Entropy **2014**, *xx*, 1-x; doi:10.3390/

OPEN ACCESS
entropy
ISSN 1099-4300
www.mdpi.com/journal/entropy

Article

Speed Gradient and MaxEnt principles for Shannon and Tsallis entropies

Alexander L. Fradkov [1], Dmitry S. Shalymov [2]*

[1] Institute for Problems of Mechanical Engineering, Russian Academy of Sciences, 61, Bolshoy ave. V.O., 199178, Saint-Petersburg, Russia
[2] Mathematics and Mechanics Faculty, Saint-Petersburg State University, Universitetsky prospekt, 28, 198504, Saint-Petersburg, Russia

* Author to whom correspondence should be addressed; dmitry.shalymov@gmail.com

Received: xx / Accepted: xx / Published: xx

Abstract: In this paper we consider dynamics of non-stationary processes that follow the MaxEnt principle. We derive a set of equations describing dynamics of a system for Shannon and Tsallis entropies. Systems with discrete probability distribution are considered under mass conservation and energy conservation constraints. The existence and uniqueness of solution are established and asymptotic stability of the equilibrium is proved. Equations are derived based on the speed-gradient principle originated in control theory.

Keywords: Shannon entropy, Tsallis entropy, maximum entropy (MaxEnt) principle, non-linear kinetics, speed-gradient principle

PACS classifications: 65.40.Gr, 05.45.-a, 02.30.Yy, 02.50.-r

1. Introduction

The notion of entropy is widely used in modern statistical physics, thermodynamics, information theory, engineering etc.

In 1948, Claude Shannon [1] introduced the entropy of a probability distribution $P(X)$ as

$$S(X) = -\sum_i P(x_i) \log P(x_i), \tag{1}$$

where X is a discrete random variable with possible values $\{x_1, ..., x_n\}$.

In 1988, Constantino Tsallis [2] introduced a generalized Shannon entropy.

$$S(X, q) = \frac{1}{q-1}\left(1 - \sum_i P(x_i)^q\right), \tag{2}$$

where q is any real number. It was shown that Tsallis entropy tends to Shannon entropy when $q \to 1$. Tsallis entropy have found many applications in various scientific fields such as chemistry, biology, medicine, economics, geophysics, etc.

A phenomenon when system tends to its state with maximum entropy is known as the maximum entropy (MaxEnt) principle. Since seminal works of E.T. Jaynes (1957) [3,4] and until recent years [5–7] the MaxEnt principle attracts a strong interest of researchers.

Despite a large number of publications studying the maximum entropy states, the dynamics of evolution and transient behavior of the system are still not well investigated. The MaxEnt principle defines the asymptotic behavior of the system, but does not say anything about how the system moves to its asymptotic behavior.

We use the speed-gradient (SG) principle that has already been successfully applied in [8–10] to obtain equations of the dynamics for discrete systems. Applicability of the SG principle is experimentally tested in [10] for the systems of finite number of particles simulated with the molecular dynamics method. In this paper we apply SG-approach to describe dynamics of the systems that tend to maximize Tsallis entropy.

The evolution law of the system for Shannon entropy is formulated in the following general form:

$$\dot{P}(t) = -\Gamma(I - \Psi) \log P(t), \tag{3}$$

where logarithm of the vector is understood componentwise, I is an identity operator, Ψ is a symmetric matrix, $\Gamma > 0$ is a constant gain. For Tsallis entropy the evolution law is

$$\dot{P}(t) = -\Gamma \frac{q}{q-1}(I - \Psi_{Ts})P(t)^{q-1}, \tag{4}$$

By means of this approach the distribution corresponding to the maximum value of entropy is found.

There is a well known general form of the time-evolution equations for non-equilibrium systems, abbreviated as GENERIC (general equation for the non-equilibrium reversible-irreversible coupling) [11,12]. SG-principle has some similarities with GENERIC equations for the case when the goal function is set as entropy and constraints are specified by energy. Nevertheless SG-principle can be considered to be more general because almost arbitrary functional may be taken as a goal function, not only entropy functional. Connection between GENERIC and SG-principle is shown in section 2.3.

The paper is organized as follows. The next section formulates the SG principle with some illustrating examples. The 3rd section introduces Jaynes' formalism. Dynamics equation for the system that tends to maximize Shannon entropy is described in 4th section. The 5th section derives dynamics equation for Tsallis entropy. Equilibrium stability and asymptotic convergence is proved. The 6th section contains results from 5th section extended to the internal energy conservation constraint.

2. Speed-Gradient Principle

Let us consider a category of open physical systems which dynamics are described by the system of nonlinear differential equations

$$\dot{x} = f(x, u, t), \tag{5}$$

where $x \in \mathbb{C}^n$ is the system state vector, u is the vector of input (free) variables $t \geq 0$. The problem is to derive the law of variation (evolution) of $u(t)$ that satisfies some criterion of "naturalness" of its behavior to give the model features characterizing a real physical system.

A typical approach to derive such a criterion from variational principles usually starts with specifying some integral functional (for example, the action functional of the least action principle [13]). Functional minimization defines probable trajectories of the system $\{x(t), u(t)\}$ as points in the corresponding functional space. An alternative approach is based on local functions depending on a current system states. According to M. Planck [14], local principles have some advantage over integral ones because the current state and motion of the system do not depend on its later states and motions. Following [8,15], let us formulate the speed-gradient local variational principle allowing to synthesize the laws of system dynamics.

The speed-gradient principle: *of all possible motions, the system implements the ones for which input variables vary in the direction of the speed-gradient of some "goal" functional Q_t.* If the constraints are imposed on a motion of the system then the direction is a speed-gradient vector projection on the admissible directions (the ones that satisfy the constraints) set.

The SG equation of motion is formed as the feedback law in the finite form

$$u = -\Gamma \nabla_u \dot{Q}_t \tag{6}$$

or in the differential form

$$\frac{du}{dt} = -\Gamma \nabla_u \dot{Q}_t, \tag{7}$$

where $\dot{Q}_t$ is a rate of change of the goal functional along the trajectory of the system (5), Γ is a positive constant gain that may be a positive definite matrix.

Let us describe application of the SG principle in the simplest (yet the most important) case where a category of models of the dynamics (5) is specified as the relation:

$$\dot{x} = u. \tag{8}$$

The relation (8) just means that we are deriving the law of variation of velocity of the system state. In accordance with the SG principle, the goal functional (function) $Q(x)$ needs to be specified first. $Q(x)$ should be based on physics of a real system and reflect its tendency to decrease the current $Q(x(t))$ value. After that, the law of dynamics can be expressed as (6).

The SG principle is also applicable to develop models of the dynamics of distributed systems that are described on infinite-dimensional state spaces. There in particular can be a vector x in a Hilbert space $\mathcal{X}$ and a nonlinear operator $f(x, u, t)$ defined on a dense set $D_F \subset \mathcal{X}$; in this case, the solutions of equation (5) are generalized.

To illustrate how the SG principle works in physical sense we consider several examples.

2.1. Example 1: Motion of a particle in the potential field

In this case the vector $x = (x_1, x_2, x_3)^T$ consists of coordinates x_1, x_2, x_3 of a particle. Choose smooth $Q(x)$ as the potential energy of a particle and derive the speed-gradient law in the differential form. To this end, calculate the speed

$$\dot{Q} = \partial Q/\partial t = [\nabla_x Q(x)]^T u \tag{9}$$

and the speed-gradient

$$\nabla_u \dot{Q} = \nabla_x Q(x) \tag{10}$$

Then, choosing the diagonal positive definite gain matrix $\Gamma = m^{-1} I_3$, where $m > 0$ is a parameter, I_3 is the 3×3 identity matrix, we arrive at the Newton's law $\dot{u} = -m^{-1} \nabla_x Q(x)$ or

$$m\ddot{x} = -\nabla_x Q(x).$$

Note that the speed-gradient laws with non diagonal gain matrices Γ can be incorporated if a non-Euclidean metric in the space of inputs is introduced by the matrix Γ^{-1}. Admitting dependence of the metric matrix Γ on x one can obtain evolution laws for complex mechanical systems described by Lagrangian or Hamiltonian formalism.

The SG-principle applies not only to finite dimensional systems, but also to infinite dimensional (distributed) ones. Particularly, x may be a vector of a functional space $\mathcal{X}$ and $f(x, u, t)$ may be a nonlinear differential operator (in such a case the solutions of (5) should be understood as generalized ones). We will omit mathematical details for simplicity.

2.2. Example 2: Wave, diffusion and heat transfer equations

Let $x = x(r),\ r = (r_1, r_2, r_3)^T \in \Omega$ be the temperature field or the concentration of a substance field defined in the domain $\Omega \in R^3$. Choose the goal functional evaluating non-uniformity of the field as follows

$$Q(x,t) = \frac{1}{2}\int_\Omega |\nabla_r x(r,t)|^2 dr,$$

where $\nabla_r x(r,t)$ is the spatial gradient of the field and boundary conditions are assumed zero for simplicity. Calculation of the speed $\dot{Q}(x,t)$ and then speed-gradient of Q_t yields

$$\frac{d}{dt}Q = -\int_\Omega \Delta x(r,t) u(r,t) dr, \quad \nabla_u \dot{Q} = \Delta x(r,t),$$

where $\Delta = \sum_{i=1}^3 \frac{\partial^2}{\partial r_i^2}$ is the Laplace operator. Therefore the speed-gradient law in differential form (7) is

$$\frac{\partial^2}{\partial t^2} x(r,t) = -\gamma \Delta x(r,t),$$

which corresponds to the D'Alembert wave equation. The SG-law in finite form (6) reads

$$\frac{\partial x}{\partial t}(r,t) = -\gamma \Delta x(r,t)$$

and coincides with the diffusion or heat transfer equation.

Note that the differential form of the speed-gradient laws often corresponds to reversible processes while the finite form generates irreversible ones. For modelling of more complex dynamics a combination of finite and differential SG-laws may be useful.

In a similar way dynamical equations for many other mechanical, electrical and thermodynamic systems can be recovered. The SG-principle applies to a broad class of physical systems subjected to potential and/or dissipative forces. This paper is aimed at application of the SG-principle to entropy-driven systems.

2.3. GENERIC and SG-principle

GENERIC time-evolution equation can be written as

$$\frac{dx}{dt} = L(x)\frac{\partial E(x)}{\partial x} + M(x)\frac{\partial S(x)}{\partial x}, \tag{11}$$

where x represents a state of a system, E is a total energy functional and S is entropy functional.

Let us show connection between GENERIC and SG equations. Suppose we have to maximize entropy function $S(x)$ of a system having constraint for a total energy $E(x) = E = const$. Lagrange function can be defined as

$$\Lambda(x, \lambda) = S(x) + \lambda\,(E(x) - E) \tag{12}$$

Let us use SG-principle eq. (6) for Lagrangian (12), i.e. $Q_t = \Lambda$.

$$u = -\Gamma\nabla_u\dot{\Lambda}(x, \lambda) = -\Gamma\nabla_u\left[\dot{S} + \lambda\dot{E}\right] \tag{13}$$

According to (9) and (10) we can rewrite eq. (13) as

$$u = -\Gamma\left[\nabla_x S(x) + \lambda\nabla_x E(x)\right] \tag{14}$$

We can see that dynamics equation obtained from SG-principle (14) coincides with GENERIC equation (11) for $\frac{dx}{dt} = u, \quad L(x) = -\Gamma\lambda$ and $M(x) = \Gamma$.

GENERIC is based on two 'potentials' of total energy and entropy. SG-principle can use any functional that has to be maximized (minimized) as a goal function. It may be not only Lagrangian (12) or entropy functional. SG-principle can consider various functionals and constraints. So it can be treated as a more general approach. Nevertheless GENERIC is also a general equation. It uses parametrized matrices $L(x)$ and $M(x)$ that make it possible to use GENERIC for a wide range of time-evolution systems.

3. Jaynes's Maximum Entropy Principle

The approach proposed by Jaynes [3,4] became the foundation for statistical physics nowadays. Its main ideas are described below.

Let $P(x)$ be a probability distribution for a discrete random variable X. This is an unknown distribution that needs to be defined on the basis of a certain system information. Let us suppose that there is the information about some average values $\overline{H}_m$ which are known a priori:

$$\overline{H}_m = \sum_{i=1}^{n} H_m(x_i)P(x_i), \; m = 1, ..., M. \tag{15}$$

The next equality is also true for probability distribution

$$\sum_{i=1}^{n} P(x_i) = 1. \tag{16}$$

Conditions (15) and (16) in general can be insufficient to derive $P(x)$. In this case, according to Jaynes, applying maximization of information entropy S_I (1) is the most objective method to define the distribution.

Maximum search with additional conditions (15) and (16) is performed by using Lagrange multipliers; it leads to

$$P(X) = \frac{1}{Z(\lambda_1, .., \lambda_m)} exp\left(\sum_{m=1}^{M} \lambda_m H_m\right), \tag{17}$$

$$Z(\lambda_1, .., \lambda_m) = \sum_{i=1}^{n} exp\left(\sum_{m=1}^{M} \lambda_m H_m\right), \tag{18}$$

where λ_m can be derived from conditions (15).

In case of equilibrium these formulas show that the maximum of information entropy coincides with the Boltzmann-Gibbs entropy and can be identified with the thermodynamic entropy.

4. Maximization of Shannon entropy with the Speed-Gradient method

Consider a discrete system which consists of N identical particles distributed over m cells.

In case when the mass conservation constraint holds it is true that $\sum_{i=1}^{m} N_i = N$. It can be normalized as

$$\sum_{i=1}^{m} \frac{N_i}{N} = 1 \tag{19}$$

Particles can migrate from one cell to another. We are interested in both the steady-state and the transient behavior of the system. Due to MaxEnt principle it is true that the limit behavior of the system maximizes its entropy for the steady-state in case when nothing else is known [4].

To get a transient mode behavior first we apply the SG-principle choosing Shannon entropy (1) as the goal function to be maximized.

For simplicity we assume that the motion is continuous in time and the numbers N_i are changing continuously. Then the law of motion can be represented as

$$\dot{N}_i = u_i, i = 1, ..., m, \tag{20}$$

where $u_i = u_i(t)$ are control functions (6) which has to be determined.

This approach has already been introduced in [9]. So we omit the details of derivation of equations and introduce the final form of the system dynamics:

$$N_i = \frac{\gamma}{m} \sum_{i=1}^{m} \ln N_i - \gamma \ln N_i, \;\; i = 1, .., m, \;\; \gamma > 0 \tag{21}$$

In case when internal energy constraint also holds:

$$E = \sum_{i=1}^{m} \frac{N_i}{N} E_i, \tag{22}$$

where E_i is the energy of particle in the ith cell and the total energy does not change, then the evolution law has the form:

$$\dot{\mathcal{N}} = -\Gamma(I - \Psi_S)\log\mathcal{N}, \tag{23}$$

where logarithm of the vector is understood componentwise, $\mathcal{N} = (N_1, ..., N_m)^T$, I is an identity matrix, Ψ_S is a symmetric $m \times m$ matrix defined as follows:

$$\psi_{ij} = \frac{1}{m} + \tilde{E}_i\tilde{E}_j, \quad i, j = 1, .., m,$$

where $\tilde{E}_i = E_i - \frac{1}{m}\sum_{i=1}^{m} E_i$ is a vector of energies.

It is shown that the limit probability distribution is unique and can be obtained from Jaynes's MaxEnt principle [9].

5. Speed-Gradient dynamics of Tsallis entropy maximization process

We extend approach introduced in previous section to the case of Tsallis entropy.

To get evolution law for the mass conservation constraint (19) we apply the SG-principle choosing Tsallis entropy as the goal function to be maximized

$$S(X, q) = \frac{1}{q-1}\left(1 - \sum_{i=1}^{m}\left(\frac{N_i}{N}\right)^q\right), \tag{24}$$

where q is any real number and $X = (N_1, ..., N_m)^T$ is the state vector of the system.

We will also use the law of motion in the form (20).

Evaluation scheme is as follows. First the speed of entropy change is evaluated:

$$\dot{S}(q) = \frac{q}{(1-q)N^q}\sum_{i=1}^{m} N_i^{q-1}\dot{N}_i$$

Then evaluate the gradient of the speed with respect to the vector of controls u_i considered as frozen parameters.

$$\nabla_{u_i}\dot{S}(q) = \frac{q}{(1-q)N^q}N_i^{q-1}$$

And finally define actual controls proportionally to the projection of the speed-gradient to the surface of constraints (19) according to SG eq. (6).

$$u_i = \dot{N}_i = \Gamma\left(\frac{q}{(1-q)N^q}N_i^{q-1} + \lambda\right)$$

Now we can evaluate Lagrange multiplier λ:

$$\sum_{i=1}^{m}\dot{N}_i = 0 \Rightarrow \lambda = \frac{1}{m}\frac{q}{(q-1)N^q}\sum_{i=1}^{m} N_i^{q-1}$$

The final form of the system dynamics law is as follows:

$$\dot{N}_i = \Gamma\left(\frac{q}{(1-q)N^q}N_i^{q-1} + \frac{1}{m}\frac{q}{(q-1)N^q}\sum_{i=1}^{m}N_i^{q-1}\right) =$$

$$\Gamma\frac{q}{(q-1)N^q}\left(\frac{1}{m}\sum_{i=1}^{m}N_i^{q-1} - N_i^{q-1}\right) \quad (25)$$

Let us find the equilibrium mode which corresponds to asymptotic behavior of the variables N_i. In this mode $\dot{N}_i = 0$. Based on (25) it means that $mN_i^{q-1} = \sum_{i=1}^{m} N_i^{q-1}$ which is possible only when all N_i are equal. According to constraint (19) we have that $N_i = \frac{N}{m}$. This result corresponds to the maximum state of classical entropy and agrees with thermodynamics.

5.1. Equilibrium stability

Let us examine the stability of the equilibrium mode. We introduce Lyapunov function

$$V(X,q) = S_{max}(q) - S(q),$$

where $S_{max}(q)$ is a maximum possible value for Tsallis entropy with parameter q.

Evaluation of $\dot{V}$ yields

$$\dot{V}(q) = -\dot{S}(q) = \frac{q}{(q-1)N^q}\sum_{i=1}^{m}N_i^{q-1}\dot{N}_i =$$

$$= \Gamma(\frac{q}{1-q})^2\frac{1}{mN^{2q}}\left(\left(\sum_{i=1}^{m}N_i^{q-1}\right)^2 - m\sum_{i=1}^{m}(N_i^{q-1})^2\right)$$

Consider Cauchy-Bunyakovsky-Schwarz (CBS) inequality for two vectors $a, b \in \mathbb{R}^m$:

$$\left(\sum_{i=1}^{m}a_ib_i\right)^2 \leq \sum_{i=1}^{m}a_i^2\sum_{i=1}^{m}b_i^2 \quad (26)$$

Based on CBS inequality for vectors $a = (1, .., 1)$ and $b = (N_1^{q-1}, .., N_m^{q-1})$ we have that $\dot{V}(q) \leq 0$. Equality $\dot{V}(q) = 0$ holds if and only if all the values N_i are equal. This is the maximum of entropy state. Thus the law (25) provides global asymptotic stability of the maximum entropy state. The physical meaning of this law is nothing but moving along the direction of the maximum entropy production rate (direction of the fastest entropy growth).

6. Internal energy constraint

The case of more than one constraint can be treated in the same fashion. Suppose the energy conservation law (22) holds in addition to the mass conservation law (19).

Then the evolution law should have the form

$$u_i = -\Gamma\frac{q}{1-q}\frac{N_i^{q-1}}{N^q} + \lambda_1 E_i + \lambda_2, \quad (27)$$

where λ_1 and λ_2 are Lagrange multipliers that can be defined by substitution of (27) into (19) and (22).

The solutions for λ_1 and λ_2 are given by formulas

$$\begin{cases} \lambda_1 & = \Gamma \frac{q}{(q-1)N^q} \frac{\left(\sum_{i=1}^m N_i^{q-1} \sum_{i=1}^m E_i - m \sum_{i=1}^m N_i^{q-1} E_i\right)}{m \sum_{i=1}^m E_i^2 - (\sum_{i=1}^m E_i)^2} \\ \lambda_2 & = \Gamma \frac{q}{(q-1)N^q} \frac{\left(\sum_{i=1}^m N_i^{q-1} E_i \sum_{i=1}^m E_i - \sum_{i=1}^m N_i^{q-1} \sum_{i=1}^m E_i^2\right)}{m \sum_{i=1}^m E_i^2 - (\sum_{i=1}^m E_i)^2} \end{cases} \tag{28}$$

This solution is defined for $(m \sum_{i=1}^m E_i^2 - (\sum_{i=1}^m E_i)^2) \neq 0$. It holds for all cases except a degenerate case when all E_i are equal.

General form of evolution law can be obtained by substitution of λ_1 and λ_2 from (28) into equation (27). In abbreviated form we represent this law as

$$\dot{\mathcal{N}} = -\Gamma' \frac{q}{1-q} (I - \Psi_{Ts}) \mathcal{N}^{q-1}, \tag{29}$$

where $\mathcal{N} = (N_1, ..., N_m)^T$, I is the $m \times m$ identity matrix, $\Gamma' = \frac{\Gamma}{N^q}$, Ψ_{Ts} is a symmetric $m \times m$ matrix defined as follows:

$$\psi_{ij} = \frac{1}{m} + \frac{\tilde{E}_i \tilde{E}_j}{\|\overline{E}\|^2 - \frac{1}{m}(\mathbf{1}\overline{E})^2},$$

where $\tilde{E}_i = E_i + \frac{1}{m} \sum_{i=1}^m E_i$, $\mathbf{1} = (1, ..., 1)$ and $\overline{E} = (E_1, ..., E_m)^T$ is a vector of energies.

6.1. Equilibrium stability

As before it can be shown that $V(X, q) = S_{max}(q) - S(X, q)$ is Lyapunov function and there is a unique stable equilibrium state of the system in non-degenerate cases. Let us demonstrate it.

$$\dot{V}(q) = -\dot{S}(q) = \frac{q}{(q-1)N^q} \sum_{i=1}^m N_i^{q-1} u_i \tag{30}$$

We substitute λ_1 and λ_2 from (28) into equation (27) and the expression for u_i we substitute into (30). Result expression is

$$\dot{V}(q) = \frac{\Gamma}{m} \left(\frac{q}{(q-1)N^q}\right)^2 \left(\frac{\left(m \sum E_i N_i^{q-1} - \sum E_i \sum N_i^{q-1}\right)^2}{m \sum E_i^2 - (\sum E_i)^2} - \left(m \sum \left(N_i^{q-1}\right)^2 - \left(\sum N_i^{q-1}\right)^2 \right) \right) \tag{31}$$

We introduce a new scalar product function for two vectors as

$$\langle f, g \rangle = m \sum_i^m f_i g_i - \sum_i^m f_i \sum_i^m g_i \tag{32}$$

For scalar product (32) the CBS inequality is true:

$$\langle f, g \rangle^2 \leq \langle f, f \rangle \langle g, g \rangle \tag{33}$$

An equality in (33) takes place only when $\exists \lambda, \mu \in \mathbb{R} : f = \lambda g + \mu$.

Using inequality (33) for vectors $f = (N_1^{q-1}, ..., N_m^{q-1})^T$ and $g = (E_1, ..., E_m)^T$ we get for (31) that $\dot{V}(q) \leq 0$. And $\dot{V}(q) = 0$ occurs for the only one case when $\exists \lambda, \mu \in \mathbb{R} : N_i^{q-1} = \lambda E_i + \mu$ for all i. Due to (27) at the equilibrium state of the system the following equalities hold:

$$\Gamma \frac{q}{q-1} \frac{N_i^{q-1}}{N^q} + \lambda_1 E_i + \lambda_2 = 0 \tag{34}$$

Which means that for equilibrium state the final distribution is unique and satisfy equation

$$N_i^{q-1} = \lambda E_i + \mu,$$

where $\lambda = \lambda_1 \frac{1-q}{\Gamma q} N^q$, $\mu = \lambda_2 \frac{1-q}{\Gamma q} N^q$ for λ_1 and λ_2 defined in (28).

6.2. Correspondence to the Tsallis distribution

Let us show that distribution in (34) corresponds to the Tsallis distribution which is equilibrium distribution for Tsallis MaxEnt with a given set of constraints. This distribution is defined in [2] as:

$$p_i = \frac{1}{Z_q} (1 - \beta(q-1)E_i)^{\frac{1}{q-1}} \tag{35}$$

where β is a special Lagrange multiplier and $Z_q = \sum_i (1 - \beta(q-1)E_i)^{\frac{1}{q-1}}$ stands for the normalization constant.

From (34) we get that

$$N_i = \left((\lambda_1 E_i + \lambda_2) \frac{1-q}{\Gamma q} N^q \right)^{\frac{1}{q-1}} \tag{36}$$

Let us substitute N_i from (36) into (19). We get

$$\frac{1}{N} \left(\frac{1-q}{\Gamma q} N^q \right)^{\frac{1}{q-1}} \sum_{i=1}^{m} (\lambda_1 E_i + \lambda_2)^{\frac{1}{q-1}} = 1 \Rightarrow$$

$$\left(\frac{1-q}{\Gamma q} N^q \right)^{\frac{1}{q-1}} = \frac{N}{\sum_{i=1}^{m} (\lambda_1 E_i + \lambda_2)^{\frac{1}{q-1}}} \tag{37}$$

After substitution (37) into (36) we get that

$$\frac{N_i}{N} = \frac{(\lambda_1 E_i + \lambda_2)^{\frac{1}{q-1}}}{\sum_{i=1}^{m} (\lambda_1 E_i + \lambda_2)^{\frac{1}{q-1}}} \tag{38}$$

Let us denote

$$\frac{\lambda_1}{\lambda_2} = -\beta(q-1) \tag{39}$$

As $\lambda_1 E_i + \lambda_2 = \lambda_2 \left(1 + \frac{\lambda_1}{\lambda_2} E_i\right)$ and taking into account eq. (39), the eq. (37) can be transformed to

$$\frac{N_i}{N} = \frac{(1 - \beta(q-1))^{\frac{1}{q-1}}}{\sum_{i=1}^{m} (1 - \beta(q-1))^{\frac{1}{q-1}}}, \tag{40}$$

where $\beta = -\frac{1}{q-1}\frac{\lambda_1}{\lambda_2}$. We can see that (40) coincides with Tsallis distribution (35). As mentioned in [16], β in eq. (35) is not the Lagrange multiplier associated to the internal energy constraint (which is λ_1 in our notation). Following by notation of C. Tsallis (see eq.(10) in [2]) we have that $\lambda_1 = -\lambda_2\beta(q-1)$. It explains the variable substitution in (39).

It is evident that (40) satisfies the normalization constraint (19). Let us check that the second constraint for energy (22) is also satisfied.

Let us substitute N_i from (36) into (22). Then we get

$$\frac{1}{N}\left(\frac{1-q}{\Gamma q}N^q\right)^{\frac{1}{q-1}}\sum_{i=1}^{m}(\lambda_1 E_i+\lambda_2)^{\frac{1}{q-1}}E_i = E \Rightarrow$$
$$\left(\frac{1-q}{\Gamma q}N^q\right)^{\frac{1}{q-1}} = \frac{NE}{\sum_{i=1}^{m}(\lambda_1 E_i+\lambda_2)^{\frac{1}{q-1}}E_i} \quad (41)$$

After substitution (41) into (36) we get that

$$N_i = \left((\lambda_1 E_i+\lambda_2)^{\frac{1}{q-1}}\right)\left(\frac{NE}{\sum_{i=1}^{m}(\lambda_1 E_i+\lambda_2)^{\frac{1}{q-1}}E_i}\right) \Rightarrow$$
$$E_iN_i = \frac{(\lambda_1 E_i+\lambda_2)^{\frac{1}{q-1}}NE}{\sum_{i=1}^{m}(\lambda_1 E_i+\lambda_2)^{\frac{1}{q-1}}} \Rightarrow \sum_{i=1}^{m}N_iE_i = NE$$

Which means that internal energy constraint (22) is true for (40).

7. Conclusions

We have investigated non-stationary states of processes that follow the MaxEnt principle for Shannon and Tsallis entropies. We have derived equations (21), (23), (25) and (29) which describe dynamics of the system that tends to the state with maximum entropy. Systems with discrete probability distribution of states were considered under mass conservation and energy conservation constraints. We have shown that the limit distribution is unique and corresponds to Gibbs-Jaynes in the case of Shannon entropy and Tsallis distribution for Tsallis entropy.

Both Shannon and Tsallis entropies can be also defined for continuous probability distributions. Methods described in this paper are possible to extend for probability density functions (pdf).

The key point of our approach is using the SG-method with the goal function chosen as the entropy of the system. SG principle generates equations for the transient (non-stationary) states of the system operation, i.e. it gives an answer to the question of **How the system will evolve?** This fact distinguishes the SG principle from the principle of maximum entropy, the principle of maximum Fisher information and others characterizing the steady-state processes and providing an answer to the questions of **To where?** and **How far?**

References

1. Shannon, C. A mathematical theory of communication. *Bell System Technical Journal* **1948**, *27*, 379–423.

2. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics* **1988**, *52*, 479–487.
3. Jaynes, E. The minimum entropy production principle. *Ann. Rev. Phys. Chem* **1980**, *31*, 579–601.
4. Jaynes, E. Information theory and statistical mechanics. *Phys. Rev* **1957**, *106*, 620–630.
5. Chen, Y. The rank-size scaling law and entropy-maximizing principle. *Physica A: Statistical Mechanics and its Applications* **2012**, *391*, 767–778.
6. Martyushev, L.M. Entropy and Entropy Production: Old Misconceptions and New Breakthroughs. *Entropy* **2013**, *15*, 1152–1170.
7. Fanelli, D.; Ninno, G.D.; Turchi, A. On the paradoxical behavior of a cyclic device working with a non-Boltzmannian fluid. *Commun Nonlinear Sci Numer Simulat* **2012**, pp. 4174–4180.
8. Fradkov, A.L. *Cybernetical physics: from control of chaos to quantum control*; Springer–Verlag, 2007.
9. Fradkov, A.L. Speed-gradient entropy principle for nonstationary processes. *Entropy* **2008**, *10*, 757–764.
10. Fradkov, A.; Krivtsov, A. Speed-gradient principle for description of transient dynamics in systems obeying maximum entropy principle. Bayesian inference and maximum entropy methods in science and engineering. AIP Conference Proceedings, 2010, Vol. 1305, pp. 399–406.
11. Grmela, M.; Öttinger, H. Dynamics and thermodynamics of complex fluids. I. Development of a general formalism. *Physical Review E* **1997**, *56 (6)*, 6620.
12. Öttinger, H.; Grmela, M. Dynamics and thermodynamics of complex fluids. II. Illustrations of a general formalism. *Physical Review E* **1997**, *56 (6)*, 6633.
13. Lanczos, K. *The Variational Principles of Mechanics*; Toronto, 1962.
14. Planck, M. The unity of physical world views. *Physikalische Zeitschrift* **1909**, *10*, 62–75.
15. Fradkov, A.; Miroshnik, I.; Nikiforov, V. *Nonlinear and Adaptive Control of Complex Systems*; Kluwer Acad. Publ., 1999.
16. Tsallis, C.; Mendes, R.; Plastino, A. The role of constraints within generalized nonextensive statistics. *Physica A* **1998**, *261*, 534–554.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).