): Moths caught in light trap at Orono, Maine, Uk by Dirks

TABLE 3

Octave	No. of individuals Per species (i)	Log_2i	No. of species per octave
-	<1	-	>27.5
A	1 - 2	0 - 1	42
B	2 - 4	1 - 2	39.5
C	4 – 8	2 – 3	27.5
D	8 – 16	3 – 4	28
E	16 – 32	4 – 5	23
F	32 – 64	5 – 6	19.5
G	64 – 128	6 – 7	9
H	128 – 256	7 – 8	4
I	256 – 512	8 – 9	4
J	512 – 1024	9 – 10	0
K	1024 – 2048	10 – 11	0
L	2048 – 4096	11 – 12	1
M	4096 – 8192	12 – 13	1

(c) Female moths caught in light trap at Orono, Maine, UK by Dirks

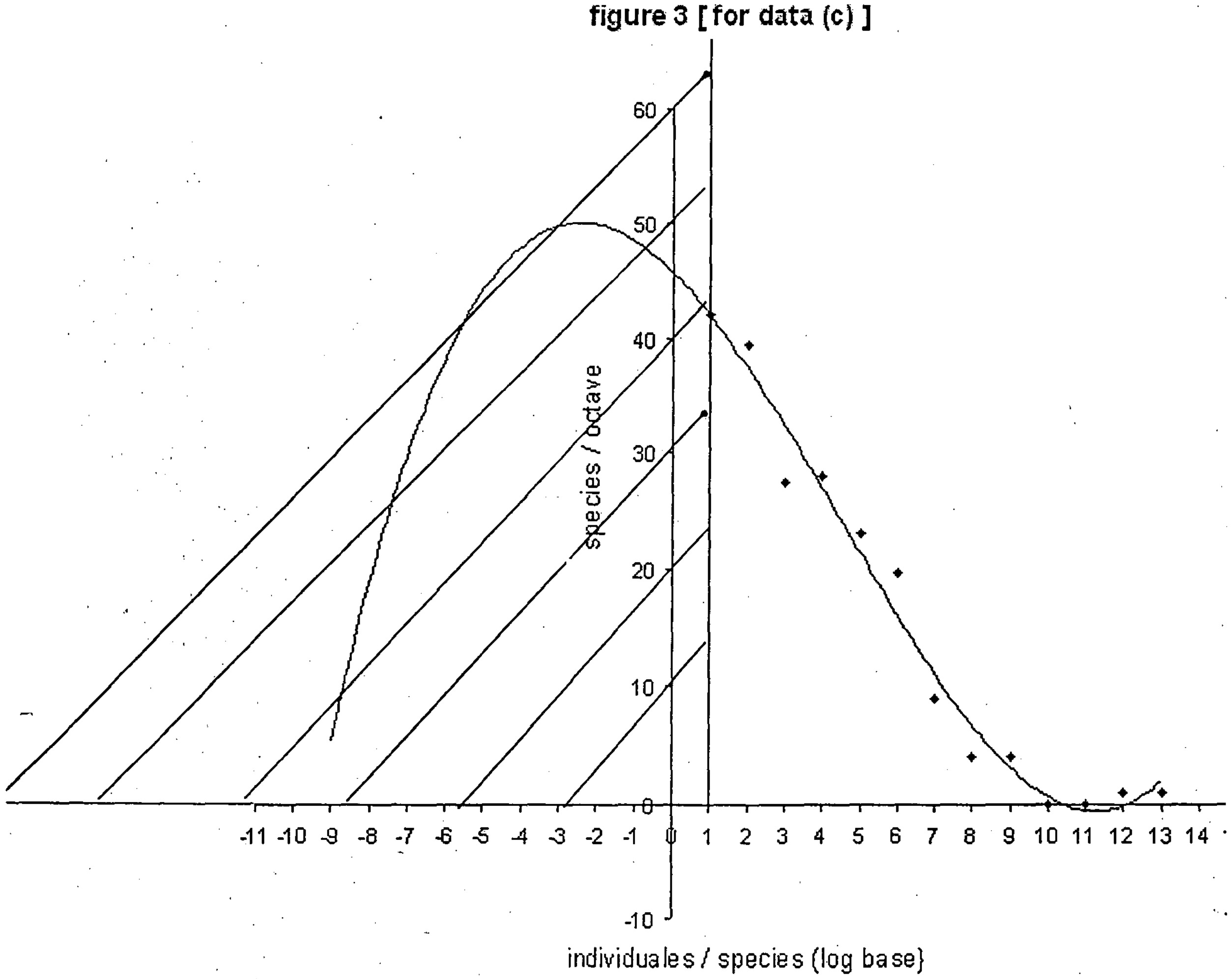

Fig (3): Female moths caught in light trap at Orono, Maine, UK by Dirks

TABLE 4

Octave	No. of individuals Per species (i)	$\log_2 i$	No. of species per octave
-	<1	-	>17.5
A	1 - 2	0 - 1	23
B	2 - 4	1 - 2	27.5
C	4 – 8	2 – 3	36
D	8 – 16	3 – 4	27.5
E	16 – 32	4 – 5	33
F	32 – 64	5 – 6	31
G	64 – 128	6 – 7	13.5
H	128 – 256	7 – 8	19
I	256 – 512	8 – 9	5
J	512 – 1024	9 – 10	6
K	1024 – 2048	10 – 11	0
L	2048 – 4096	11 – 12	1

(d) Moths caught in light trap at Rothamasted, England by Williams

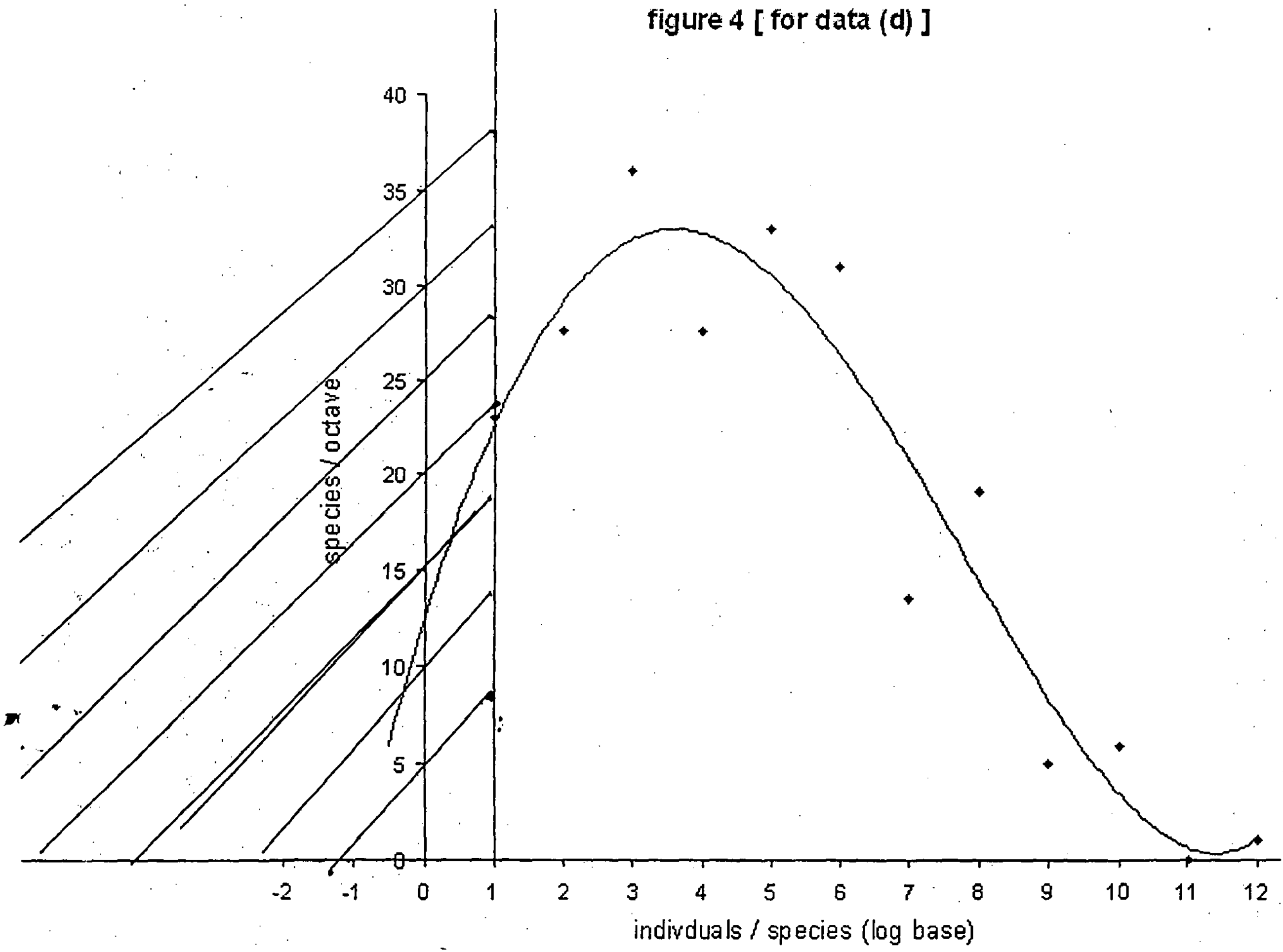

Fig (4): Moths caught in light trap at Rothamasted, England by Williams

4.4 Summary and concluding remarks

We have been concerned in this dissertation with a statistical mechanical approach to the modeling of ecosystems. This approach suggests itself naturally because of the similarities between an ecosystem and a physical system, both of which consist of enormous numbers of entities. While the set of equations that determine the dynamics can be written down in both the cases, it is impractical and not even very illuminating to solve them. In such a situation it is useful to determine the gross features of the manner in which the system might evolve. For instance, one could seek to determine if there are any conserved quantities that characterize the system (e.g., the total energy of the system/biomass), and the manner in which entities of the system are distributed over specific velocity ranges/species). That such features can be unraveled was first shown by Kerner [2], who also obtained an analytic expression for the distribution of individuals over different species; this may be called the Kerner distribution. This is the work that has been covered in chapter 2. Essential features of the Kerner method are: (a) the dynamics of the ecosystem is assumed to be given by Volterra's equations for k-species (k very large), and (b) the community matrix is assumed to be anti-symmetric.

Chapter 3 deals with the work of several authors. The work of Gunasekran and Pande was motivated by the fact that the field studies carried out by Saunders [8], Preston and Norris [5], Dirks [9], and Williams [10], etc. yield not the distribution obtained by Kerner, but a log-normal distribution. In an elegant manner, they were able to show that if one assumes the dynamics of the ecosystem to be given by the Gompertz model (rather than Volterra's equations), then one is led to the log-normal model. Like Kerner, these authors also used the property of anti-symmetry of the community matrix, and found that the system admits of one conserved quantity. Sitaram and Varma [6] also used the Kerner method for the Gompertz model, but showed that the system admits not one but k/2 constants of motion ("temperatures"). This set the stage for the work of Singh and Pande [7], who gave the most general treatment of the statistical mechanics of the Gompertz model by applying the Kerner method to it without the assumption that the community matrix is anti-symmetric. The final result for the distribution of individuals over different species in all these studies was a log-normal distribution.

We now comment on the fact that in each of the four plots of log-normal distribution that have been reproduced in figures (1-4), the full curve that characterizes a Gaussian distribution is not visible: the curve is decapitated on the left side. This has to do with the relation of a (random) sample to the "universe" it is supposed to represent. Because of the commonness and rarity of different species, a part of the universe will remain "hidden" (under a "veil", as it were) in any sample. The veil line will occur at different locations in different samples: it may coincide with the crest of the curve or it may be at any point on the ascending or descending slopes. If one regards the universe as fixed, one can bring the veil line to any point one wishes by continuing to collect. Ultimately one can collect the whole universe, and then the sample will coincide with the universe which, it is believed, will be a log-normal one.

Finally, a reasonable question to ask is: Among all the distributions that are known, why should the log-normal distribution be the one that is favoured by nature? It seems to us that the answer to this question lies in the observation that evolutionary changes in an ecosystem take place at an infinitesimally small rate, which is precisely the condition for a physical system to pass from one equilibrium state to another. The role of equilibrium states can not be over-emphasized in any statistical mechanical approach. This explains, in a sense, as to why the statistical mechanical approach for a physical system works for an ecosystem. And now about the question: why log-normal rather than normal distribution? One might say that nature prefers the log-normal one because then the slow rate of evolution is in-built into the distribution itself (log is a slowly varying function of its argument).

Bibliography

BIBLIOGRAPHY

[1] Odum, E. P. *Fundamentals of Ecology*, Natraj Publishers, Dehradun, India.

[2] Kerner, E. H. (1957). *Bulletin of Mathematical Biophysics,* **19:**121-146.

[3] Pielou, E. C. (1977) *Mathematical ecology, New York, John-Wiley and Sons.*

[4] Gunasekaran, N. and Pande, L. K. (1982). *Journal of Theoretical Biology*, **98**: 301- 305

[5] Preston, F.W. (1948) *Ecology* **29**: 254-283.

[6] Sitaram, B.R. and Varma, V.S. (1984) *Journal of Theoretical Biology,* **110**: 253-256.

[7] Singh. Balwant and Pande, L. K. (1991) *Proc. Nat. Acad. Sci. India,* **61 (A), 1V**: 543- 555.

[8] Saunders, A. A. (1936) *New York State Museum Handbook* **16**, Albany, N. Y.

[9] Dirks, C. O. (1937) *The Maine Agriculture Experiment Station, Bulletin* **389**, Orono, Me.

[10] Corbet, A. S., Fisher, R. A. and Williams, C.B. (1943). *Journal of Animal Ecology,* **12**: 42-58.

[11] Takenake, S. 1941. *Jap. Jour. Med. Sci., Part iii, Biophysics,* **7**: 129-140.

[12] Volterra, V. (1978) in *The Golden Age of Theoretical Ecology*. Ed.Scudo, F.M. and Ziegler, J. R, Springer- Verlag, Heidelberg.

[13] Gibbs, J.W. 1902. *Elementary Principles in Statistical Mechanics*, New Haver; Yale University Press

[14] Spiegel, M. R. *Theory and Problem of Statistics in SI Unit*, McGraw-Hill book Company, Singapore.

[15] Gompertz, B. (1825) *Phil. Trans. Roy. Soc.* **115**: 513

[16] Bhat. N. and Pande, L.K. (1980). *Journal of Theoretical Biology,* **83:** 321-344.

[17] Schrodinger, E. 1952. "*Statistical Thermodynamics.*" Cambridge: University Press.

[18] Rosen, R. (1970), *Dynamical System Theory in Biology*, Vol. I., Wiley Inter Sciences Series on Biomedical Engineering, New York.

[19] Whittaker, E. T. and Watson, G. N. (1969), *A Course of Modern Analysis*, Cambridge University Press. London

[20] Preston, F. W. (1962). *Ecology* **43:** 185-215; **43:** 410-432

[21] Ayres, F. (1962), *Theory and Problems of Matrices*, Tata Mcgraw-Hill Pub.Co. Ltd. New Delhi.

[22] Lotka, A.J. (1925), *Elements of Physical Biology.* Baltimore; Williams and Wilkins Co.

[23] Maynard Smith, J. (1975). *Models in Ecology.* London: Cambridge University Press

[24] Pande, L. K. (1978) *J. Theor. Biol.* **83:** 321

[25] Gomatam, J. (1974). *Bulletin of Mathematical Biology,* **36:**355-364.

[26] Kerner, E. H. (1974). *Bulletin of Mathematical Biology,* **36:** 477-488.