

Article

General H -theorem and Entropies that Violate the Second Law

Alexander N. Gorban

Department of Mathematics, University of Leicester, University Road, Leicester, LE1 7RH, UK;

E-Mail: ag153@le.ac.uk

Received: 9 March 2014; in revised form: 15 April 2014 / Accepted: 24 April 2014 /

Published: 29 April 2014

Abstract: H -theorem states that the entropy production is nonnegative and, therefore, the entropy of a closed system should monotonically change in time. In information processing, the entropy production is positive for random transformation of signals (the information processing lemma). Originally, the H -theorem and the information processing lemma were proved for the classical Boltzmann-Gibbs-Shannon entropy and for the correspondent divergence (the relative entropy). Many new entropies and divergences have been proposed during last decades and for all of them the H -theorem is needed. This note proposes a simple and general criterion to check whether the H -theorem is valid for a convex divergence H and demonstrates that some of the popular divergences obey no H -theorem. We consider systems with n states A_i that obey first order kinetics (master equation). A convex function H is a Lyapunov function for all master equations with given equilibrium if and only if its conditional minima properly describe the equilibria of pair transitions $A_i \rightleftharpoons A_j$. This theorem does not depend on the principle of detailed balance and is valid for general Markov kinetics. Elementary analysis of pair equilibria demonstrate that the popular Bregman divergences like Euclidian distance or Itakura-Saito distance in the space of distribution cannot be the universal Lyapunov functions for the first-order kinetics and can increase in Markov processes. Therefore, they violate the second law and the information processing lemma. In particular, for these measures of information (divergences) random manipulation with data may add information to data. The main results are extended to nonlinear generalized mass action law kinetic equations.

Keywords: Markov process; Lyapunov function; non-classical entropy; information processing; quasiconvexity; directional convexity; Schur convexity

1. The Problem

The first non-classical entropy was proposed by Rényi in 1960 [1]. In the same paper he discovered the very general class of divergences, the so-called f -divergences (or Csiszár-Morimoto divergences because of the works of Csiszár [2] and Morimoto [3] published simultaneously in 1963):

$$H_h(p) = H_h(P||P^*) = \sum_i p_i^* h\left(\frac{p_i}{p_i^*}\right) \quad (1)$$

where $P = (p_i)$ is a probability distribution, P^* is an equilibrium distribution, $h(x)$ is a convex function defined on the open ($x > 0$) or closed $x \geq 0$ semi-axis. We use here the notation $H_h(P||P^*)$ to stress the dependence of H_h on both p_i and p_i^* .

These divergences have the form of the relative entropy or, in the thermodynamic terminology, the (negative) free entropy, the Massieu-Planck functions [4], or F/RT where F is the free energy. They measure the deviation of the current distribution P from the equilibrium P^* .

After 1961, many new entropies and divergences were invented and applied to real problems, including Burg entropy [5], Cressie-Red family of power divergences [6], Tsallis entropy [7,8], families of α -, β - and γ -divergences [9] and many others (see the review papers [10,11]). Many of them have the f -divergence form, but some of them do not. For example, the squared Euclidean distance from P to P^* is not, in general, a f -divergence unless all p_i^* are equal (equidistribution). Another example gives the Itakura-Saito distance:

$$\sum_i \left(\frac{p_i}{p_i^*} - \ln \frac{p_i}{p_i^*} - 1 \right) \quad (2)$$

The idea of Bregman divergences [12] provides a new general source of divergences different from the f -divergences. Any strictly convex function F in an closed convex set V satisfies the Jensen inequality

$$D_F(p, q) = F(p) - F(q) - (\nabla_q F(q), p - q) > 0 \quad (3)$$

if $p \neq q$, $p, q \in V$. This positive quantity $D_F(p, q)$ is the Bregman divergence associated with F . For example, for a positive quadratic form $F(x)$ the Bregman distance is just $D_F(p, q) = F(p - q)$. In particular, if F is the squared Euclidean length of x then $D_F(p, q)$ is the squared Euclidean distance. If F is the Burg entropy, $F(x) = -\sum_i \ln p_i$, then $D_F(p, q)$ is the Itakura-Saito distance. The Bregman divergences have many attractive properties. For example, the mean vector minimizes the expected Bregman divergence from the random vector [13]. The Bregman divergences are convenient for numerical optimization because generalized Pythagorean identity [14]. Nevertheless, for information processing and for many physical applications one more property is crucially important.

The divergence between the current distribution and equilibrium should monotonically decrease in Markov processes. It is the ultimate requirement for use of the divergence in information processing and in non-equilibrium thermodynamics and kinetics. In physics, the first result of this type was Boltzmann's H -theorem proven for nonlinear kinetic equation. In information theory, Shannon [15] proved this theorem for the entropy ("the data processing lemma") and Markov chains.

In his well-known paper [1], Rényi also proved that $H_h(P||P^*)$ monotonically decreases in Markov processes (he gave the detailed proof for the classical relative entropy and then mentioned that for the f -divergences it is the same). This result, elaborated further by Csiszár [2] and Morimoto [3],

embraces many later particular H -theorems for various entropies including the Tsallis entropy and the Rényi entropy (because it can be transformed into the form (1) by a monotonic function, see for example [11]). The generalized data processing lemma was proven [16,17]: for every two positive probability distributions P, Q the divergence $H_h(P\|Q)$ decreases under action of a stochastic matrix $A = (a_{ij})$

$$H_h(AP\|AQ) \leq \bar{\alpha}(A)H_h(P\|Q) \tag{4}$$

where

$$\bar{\alpha}(A) = \frac{1}{2} \max_{i,k} \left\{ \sum_j |a_{ij} - a_{kj}| \right\} \tag{5}$$

is the ergodicity contraction coefficient, $0 \leq \bar{\alpha}(A) \leq 1$. Here, neither Q nor P must be the equilibrium distribution: divergence between any two distributions decreases in Markov processes.

Under some additional conditions, the property to decrease in Markov processes characterizes the f -divergences [18,19]. For example, if a divergence decreases in all Markov processes, does not change under permutation of states and can be represented as a sum over states (has the trace form), then it is the f -divergence [11,18].

The dynamics of distributions in the continuous time Markov processes is described by the master equation. Thus, the f -divergences are the Lyapunov functions for the master equation. The important property of the divergences $H_h(P\|P^*)$ is that they are *universal* Lyapunov functions. That is, they depend on the current distribution P and on the equilibrium P^* but do not depend on the transition probabilities directly.

For each new divergence we have to analyze its behavior in Markov processes and to prove or refute the H -theorem. For this purpose, we need a simple and general criterion. It is desirable to avoid any additional requirements like the trace form or symmetry. In this paper we develop this criterion.

It is obvious that the equilibrium P^* is a global minimum of any universal Lyapunov function $H(P)$ in the simplex of distributions (see the model equation below). In brief, the *general H-theorem* states that a convex function $H(P)$ is a universal Lyapunov function for the master equation if and only if its conditional minima correctly describe the partial equilibria for pairs of transitions $A_i \rightleftharpoons A_j$. These partial equilibria are given by proportions $p_i/p_i^* = p_j/p_j^*$. They should be solutions to the problem

$$\begin{aligned} H(P) \rightarrow \min \quad & \text{subject to } p_k \geq 0 \ (k = 1, \dots, n), \\ \sum_{k=1}^n p_k = 1, \quad & \text{and given values of } p_l \ (l \neq i, j) \end{aligned} \tag{6}$$

These solutions are minima of $H(P)$ on segments $p_i + p_j = 1 - \sum_{l \neq i,j} p_l, p_{i,j} \geq 0$. They depend on $n - 2$ parameters $p_l \geq 0 \ (l \neq i, j, \sum_{l \neq i,j} p_l < 1)$.

Using this general H -theorem we analyze several Bregman divergences that are not f -divergences and demonstrate that they do not allow the H -theorem even for systems with three states. We present also the generalizations of the main results for Generalized Mass Action Law (GMAL) kinetics.

2. Three Forms of Master Equation and the Decomposition Theorem

We consider continuous time Markov chains with n states A_1, \dots, A_n . The *Kolmogorov equation* or *master equation* for the probability distribution P with the coordinates p_i (we can consider P as a vector-column $P = [p_1, \dots, p_n]^T$) is

$$\frac{dp_i}{dt} = \sum_{j, j \neq i} (q_{ij}p_j - q_{ji}p_i) \quad (i = 1, \dots, n) \quad (7)$$

where q_{ij} ($i, j = 1, \dots, n, i \neq j$) are nonnegative. In this notation, q_{ij} is the *rate constant* for the transition $A_j \rightarrow A_i$. Any set of nonnegative coefficients q_{ij} ($i \neq j$) corresponds to a master equation. Therefore, the class of the master equations can be represented as a nonnegative orthant in $\mathbb{R}^{n(n-1)}$ with coordinates q_{ij} ($i \neq j$). Equations of the same class describe any first order kinetics in perfect mixtures. The only difference between the general first order kinetics and master equation for the probability distribution is in the balance conditions: the sum of probabilities should be 1, whereas the sum of variables (concentrations) for the general first order kinetics may be any positive number.

It is useful to mention that the *model equation* with equilibrium P^* and relaxation time τ

$$\frac{dp_i}{dt} = \frac{1}{\tau}(p_i^* - p_i) \quad (i = 1, \dots, n) \quad (8)$$

is a particular case of master equation for normalized variables p_i ($p_i \geq 0, \sum_i p_i = 1$). Indeed, let us take in Equation (7) $q_{ij} = \frac{1}{\tau}p_i^*$.

The *graph of transitions* for a Markov chain is a directed graph. Its vertices correspond to the states A_i and the edges correspond to the transitions $A_j \rightarrow A_i$ with the positive transition coefficients, $q_{ij} > 0$. The digraph of transitions is *strongly connected* if there exists an oriented path from any vertex A_i to every other vertex A_j ($i \neq j$). The continuous-time Markov chain is *ergodic* if there exists a unique strictly positive equilibrium distribution P^* ($p_i^* > 0, \sum_i p_i^* = 1$) for master Equation (7) [20,21]. Strong connectivity of the graph of transitions is necessary and sufficient for ergodicity of the corresponding Markov chain.

A digraph is *weakly connected* if the underlying undirected graph obtained by replacing directed edges by undirected ones is connected. The maximal weakly connected components of a digraph are called connected (or weakly connected) components. The maximal strongly connected subgraphs are called strong components. The necessary and sufficient condition for the *existence* of a strongly positive equilibrium for master Equation (7) is: *the weakly connected components of the transition graph are its strong components*. An equivalent form of this condition is: *if there exists a directed path from A_i to A_j , then there exists a directed path from A_j to A_i* . In chemical kinetics this condition is sometimes called the “weak reversibility” condition [22,23]. This implies that the digraph is the union of disjoint strongly connected digraphs. For each strong component of the transition digraph the normalized equilibrium is unique and the equilibrium for the whole graph is a convex combination of positive normalized equilibria for its strong components. If m is the number of these components then the set of normalized positive equilibria of master equation ($P^*: p_i^* > 0, \sum_i p_i^* = 1$) is a relative interior of a $m - 1$ -dimensional polyhedron in the unit simplex Δ_n . The set of non-normalized positive equilibria ($P^*: p_i^* > 0$) is a relative interior of a m -dimensional cone in the positive orthant \mathbb{R}_+^n .

The Markov chain in Equation (7) is *weakly ergodic* if it allows the only conservation law: the sum of coordinates, $\sum_i p_i \equiv const$. Such a system forgets its initial condition: the distance between any two trajectories with the same value of the conservation law tends to zero when time goes to infinity. Among all possible norms, the l_1 distance ($\|P - Q\|_{l_1} = \sum_i |p_i - q_i|$) plays a special role: it does not increase in time for any first order kinetic system in master Equation (7) and strongly monotonically decreases to zero for normalized probability distributions ($\sum_i p_i = \sum_i q_i = 1$) and weakly ergodic chains. The difference between weakly ergodic and ergodic systems is in the obligatory existence of a *strictly positive* equilibrium for an ergodic system. A Markov chain is weakly ergodic if and only if for each two vertices A_i, A_j ($i \neq j$) we can find such a vertex A_k that is reachable by oriented paths both from A_i and from A_j . This means that the following structure exists [24]:

$$A_i \rightarrow \dots \rightarrow A_k \leftarrow \dots \leftarrow A_j . \tag{9}$$

One of the paths can be degenerated: it may be $i = k$ or $j = k$.

Now, let us restrict our consideration to the set of the Markov chains with the given positive equilibrium distribution P^* ($p_i^* > 0$). We do not assume that this distribution is compulsory unique. The transition graph should be the union of disjoint strongly connected digraphs (in particular, it may be strongly connected). Using the known positive equilibrium P^* we can rewrite master Equation (7) in the following form

$$\frac{dp_i}{dt} = \sum_{j, j \neq i} q_{ij} p_j^* \left(\frac{p_j}{p_j^*} - \frac{p_i}{p_i^*} \right) \tag{10}$$

where p_i^* and q_{ij} are connected by the *balance equation*

$$\sum_{j, j \neq i} q_{ij} p_j^* = \left(\sum_{j, j \neq i} q_{ji} \right) p_i^* \text{ for all } i = 1, \dots, n \tag{11}$$

For the next transformation of master equation we join the mutually reverse transitions in pairs $A_i \rightleftharpoons A_j$ in pairs (say, $i > j$) and introduce the *stoichiometric vectors* γ^{ji} with coordinates:

$$\gamma_k^{ji} = \begin{cases} -1 & \text{if } k = j, \\ 1 & \text{if } k = i, \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

Let us rewrite the master Equation (7) in the *quasichemical form*:

$$\frac{dP}{dt} = \sum_{i > j} (w_{ij}^+ - w_{ij}^-) \gamma^{ji} \tag{13}$$

where $w_{ij}^+ = q_{ij} p_j^* \frac{p_j}{p_j^*}$ is the rate of the transitions $A_j \rightarrow A_i$ and $w_{ij}^- = q_{ji} p_i^* \frac{p_i}{p_i^*}$ is the rate of the reverse process $A_j \leftarrow A_i$ ($i > j$).

The reversible systems with detailed balance form an important class of first order kinetics. The *detailed balance* condition reads [25]: at equilibrium, $w_{ij}^+ = w_{ij}^-$, i.e.,

$$q_{ij} p_j^* = q_{ji} p_i^* (= w_{ij}^*) \quad i, j = 1, \dots, n \tag{14}$$

Here, w_{ij}^* is the *equilibrium flux* from A_i to A_j and back.

For the systems with detailed balance the quasichemical form of the master equation is especially simple:

$$\frac{dP}{dt} = \sum_{i>j} w_{ij}^* \left(\frac{p_j}{p_j^*} - \frac{p_i}{p_i^*} \right) \gamma^{ji} \tag{15}$$

It is important that any set of nonnegative equilibrium fluxes w_{ij}^* ($i > j$) defines by Equation (15) a system with detailed balance with a given positive equilibrium P^* . Therefore, the set of all systems with detailed balance presented by Equation (15) and a given equilibrium may be represented as a nonnegative orthant in $\mathbb{R}^{\frac{n(n-1)}{2}}$ with coordinates w_{ij}^* ($i > j$).

The *decomposition theorem* [26,27] states that for any given positive equilibrium P^* and any positive distribution P the set of possible values dP/dt for Equations (13) under the balance condition (11) coincides with the set of possible values dP/dt for Equations (15) under detailed balance condition (14).

In other words, for every general system of the form (13) with positive equilibrium P^* and any given non-equilibrium distribution P there exists a system with detailed balance of the form (15) with the same equilibrium and the same value of the velocity vector dP/dt at point P . Therefore, the sets of the universal Lyapunov function for the general master equations and for the master equations with detailed balance coincide.

3. General H-Theorem

Let $H(P)$ be a convex function on the space of distributions. It is a Lyapunov function for a master equations with the positive equilibrium P^* if $dH(P(t))/dt \leq 0$ for any positive normalized solution $P(t)$. For a system with detailed balance given by Equation (15)

$$\frac{dH(P(t))}{dt} = - \sum_{i>j} w_{ij}^* \left(\frac{p_j}{p_j^*} - \frac{p_i}{p_i^*} \right) \left(\frac{\partial H(P)}{\partial p_j} - \frac{\partial H(P)}{\partial p_i} \right) \tag{16}$$

The inequality $dH(P(t))/dt \leq 0$ is true for all nonnegative values of w_{ij}^* if and only if it holds for any term in Equation (16) separately. That is, for any pair i, j ($i > j$) the convex function $H(P)$ is a Lyapunov function for the system (15) where one and only one w_{ij}^* is not zero.

A convex function on a straight line is a Lyapunov function for a one-dimensional system with single equilibrium if and only if the equilibrium is a minimizer of this function. This elementary fact together with the previous observation gives us the criterion for universal Lyapunov functions for systems with detailed balance. Let us introduce the *partial equilibria criterion*:

Definition 1 (Partial equilibria criterion). *A convex function $H(P)$ on the simplex Δ_n of probability distributions satisfies the partial equilibria criterion with a positive equilibrium P^* if the proportion $p_i/p_i^* = p_j/p_j^*$ give the minimizers in the problem (6).*

Proposition 1. *A convex function $H(P)$ on the simplex Δ_n of probability distributions is a Lyapunov function for all master equations with the given equilibrium P^* that obey the principle of detailed balance if and only if it satisfies the partial equilibria criterion with the equilibrium P^* .*

Combination of this Proposition with the decomposition theorem [26] gives the same criterion for general master equations without hypothesis about detailed balance

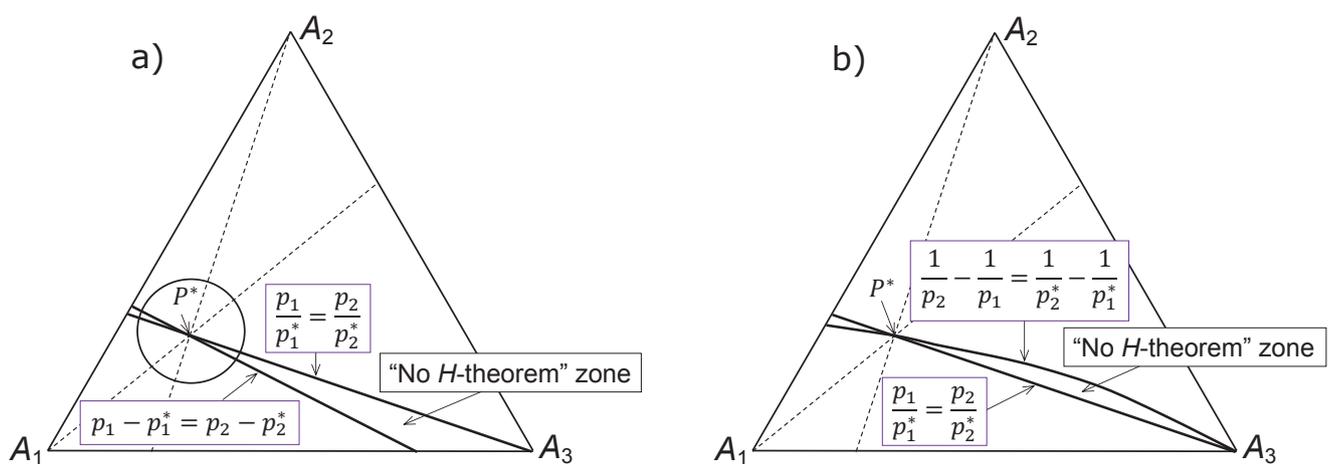
Proposition 2. A convex function $H(P)$ on the simplex Δ_n of probability distributions is a Lyapunov function for all master equations with the given equilibrium P^* if and only if it satisfies the partial equilibria criterion with the equilibrium P^* .

These two propositions together form the general H -theorem.

Theorem 1. The partial equilibria criterion with a positive equilibrium P^* is a necessary condition for a convex function to be the universal Lyapunov function for all master equations with detailed balance and equilibrium P^* and a sufficient condition for this function to be the universal Lyapunov function for all master equations with equilibrium P^* .

Let us stress that here the partial equilibria criterion provides a necessary condition for systems with detailed balance (and, therefore, for the general systems without detailed balance assumption) and a sufficient condition for the general systems (and, therefore, for the systems with detailed balance too).

Figure 1. The triangle of distributions for the system with three states A_1, A_2, A_3 and the equilibrium $p_1^* = \frac{4}{7}, p_2^* = \frac{2}{7}, p_3^* = \frac{1}{7}$. The lines of partial equilibria $A_i \rightleftharpoons A_j$ given by the proportions $p_i/p_i^* = p_j/p_j^*$ are shown, for $A_1 \rightleftharpoons A_2$ by solid straight lines (with one end at the vertex A_3), for $A_2 \rightleftharpoons A_3$ and for $A_1 \rightleftharpoons A_3$ by dashed lines. The lines of conditional minima of $H(P)$ in problem (6) are presented for the partial equilibrium $A_1 \rightleftharpoons A_2$ (a) for the squared Euclidean distance (a circle here is an example of the $H(P)$ level set) and (b) for the Itakura-Saito distance. Between these lines and the line of partial equilibria the “no H -theorem zone” is situated. In this zone, $H(P)$ increases in time for some master equations with equilibrium P^* . Similar zones (not shown) exist near other partial equilibrium lines too. Outside these zones, $H(P)$ monotonically decreases in time for any master equation with equilibrium P^* .



4. Examples

The simplest Bregman divergence is the squared Euclidean distance between P and P^* , $\sum_i (p_i - p_i^*)^2$. The solution to the problem (6) is: $p_i - p_i^* = p_j - p_j^*$. Obviously, it differs from the proportion required by the partial equilibria criterion $\frac{p_i}{p_j} = \frac{p_i^*}{p_j^*}$ (Figure 1a).

For the Itakura-Saito distance (2) the solution to the problem (6) is: $\frac{1}{p_i} - \frac{1}{p_i^*} = \frac{1}{p_j} - \frac{1}{p_j^*}$. It also differs from the proportion required (Figure 1b).

If the single equilibrium in 1D system is not a minimizer of a convex function H then $dH/dt > 0$ on the interval between the equilibrium and minimizer of H (or minimizers if it is not unique). Therefore, if $H(P)$ does not satisfy the partial equilibria criterion then in the simplex of distributions there exists an area bordered by the partial equilibria surface for $A_i \rightleftharpoons A_j$ and by the minimizers for the problem (6), where for some master equations $dH/dt > 0$ (Figure 1). In particular, in such an area $dH/dt > 0$ for the simple system with two mutually reverse transitions, $A_i \rightleftharpoons A_j$, and the same equilibrium.

If H satisfies the partial equilibria criterion, then the minimizers for the problem (6) coincide with the partial equilibria surface for $A_i \rightleftharpoons A_j$, and the “no H -theorem zone” vanishes.

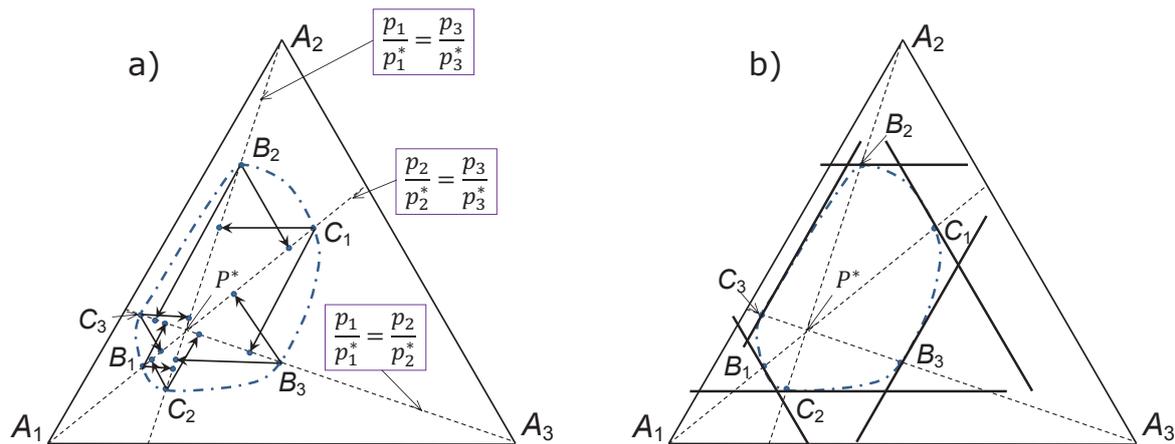
The partial equilibria criterion allows a simple geometric interpretation. Let us consider a sublevel set of $H(P)$ in the simplex Δ_n : $U_h = \{P \in \Delta_n \mid H(P) \leq h\}$. Let the level set be $L_h = \{P \in \Delta_n \mid H(P) = h\}$. For the partial equilibrium $A_i \rightleftharpoons A_j$ we use the notation E_{ij} . It is given by the equation $p_i/p_i^* = p_j/p_j^*$. The geometric equivalent of the partial equilibrium condition is: for all i, j ($i \neq j$) and every $P \in L_h \cap E_{ij}$ the straight line $P + \lambda\gamma_{ij}$ ($\lambda \in \mathbb{R}$) is a supporting line of U_h . This means that this line does not intersect the interior of U_h .

We illustrate this condition on the plane for three states in Figure 2. The level set of H is represented by the dot-dash line. It intersects the lines of partial equilibria (dashed lines) at points $B_{1,2,3}$ and $C_{1,2,3}$. For each point P from these six intersections ($P = B_i$ and $P = C_i$) the line $P + \lambda\gamma_{jk}$ ($\lambda \in \mathbb{R}$) should be a supporting line of the sublevel set (the region bounded by the dot-dash line). Here, i, j, k should all be different numbers. Segments of these lines form a hexagon circumscribed around the level set (Figure 2b).

The points of intersection $B_{1,2,3}$ and $C_{1,2,3}$ cannot be selected arbitrarily on the lines of partial equilibria. First of all, they should be the vertices of a convex hexagon with the equilibrium P^* inside. Secondly, due to the partial equilibria criterion, the intersections of the straight line $P + \lambda\gamma_{ij}$ with the partial equilibria E_{ij} are the conditional minimizers of H on this line, and therefore should belong to the sublevel set $U_{H(P)}$. If we apply this statement to $P = B_i$ and $P = C_i$, then we will get two projections of this point onto partial equilibria E_{ij} parallel to γ_{ij} (Figure 2a). These projections should belong to the hexagon with the vertices $B_{1,2,3}$ and $C_{1,2,3}$. They produce a six-ray star that should be inscribed into the level set.

In Figure 2 we present the following characterization of the level set of a Lyapunov function for the Markov chains with three states. This convex set should be circumscribed around the six-ray star (Figure 2a) and inscribed in the hexagon of the supporting lines (Figure 2b).

Figure 2. Geometry of the Lyapunov function level set. The triangle of distributions for the system with three states A_1, A_2, A_3 and the equilibrium $p_1^* = \frac{4}{7}, p_2^* = \frac{2}{7}, p_3^* = \frac{1}{7}$. The lines of partial equilibria $A_i \rightleftharpoons A_j$ given by the proportions $p_i/p_i^* = p_j/p_j^*$ are shown by dashed lines. The dash-dot line is the level set of a Lyapunov function H . It intersects the lines of partial equilibria at points $B_{1,2,3}$ and $C_{1,2,3}$. (The points B_i are close to the vertices A_i , the points C_i belong to the same partial equilibrium but on another side of the equilibrium P^* .) For each point B_i, C_i the corresponding partial equilibria of two transitions $A_i \rightleftharpoons A_j$ ($j \neq i$) are presented (a). These partial equilibria should belong to the sublevel set of H . They are the projections of B_i, C_i onto the lines of partial equilibria $A_i \rightleftharpoons A_j$ ($j \neq i$) with projecting rays parallel to the sides $[A_i, A_j]$ of the triangle (i.e., to the stoichiometric vectors γ^{ji} (12)). The six-ray star with vertices B_i, C_i should be inside the dash-dot contour (a). Therefore, the projection of B_i onto the partial equilibrium $A_i \rightleftharpoons A_j$ should belong to the segment $[C_k, P^*]$ and the projection of C_i onto the partial equilibrium $A_i \rightleftharpoons A_j$ should belong to the segment $[B_k, P^*]$ (a). The lines parallel to the sides A_j, A_k of the triangle should be supporting lines of the level set of H at points B_i, C_i (i, j, k are different numbers) (b). Segments of these lines form a circumscribed hexagon around the level set (b).



All the f -divergences given by Equation (1) satisfy the partial equilibria criterion and are the universal Lyapunov functions but the reverse is not true: the class of universal Lyapunov functions is much wider than the set of the f -divergences. Let us consider the set “PEC” of convex functions $H(P||P^*)$, which satisfy the partial equilibria criterion. It is closed with respect to the following operations

- Conic combination: if $H_j(P||P^*) \in PEC$ then $\sum_j \alpha_j H_j(P||P^*) \in PEC$ for nonnegative coefficients $\alpha_j \geq 0$.
- Convex monotonic transformation of scale: if $H(P||P^*) \in PEC$ then $F(H(P||P^*)) \in PEC$ for any convex monotonically increasing function of one variable F .

Using these operations we can construct new universal Lyapunov functions from a given set. For example,

$$\frac{1}{2} \sum_i \frac{(p_i - p_i^*)^2}{p_i^*} + \prod_j \exp \frac{(p_j - p_j^*)^2}{2p_j^*}$$

is a universal Lyapunov function that does not have the f -divergence form because the first sum is an f -divergence given by Equation (1) with $h(x) = \frac{1}{2}(x - 1)^2$ and the product is the exponent of this f -divergence (exp is convex and monotonically increasing function).

The following function satisfies the partial equilibria criterion for every $\varepsilon > 0$.

$$\frac{1}{2} \sum_i \frac{(p_i - p_i^*)^2}{p_i^*} + \frac{\varepsilon}{4n^2} \prod_{i,j,i \neq j} (p_i p_j^* - p_j p_i^*)^2 \quad (17)$$

It is convex for $0 < \varepsilon < 1$. (Just apply the Gershgorin theorem [28] to the Hessian and use that all $p_i, p_i^* \leq 1$.) Therefore, it is a universal Lyapunov function for master equation in Δ_n if $0 < \varepsilon < 1$. The partial equilibria criterion together with the convexity condition allows us to construct many such examples.

5. General H -Theorem for Nonlinear Kinetics

5.1. Generalized Mass Action Law

Several formalisms are developed in chemical kinetics and non-equilibrium thermodynamics for the construction of general kinetic equations with a given “thermodynamic Lyapunov functional”. The motivation of this approach “from thermodynamics to kinetics” is simple [29,30]: (i) the thermodynamic data are usually more reliable than data about kinetics and we know the thermodynamic functions better than the details of kinetic equations, and (ii) positivity of entropy production is a fundamental law and we prefer to respect it “from scratch”, by the structure of kinetic equations.

GMAL is a method for the construction of dissipative kinetic equations for a given thermodynamic potential H . Other general thermodynamic approaches [31–33] give similar results for a given stoichiometric algebra. Below we introduce GMAL following [29,30,34].

The *list of components* is a finite set of symbols A_1, \dots, A_n .

A *reaction mechanism* is a finite set of the *stoichiometric equations* of elementary reactions:



where $\rho = 1, \dots, m$ is the reaction number and the *stoichiometric coefficients* $\alpha_{\rho i}, \beta_{\rho i}$ are nonnegative numbers. Usually, these numbers are assumed to be integer but in some applications the construction should be more flexible and admit real nonnegative values. Let α_ρ, β_ρ be the vectors with coordinates $\alpha_{\rho i}, \beta_{\rho i}$ correspondingly.

A *stoichiometric vector* γ_ρ of the reaction in Equation (18) is a n -dimensional vector $\gamma_\rho = \beta_\rho - \alpha_\rho$ with coordinates

$$\gamma_{\rho i} = \beta_{\rho i} - \alpha_{\rho i} \quad (19)$$

that is, “gain minus loss” in the ρ th elementary reaction. We assume $\alpha_\rho \neq \beta_\rho$ to avoid trivial reactions with zero γ_ρ .

One of the standard assumptions is existence of a strictly positive stoichiometric conservation law, a vector $b = (b_i), b_i > 0$ such that $\sum_i b_i \gamma_{\rho i} = 0$ for all ρ . This may be the conservation of mass, of the total probability, or of the total number of atoms, for example.

A nonnegative extensive variable N_i , the *amount* of A_i , corresponds to each component. We call the vector N with coordinates N_i “the composition vector”. The concentration of A_i is an intensive variable $c_i = N_i/V$, where $V > 0$ is the volume. The vector $c = N/V$ with coordinates c_i is the vector of concentrations.

Let us consider a domain U in n -dimensional real vector space E with coordinates N_1, \dots, N_n . For each N_i , a dimensionless entropy (or *free entropy*, for example, Massieu, Planck, or Massieu-Planck potential that corresponds to the selected conditions [4]) $S(N)$ is defined in U . “Dimensionless” means that we use S/R instead of physical S . This choice of units corresponds to the informational entropy ($p \ln p$ instead of $k_B p \ln p$).

The dual variables, potentials, are defined as the partial derivatives of $H = -S$:

$$\check{\mu}_i = \frac{\partial H}{\partial N_i}, \quad \check{\mu} = \nabla_N H \quad (20)$$

This definition differs from the chemical potentials [4] by the factor $1/RT$. We keep the same sign as for the chemical potentials, and this differs from the standard Legendre transform for S . (It is the Legendre transform for the function $H = -S$.) The standard condition for the reversibility of the Legendre transform is strong positive definiteness of the Hessian of H .

For each reaction, a *nonnegative* quantity, reaction rate r_ρ is defined. We assume that this quantity has the following structure (compare with Equations (4), (7), and (14) in [32] and Equation (4.10) in [33]):

$$r_\rho = \varphi_\rho \exp(\alpha_\rho, \check{\mu}) \quad (21)$$

where $(\alpha_\rho, \check{\mu}) = \sum_i \alpha_{\rho i} \check{\mu}_i$ is the standard inner product. Here and below, $\exp(\cdot, \cdot)$ is the exponent of the standard inner product. The *kinetic factor* $\varphi_\rho \geq 0$ is an intensive quantity and the expression $\exp(\alpha_\rho, \check{\mu})$ is the *Boltzmann factor* of the ρ th elementary reaction.

In the standard formalism of chemical kinetics the reaction rates are intensive variables and in kinetic equations for N an additional factor—the volume—appears. For heterogeneous systems, there may be several “volumes” (including interphase surfaces).

A nonnegative extensive variable N_i , the amount of A_i , corresponds to each component. We call the vector N with coordinates N_i “the composition vector”. The concentration of A_i is an intensive variable $c_i = N_i/V$, where $V > 0$ is the volume. If the system is heterogeneous then there are several “volumes” (volumes, surfaces, *etc.*), and in each volume there are the composition vector and the vector of concentrations [30,34]. Here we will consider homogeneous systems.

The kinetic equations for a homogeneous system in the absence of external fluxes are

$$\frac{dN}{dt} = V \sum_\rho r_\rho \gamma_\rho = V \sum_\rho \gamma_\rho \varphi_\rho \exp(\alpha_\rho, \check{\mu}) \quad (22)$$

If the volume is not constant then the equations for concentrations include \dot{V} and have different form (this is typical for combustion reactions, for example).

The classical Mass Action Law gives us an important particular case of GMAL given by Equation (21). Let us take the perfect free entropy

$$S = - \sum_i N_i \left(\ln \left(\frac{c_i}{c_i^*} \right) - 1 \right) \quad (23)$$

where $c_i = N_i/V \geq 0$ are concentrations and $c_i^* > 0$ are the standard equilibrium concentrations.

For the perfect entropy function presented in Equation (23)

$$\check{\mu}_i = \ln \left(\frac{c_i}{c_i^*} \right), \quad \exp(\alpha_\rho, \check{\mu}) = \prod_i \left(\frac{c_i}{c_i^*} \right)^{\alpha_{\rho i}} \tag{24}$$

and for the GMAL reaction rate function given by (21) we get

$$r_\rho = \varphi_\rho \prod_i \left(\frac{c_i}{c_i^*} \right)^{\alpha_{\rho i}} \tag{25}$$

The standard assumption for the Mass Action Law in physics and chemistry is that φ and c^* are functions of temperature: $\varphi_\rho = \varphi_\rho(T)$ and $c_i^* = c_i^*(T)$. To return to the kinetic constants notation and in particular to first order kinetics in the quasichemical form presented in Equation (13), we should write:

$$\frac{\varphi_\rho}{\prod_i c_i^{*\alpha_{\rho i}}} = k_\rho \tag{26}$$

5.2. General Entropy Production Formula

Thus, the following entities are given: the set of components A_i ($i = 1, \dots, n$), the set of m elementary reactions presented by stoichiometric equations (18), the thermodynamic Lyapunov function $H(N, V, \dots)$ [4,30,35], where dots (marks of omission) stand for the quantities that do not change in time under given conditions, for example, temperature for isothermal processes or energy for isolated systems. The GMAL presents the reaction rate r_ρ in Equation (21) as a product of two factors: the Boltzmann factor and the kinetic factor. Simple algebra gives for the time derivative of H :

$$\begin{aligned} \frac{dH}{dt} &= \sum_i \frac{\partial H}{\partial N_i} \frac{dN_i}{dt} \\ &= \sum_i \check{\mu}_i V \sum_\rho \gamma_{\rho i} \varphi_\rho \exp(\alpha_\rho, \check{\mu}) \\ &= V \sum_\rho (\gamma_\rho, \check{\mu}) \varphi_\rho \exp(\alpha_\rho, \check{\mu}) \end{aligned} \tag{27}$$

An auxiliary function $\theta(\lambda)$ of one variable $\lambda \in [0, 1]$ is convenient for analysis of dS/dt (see [29,34,36]):

$$\theta(\lambda) = \sum_\rho \varphi_\rho \exp[(\check{\mu}, (\lambda\alpha_\rho + (1 - \lambda)\beta_\rho))] \tag{28}$$

With this function, \dot{H} defined by Equation (27) has a very simple form:

$$\frac{dH}{dt} = -V \left. \frac{d\theta(\lambda)}{d\lambda} \right|_{\lambda=1} \tag{29}$$

The auxiliary function $\theta(\lambda)$ allows the following interpretation. Let us introduce the deformed stoichiometric mechanism with the stoichiometric vectors,

$$\alpha_\rho(\lambda) = \lambda\alpha_\rho + (1 - \lambda)\beta_\rho, \quad \beta_\rho(\lambda) = \lambda\beta_\rho + (1 - \lambda)\alpha_\rho \tag{30}$$

which is the initial mechanism when $\lambda = 1$, the reverted mechanism with interchange of α and β when $\lambda = 0$, and the trivial mechanism (the left and right hand sides of the stoichiometric equations coincide) when $\lambda = 1/2$. Let the deformed reaction rate be $r_\rho(\lambda) = \varphi_\rho \exp(\alpha_\rho(\lambda), \check{\mu})$ (the genuine kinetic factor is combined with the deformed Boltzmann factor). Then $\theta(\lambda) = \sum_\rho r_\rho(\lambda)$.

It is easy to check that $\theta''(\lambda) \geq 0$ and, therefore, $\theta(\lambda)$ is a convex function.

The inequality

$$\theta'(1) \geq 0 \quad (31)$$

is *necessary and sufficient* for accordance between kinetics and thermodynamics (decrease of free energy or positivity of entropy production). This inequality is a condition on the kinetic factors. Together with the positivity condition $\varphi_\rho \geq 0$, it defines a convex cone in the space of vectors of kinetic factors φ_ρ ($\rho = 1, \dots, m$). There exist two less general and more restrictive *sufficient* conditions: detailed balance and complex balance (known also as semidetailed or cyclic balance).

5.3. Detailed Balance

The detailed balance condition consists of two assumptions: (i) for each elementary reaction $\sum_i \alpha_{\rho i} A_i \rightarrow \sum_i \beta_{\rho i} A_i$ in the mechanism (18) there exists a reverse reaction $\sum_i \beta_{\rho i} A_i \leftarrow \sum_i \alpha_{\rho i} A_i$. Let us join these reactions in pairs

$$\sum_i \alpha_{\rho i} A_i \rightleftharpoons \sum_i \beta_{\rho i} A_i \quad (32)$$

After this joining, the total number of stoichiometric equations decreases. We distinguish the reaction rates and kinetic factors for direct and inverse reactions by the upper plus or minus:

$$r_\rho^+ = \varphi_\rho^+ \exp(\alpha_\rho, \check{\mu}), \quad r_\rho^- = \varphi_\rho^- \exp(\beta_\rho, \check{\mu}) \quad (33)$$

The kinetic equations take the form

$$\frac{dN}{dt} = V \sum_\rho (r_\rho^+ - r_\rho^-) \gamma_\rho \quad (34)$$

The condition of detailed balance in GMAL is simple and elegant:

$$\varphi_\rho^+ = \varphi_\rho^- \quad (35)$$

For the systems with detailed balance we can take $\varphi_\rho = \varphi_\rho^+ = \varphi_\rho^-$ and write for the reaction rate:

$$r_\rho = r_\rho^+ - r_\rho^- = \varphi_\rho (\exp(\alpha_\rho, \check{\mu}) - \exp(\beta_\rho, \check{\mu})) \quad (36)$$

M. Feinberg called this kinetic law the ‘‘Marselin-De Donder’’ kinetics [37].

Under the detailed balance conditions, the auxiliary function $\theta(\lambda)$ is symmetric with respect to change $\lambda \mapsto (1 - \lambda)$. Therefore, $\theta(1) = \theta(0)$ and, because of convexity of $\theta(\lambda)$, the inequality holds: $\theta'(1) \geq 0$. Therefore, $\dot{H} \leq 0$ and kinetic equations obey the second law of thermodynamics.

The explicit formula for $\dot{H} \leq 0$ has the well known form since Boltzmann proved his H -theorem in 1872:

$$\frac{dH}{dt} = -V \sum_\rho (\ln r_\rho^+ - \ln r_\rho^-) (r_\rho^+ - r_\rho^-) \leq 0 \quad (37)$$

A convenient equivalent form of $\dot{H} \leq 0$ is proposed in [38]:

$$\frac{dH}{dt} = -V \sum_{\rho} (r_{\rho}^{+} + r_{\rho}^{-}) \mathbb{A}_{\rho} \tanh \frac{\mathbb{A}_{\rho}}{2} \leq 0 \quad (38)$$

where

$$\mathbb{A}_{\rho} = -(\gamma_{\rho}, \check{\mu}) (= -(\gamma_{\rho}, \mu)/RT, \text{ where } \mu \text{ is the chemical potential})$$

is a normalized *affinity*. In this formula, the kinetic information is collected in the nonnegative factors, the sums of reaction rates $(r_{\rho}^{+} + r_{\rho}^{-})$. The purely thermodynamic multiplier $\mathbb{A} \tanh(\mathbb{A}/2) \geq 0$ is positive for non-zero \mathbb{A} . For small $|\mathbb{A}|$, the expression $\mathbb{A} \tanh(\mathbb{A}/2)$ behaves like $\mathbb{A}^2/2$ and for large $|\mathbb{A}|$ it behaves like the absolute value, $|\mathbb{A}|$.

The detailed balance condition reflects “microreversibility”, that is, time-reversibility of the dynamic microscopic description and was first introduced by Boltzmann in 1872 as a consequence of the reversibility of collisions in Newtonian mechanics.

5.4. Complex Balance

The complex balance condition was invented by Boltzmann in 1887 for the Boltzmann equation [39] as an answer to the Lorentz objections [40] against Boltzmann’s proof of the H -theorem. Stueckelberg demonstrated in 1952 that this condition follows from the Markovian microkinetics of fast intermediates if their concentrations are small [41]. Under this asymptotic assumption this condition is just the probability balance condition for the underlying Markov process. (Stueckelberg considered this property as a consequence of “unitarity” in the S -matrix terminology.) It was known as the semidetailed or cyclic balance condition. This condition was rediscovered in the framework of chemical kinetics by Horn and Jackson in 1972 [42] and called the complex balance condition. Now it is used for chemical reaction networks in chemical engineering [43]. Detailed analysis of the backgrounds of the complex balance condition is given in [34].

Formally, the complex balance condition means that $\theta(1) \equiv \theta(0)$ for all values of $\check{\mu}$. We start from the initial stoichiometric equations (18) without joining the direct and reverse reactions. The equality $\theta(1) \equiv \theta(0)$ reads

$$\sum_{\rho} \varphi_{\rho} \exp(\check{\mu}, \alpha_{\rho}) = \sum_{\rho} \varphi_{\rho} \exp(\check{\mu}, \beta_{\rho}) \quad (39)$$

Let us consider the family of vectors $\{\alpha_{\rho}, \beta_{\rho}\}$ ($\rho = 1, \dots, m$). Usually, some of these vectors coincide. Assume that there are q different vectors among them. Let y_1, \dots, y_q be these vectors. For each $j = 1, \dots, q$ we take

$$R_j^{+} = \{\rho \mid \alpha_{\rho} = y_j\}, \quad R_j^{-} = \{\rho \mid \beta_{\rho} = y_j\} \quad (40)$$

We can rewrite Equation (39) in the form

$$\sum_{j=1}^q \exp(\check{\mu}, y_j) \left[\sum_{\rho \in R_j^{+}} \varphi_{\rho} - \sum_{\rho \in R_j^{-}} \varphi_{\rho} \right] = 0 \quad (41)$$

The Boltzmann factors $\exp(\tilde{\mu}, y_j)$ are linearly independent functions. Therefore, the natural way to meet these conditions is: for any $j = 1, \dots, q$

$$\sum_{\rho \in R_j^+} \varphi_\rho - \sum_{\rho \in R_j^-} \varphi_\rho = 0 \tag{42}$$

This is the general *complex balance condition*. This condition is sufficient for the inequality $\dot{H} = \theta'(1) \leq 0$, because it provides the equality $\theta(1) = \theta(0)$ and $\theta(\lambda)$ is a convex function.

It is easy to check that for the first order kinetics given by Equation (10) (or Equation (13)) with positive equilibrium, the complex balance condition is just the balance equation (11) and always holds.

5.5. Cyclic Decomposition of the Systems with Complex Balance

The complex balance conditions defined by Equation (42) allow a simple geometric interpretation. Let us introduce the digraph of transformation of complexes. The vertices of this digraph correspond to the formal sums (y, A) (“complexes”), where A is the vector of components, and $y \in \{y_1, \dots, y_q\}$ are vectors α_ρ or β_ρ from the stoichiometric equations of the elementary reactions (18). The edges of the digraph correspond to the elementary reactions with non-zero kinetic factor.

Let us assign to each edge $(\alpha_\rho, A) \rightarrow (\beta_\rho, A)$ the auxiliary *current*—the kinetic factor φ_ρ . For these currents, the complex balance condition presented by Equation (42) is just Kirchhoff’s first rule: the sum of the input currents is equal to the sum of the output currents for each vertex. (We have to stress that these auxiliary currents are not the actual rates of transformations.)

Let us use for the vertices the notation $\Theta_j: \Theta_j = (y_j, A)$, $(j = 1, \dots, q)$ and denote φ_{l_j} the fluxes for the edge $\Theta_j \rightarrow \Theta_l$.

The *simple cycle* is the digraph $\Theta_{i_1} \rightarrow \Theta_{i_2} \rightarrow \dots \rightarrow \Theta_{i_k} \rightarrow \Theta_{i_1}$, where all the complexes Θ_{i_l} ($l = 1, \dots, k$) are different. We say that the simple cycle is normalized if all the corresponding auxiliary fluxes are unit: $\varphi_{i_{j+1}i_j} = \varphi_{i_1i_k} = 1$.

The graph of the transformation of complexes cannot be arbitrary if the system satisfies the complex balance condition [22].

Proposition 3. *If the system satisfies the complex balance condition (i.e. Equation (42) holds) then every edge of the digraph of transformation of complexes is included into a simple cycle.*

Proof. First of all, let us formulate Kirchhoff’s first rule (42) for subsets: if the digraph of transformation of complexes satisfies Equation (42), then for any set of complexes Ω

$$\sum_{\Theta \in \Omega, \Phi \notin \Omega} \varphi_{\Theta\Phi} = \sum_{\Theta \in \Omega, \Phi \in \Omega} \varphi_{\Phi\Theta} \tag{43}$$

where $\varphi_{\Phi\Theta}$ is the positive kinetic factor for the reaction $\Theta \rightarrow \Phi$ if it belongs to the reaction mechanism (i.e., the edge $\Theta \rightarrow \Phi$ belongs to the digraph of transformations) and $\varphi_{\Phi\Theta} = 0$ if it does not. Equation (43) is just the result of summation of Equations (42) for all $(y_j, A) = \Theta \in \Omega$.

We say that a state Θ_j is reachable from a state Θ_k if $k = i$ or there exists a non-empty chain of transitions with non-zero coefficients that starts at Θ_k and ends at Θ_j : $\Theta_k \rightarrow \dots \rightarrow \Theta_j$. Let $\Theta_{i\downarrow}$ be the set of states reachable from Θ_i . The set $\Theta_{i\downarrow}$ has no output edges.

Assume that the edge $\Theta_j \rightarrow \Theta_i$ is not included in a simple cycle, which means $\Theta_j \notin \Theta_{i\downarrow}$. Therefore, the set $\Omega = \Theta_{i\downarrow}$ has the input edge $(\Theta_j \rightarrow \Theta_i)$ but no output edges and cannot satisfy Equation (43). This contradiction proves the proposition. \square

This property (every edge is included in a simple cycle) is equivalent to the so-called “weak reversibility” or to the property that every weakly connected component of the digraph is its strong component.

For every graph with the system of fluxes, which obey Kirchhoff’s first rule, the cycle decomposition theorem holds. It can be extracted from many books and papers [20,26,44]. Let us recall the notion of extreme ray. A ray with direction vector $x \neq 0$ is a set $\{\lambda x\}$ ($\lambda \geq 0$). A ray l is an extreme ray of a cone \mathbf{Q} if for any $u \in l$ and any $x, y \in \mathbf{Q}$, whenever $u = (x + y)/2$, we must have $x, y \in l$. If a closed convex cone does not include a whole straight line then it is the convex hull of its extreme rays [45].

Let us consider a digraph \mathcal{Q} with vertices Θ_i , the set of edges E and the system of auxiliary fluxes along the edges $\varphi_{ij} \geq 0$ ($(j, i) \in E$). The set of all nonnegative functions on E , $\varphi : (j, i) \mapsto \varphi_{ij}$, is a nonnegative orthant $\mathbb{R}_+^{|E|}$. Kirchhoff’s first rule (Equation (42)) together with nonnegativity of the kinetic factors define a cone of the systems with complex balance $\mathcal{Q} \subset \mathbb{R}_+^{|E|}$.

Proposition 4 (Cycle decomposition of systems with complex balance). *Every extreme ray of \mathcal{Q} has a direction vector that corresponds to a simple normalized cycle $\Theta_{i_1} \rightarrow \Theta_{i_2} \rightarrow \dots \rightarrow \Theta_{i_k} \rightarrow \Theta_{i_1}$, where all the complexes Θ_{i_l} ($l = 1, \dots, k$) are different, all the corresponding fluxes are unit, $\varphi_{i_{j+1}i_j} = \varphi_{i_1i_k} = 1$, and other fluxes are zeros.*

Proof. Let a function $\phi : E \rightarrow \mathbb{R}_+$ be an extreme ray of \mathcal{Q} and $\text{supp}\phi = \{(j, i) \in E \mid \phi_{ij} > 0\}$. Due to Proposition 3 each edge from $\text{supp}\phi$ is included in a simple cycle formed by edges from $\text{supp}\phi$. Let us take one this cycle $\Theta_{i_1} \rightarrow \Theta_{i_2} \rightarrow \dots \rightarrow \Theta_{i_k} \rightarrow \Theta_{i_1}$. Denote the fluxes of the corresponding simple normalized cycle by ψ . It is a function on E : $\psi_{i_{j+1}i_j} = \psi_{i_1i_k} = 1$ and $\psi_{ij} = 0$ if $(i, j) \in E$ but $(i, j) \neq (i_{j+1}, i_j)$ and $(i, j) \neq (i_1, i_k)$ ($i, j = 1, \dots, k, i \neq j$).

Assume that $\text{supp}\phi$ includes at least one edge that does not belong to the cycle $\Theta_{i_1} \rightarrow \Theta_{i_2} \rightarrow \dots \rightarrow \Theta_{i_k} \rightarrow \Theta_{i_1}$. Then, for sufficiently small $\kappa > 0$, $\phi \pm \kappa\psi \in \mathcal{Q}$ and the vector $\phi \pm \kappa\psi$ is not proportional to ϕ . This contradiction proves the proposition. \square

This decomposition theorem explains why the complex balance condition was often called the “cyclic balance condition”.

5.6. Local Equivalence of Systems with Detailed and Complex Balance

The class of systems with detailed balance is the proper subset of the class of systems with complex balance. A simple (irreversible) cycle of the length $k > 2$ gives a simplest and famous example of the complex balance system without detailed balance condition.

For Markov chains, the complex balance systems are all the systems that have a positive equilibrium distribution presented by Equation (11), whereas the systems with detailed balance form the proper subclass of the Markov chains, the so-called reversible chains.

In nonlinear kinetics, the systems with complex balance provide the natural generalization of the Markov processes. They deserve the term “nonlinear Markov processes”, though it is occupied by a much wider notion [46]. The systems with detailed balance form the proper subset of this class.

Nevertheless, in some special sense the classes of systems with detailed balance and with the complex balance are equivalent. Let us consider a thermodynamic state given by the vector of potentials $\check{\mu}$ defined by Equation (20). Let all the reactions in the reaction mechanism be reversible (*i.e.*, for every transition $\Theta_i \rightarrow \Theta_j$ the reverse transition $\Theta_i \leftarrow \Theta_j$ is allowed and the corresponding edge belongs to the digraph of complex transformations). Calculate the right hand side of the kinetic equations (34) with the detailed balance condition given by Equation (35) for a given value of $\check{\mu}$ and all possible values of $\varphi_\rho^+ = \varphi_\rho^-$. The set of these values of \dot{N} is a convex cone. Denote this cone $\mathbf{Q}_{DB}(\check{\mu})$. For the same transition graph, calculate the right hand side of the kinetic equation (22) under the complex balance condition (42). The set of these values of \dot{N} is also a convex cone. Denote it $\mathbf{Q}_{CB}(\check{\mu})$. It is obvious that $\mathbf{Q}_{DB}(\check{\mu}) \subseteq \mathbf{Q}_{CB}(\check{\mu})$. Surprisingly, these cones coincide. In [26] we proved this fact on the basis of the Michaelis-Menten-Stueckelberg theorem [34] about connection of the macroscopic GMAL kinetics and the complex balance condition with the Markov microscopic description and under some asymptotic assumptions. Below a direct proof is presented.

Theorem 2 (Local equivalence of detailed and complex balance).

$$\mathbf{Q}_{DB}(\check{\mu}) = \mathbf{Q}_{CB}(\check{\mu}) \tag{44}$$

Proof. Because of the cycle decomposition (Proposition 4) it is sufficient to prove this theorem for simple normalized cycles. Let us use induction on the cycle length k . For $k = 2$ the transition graph is $\Theta_1 \rightleftharpoons \Theta_2$ and the detailed balance condition (35) coincides with the complex balance condition (42). Assume that for the cycles of the length below k the theorem is proved. Consider a normalized simple cycle $\Theta_1 \rightarrow \Theta_2 \rightarrow \dots \Theta_k \rightarrow \Theta_1$, $\Theta_i = (y_i, A)$. The corresponding kinetic equations are

$$\begin{aligned} \frac{dN}{dt} = & (y_2 - y_1) \exp(\check{\mu}, y_1) + (y_3 - y_2) \exp(\check{\mu}, y_2) + \dots \\ & + (y_{k-1} - y_k) \exp(\check{\mu}, y_{k-1}) + (y_k - y_1) \exp(\check{\mu}, y_k) \end{aligned} \tag{45}$$

At the equilibrium, all systems with detailed balance or with complex balance give $\dot{N} = 0$. Assume that the state $\check{\mu}$ is non-equilibrium and therefore not all the Boltzmann factors $\exp(\check{\mu}, y_i)$ are equal. Select i such that the Boltzmann factor $\exp(\check{\mu}, y_i)$ has minimal value, while for the next position in the cycle this factor becomes bigger. We can use a cyclic permutation and assume that the factor $\exp(\check{\mu}, y_1)$ is the minimal one and $\exp(\check{\mu}, y_2) > \exp(\check{\mu}, y_1)$.

Let us find a kinetic factor φ such that the reaction system consisting of two cycles, a cycle of the length 2 with detailed balance $\Theta_1 \xrightleftharpoons[\varphi]{\varphi} \Theta_2$ (here the kinetic factors are shown above and below the arrows) and a simple normalized cycle of the length $k - 1$, $\Theta_2 \rightarrow \dots \Theta_k \rightarrow \Theta_2$, gives the same \dot{N} at the state $\check{\mu}$ as the initial scheme. We obtain from Equation (45) the following necessary and sufficient condition

$$(y_1 - y_k) \exp(\check{\mu}, y_k) + (y_2 - y_1) \exp(\check{\mu}, y_1) = (y_2 - y_k) \exp(\check{\mu}, y_k) + \varphi(y_2 - y_1)(\exp(\check{\mu}, y_1) - \exp(\check{\mu}, y_2))$$

. It is sufficient to equate here the coefficients at every y_i ($i = 1, 2, k$). The result is

$$\varphi = \frac{\exp(\check{\mu}, y_k) - \exp(\check{\mu}, y_1)}{\exp(\check{\mu}, y_2) - \exp(\check{\mu}, y_1)}$$

By the induction assumption we proved that theorem for the cycles of arbitrary length and, therefore, it is valid for all reaction schemes with complex balance. \square

The cone $\mathbf{Q}_{\text{DB}}(\check{\mu})$ of the possible values of \dot{N} in Equation (34) is a polyhedral cone with finite set of extreme rays at any non-equilibrium state $\check{\mu}$ for the systems with detailed balance. Each of its extreme rays has the direction vector of the form

$$\gamma_{\rho} \text{sign}(\exp(\check{\mu}, \alpha_{\rho}) - \exp(\check{\mu}, \beta_{\rho})) \quad (46)$$

This follows from the form of the reaction rate presented by Equation (36) for the kinetic equations (34). Following Theorem 2, the cone of the possible values of \dot{N} for systems with complex balance has the same set of extreme rays. Each extreme ray corresponds to a single reversible elementary reaction with the detailed balance condition (35).

5.7. General H -Theorem for GMAL

Consider GMAL kinetics with the given reaction mechanism presented by stoichiometric equations (32) and the detailed balance condition (35). The reaction rates of the elementary reaction for the kinetic equations (34) are proportional to the nonnegative parameter φ_{ρ} in Equation (36). These m nonnegative numbers φ_{ρ} ($\rho = 1, \dots, m$) are independent in the following sense: for any set of values $\varphi_{\rho} \geq 0$ the kinetic equations (34) satisfy the H -theorem in the form of Equation (38):

$$\frac{dH}{dt} = -V \sum_{\rho} \varphi_{\rho} (\exp(\alpha_{\rho}, \check{\mu}) + \exp(\beta_{\rho}, \check{\mu})) \mathbb{A}_{\rho} \tanh \frac{\mathbb{A}_{\rho}}{2} \leq 0 \quad (47)$$

Therefore, nonnegativity is the only a priori restriction on the values of φ_{ρ} ($\rho = 1, \dots, m$).

One Lyapunov function for the GMAL kinetics with the given reaction mechanism and the detailed balance condition obviously exists. This is the thermodynamic Lyapunov function H used in GMAL construction. For ideal systems (in particular, for master equation) H has the standard form $\sum_i N_i (\ln(c_i/c_i^*) - 1)$ given by Equation (23). Usually, H is assumed to be convex and some singularities (like $c \ln c$) near zeros of c may be required for positivity preservation in kinetics ($\dot{N}_i \geq 0$ if $c_i = 0$). The choice of the thermodynamic Lyapunov function for GMAL construction is wide. We consider kinetic equations in a compact convex set U and assume H to be convex and continuous in U and differentiable in the relative interior of U with derivatives continued by continuity to U .

Assume that we select the thermodynamic Lyapunov function H and the reaction mechanism in the form (32). Are there other universal Lyapunov functions for GMAL kinetics with detailed balance and given mechanism? “Universal” here means “independent of the choice of the nonnegative kinetic factors”.

For a given reaction mechanism we introduce the partial equilibria criterion by analogy to Definition 1. Roughly speaking, a convex function F satisfies this criterion if its conditional minima correctly describe the partial equilibria of elementary reactions.

For each elementary reaction $\sum_i \alpha_{\rho i} A_i \rightleftharpoons \sum_i \beta_{\rho i} A_i$ from the reaction mechanism given by the stoichiometric equations (32) and any $X \in U$ we define an interval of a straight line

$$I_{X,\rho} = \{X + \lambda \gamma_{\rho} \mid \lambda \in \mathbb{R}\} \cap U. \quad (48)$$

Definition 2 (Partial equilibria criterion for GMAL). A convex function $F(N)$ on U satisfies the partial equilibria criterion with a given thermodynamic Lyapunov function H and reversible reaction mechanism given by stoichiometric equations (32) if

$$\operatorname{argmin}_{N \in I_{X,\rho}} H(N) \subseteq \operatorname{argmin}_{N \in I_{X,\rho}} F(N) \quad (49)$$

for all $X \in U$, $\rho = 1, \dots, m$.

Theorem 3. A convex function $F(N)$ on U is a Lyapunov function for all kinetic equations (34) with the given thermodynamic Lyapunov function H and reaction rates presented by Equation (36) (detailed balance) if and only if it satisfies the partial equilibria criterion (Definition 2).

Proof. The partial equilibria criterion is necessary because $F(N)$ should be a Lyapunov function for a reaction mechanism that consists of any single reversible reaction from the reaction mechanism (32). It is also sufficient because for the whole reaction mechanism the kinetic equations (34) are the conic combinations of the kinetic equations for single reversible reactions from the reaction mechanism (32). \square

For the general reaction systems with complex balance we can use the theorem about local equivalence (Theorem 2). Consider a GMAL reaction system with the mechanism (18) and the complex balance condition.

Theorem 4. A convex function $F(N)$ on U is a Lyapunov function for all kinetic equations (22) with the given thermodynamic Lyapunov function H and the complex balance condition (42) if it satisfies the partial equilibria criterion (Definition 2).

Proof. The theorem follows immediately from Theorem 3 about Lyapunov functions for systems with detailed balance and the theorem about local equivalence between systems with local and complex balance (Theorem 2). \square

The general H -theorems for GMAL is similar to Theorem 1 for Markov chains. Nevertheless, many non-classical universal Lyapunov functions are known for master equations, for example, the f -divergences given by Equation (1), while for a nonlinear reaction mechanism it is difficult to present a single example different from the thermodynamic Lyapunov function or its monotonic transformations. The following family of example generalizes Equation (17).

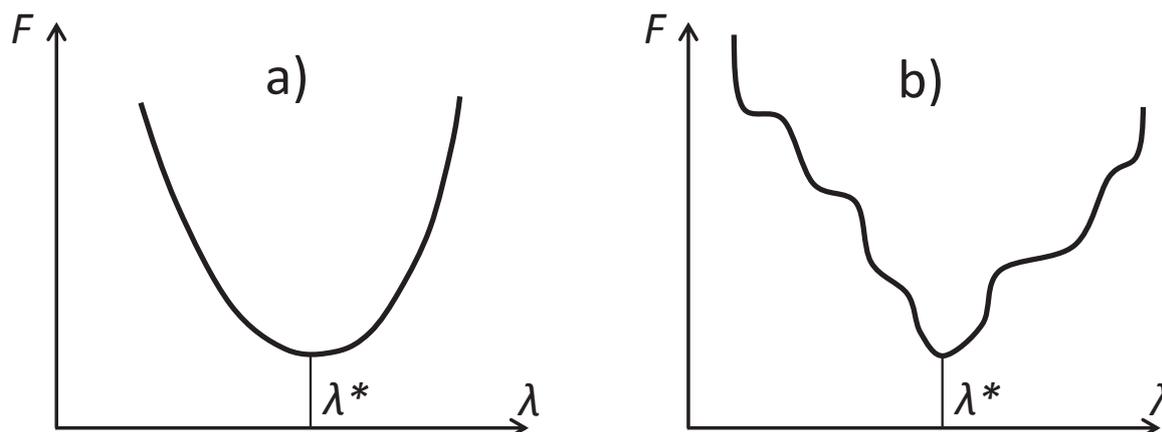
$$F(N) = H(N) + \varepsilon f(N) \prod_{\rho} (\exp(\alpha_{\rho}, \check{\mu}) - \exp(\beta_{\rho}, \check{\mu}))^2 \quad (50)$$

where $f(N)$ is a non-negative differentiable function and $\varepsilon > 0$ is a sufficiently small number. This function satisfies the conditional equilibria criterion. For continuous $H(N)$ on compact U with the spectrum of the Hessian uniformly separated from zero, this $F(N)$ is convex for sufficiently small $\varepsilon > 0$.

6. Generalization: Weakened Convexity Condition, Directional Convexity and Quasiconvexity

In all versions of the general H -theorems we use convexity of the Lyapunov functions. Strong convexity of the thermodynamic Lyapunov functions H (or even positive definiteness of its Hessian) is needed, indeed, to provide reversibility of the Legendre transform $N \leftrightarrow \nabla H$.

Figure 3. Monotonicity on both sides of the minimizer λ^* for convex (a) and non-convex but quasiconvex (b) functions



For the kinetic Lyapunov functions that satisfy the partial equilibria criterion, we use, actually, a rather weak consequence of convexity in restrictions on the straight lines $X + \lambda\gamma$ (where $\lambda \in \mathbb{R}$ is a coordinate on the real line, γ is a stoichiometric vector of an elementary reaction): if $\lambda^* = \underset{\lambda \in \mathbb{R}}{\operatorname{argmin}} F(X + \lambda\gamma)$ then on the half-lines (rays) $\lambda \geq \lambda^*$ and $\lambda \leq \lambda^*$ function $F(X + \lambda\gamma)$ is monotonic. It does not decrease for $\lambda \geq \lambda^*$ and does not increase for $\lambda \leq \lambda^*$. Of course, convexity is sufficient (Figure 3a) but a much weaker property is needed (Figure 3b).

A function F on a convex set U is *quasiconvex* [47] if all its sublevel sets are convex. It means that for every $X, Y \in U$

$$F(\lambda X + (1 - \lambda)Y) \leq \max\{F(X), F(Y)\} \text{ for all } \lambda \in [0, 1] \quad (51)$$

In particular, a function F on a segment is quasiconvex if all its sublevel sets are segments.

Among many other types of convexity and quasiconvexity (see, for example [48]) two are important for the general H -theorem. We do not need convexity of functions along all straight lines in U . It is sufficient that the function is convex on the straight lines $X + \mathbb{R}\gamma_\rho$, where γ_ρ are the stoichiometric (direction) vectors of the elementary reactions.

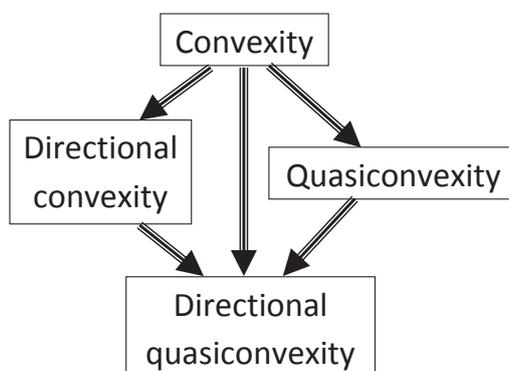
Let D be a set of vectors. A function F is D -convex if its restriction to each line parallel to a nonzero $v \in D$ is convex [49]. In our case, D is the set of stoichiometric vectors of the transitions, $D = \{\gamma_\rho \mid \rho = 1, \dots, m\}$. We can use this *directional convexity* instead of convexity in Propositions 1, 2 and Theorems 1, 3, 4.

Finally, we can relax the convexity conditions even more and postulate *directional quasiconvexity* [50] for the set of directions $D = \{\gamma_\rho \mid \rho = 1, \dots, m\}$. Propositions 1, 2 and Theorems 1, 3, 4 will be still true

if the functions are continuous, quasiconvex in restrictions on all lines $X + \mathbb{R}\gamma_\rho$ and satisfy the partial equilibria criterion.

Relations between these types of convexity are schematically illustrated in Figure 4.

Figure 4. Relations between different types of convexity



7. Discussion

Many non-classical entropies are invented and applied to various problems in physics and data analysis. In this paper, the general necessary and sufficient criterion for the existence of H -theorem is proved. It has a simple and physically transparent form: the convex divergence (relative entropy) should properly describe the partial equilibria for transitions $A_i \rightleftharpoons A_j$. It is straightforward to check this partial equilibria criterion. The applicability of this criterion does not depend on the detailed balance condition and it is valid both for the class of the systems with detailed balance and for the general first order kinetics without this assumption.

If an entropy has no H -theorem (that is, it violates the second law and the data processing lemma) then there should be unprecedentedly strong reasons for its use. Without such strong reasons we cannot employ it. Now, I cannot find an example of sufficiently strong reasons but people use these entropies in data analysis and we have to presume that they may have some hidden reasons and that these reasons may be sufficiently strong. We demonstrate that this problem arises even for such popular divergences like Euclidean distance or Itakura-Saito distance.

The general H -theorem is simply a reduction of a dynamical question (Lyapunov functionals) to a static one (partial equilibria). It is not surprising that it can be also proved for nonlinear Generalized Mass Action Law kinetics. Here kinetic systems with complex balance play the role of the general Markov chains, whereas the systems with detailed balance correspond to the reversible Markov chains. The requirement of convexity of Lyapunov functions can be relaxed to the directional convexity (in the directions of reactions) or even directional quasiconvexity.

For the reversible Markov chains presented by Equations (15) with the classical entropy production Formula (16), every universal Lyapunov function H should satisfy inequalities

$$\left(\frac{p_j}{p_j^*} - \frac{p_i}{p_i^*} \right) \left(\frac{\partial H(P)}{\partial p_j} - \frac{\partial H(P)}{\partial p_i} \right) \leq 0 \text{ for all } i, j, i \neq j \quad (52)$$

These inequalities are closely related to another generalization of convexity, the Schur convexity [51]. They turn into the definition of the Schur convexity when equilibrium is the equidistribution with

$p_i^* = 1/n$ for all i . Universal Lyapunov functions for nonlinear kinetics give one more generalization of the Schur convexity.

Introduction of many non-classical entropies leads to the “uncertainty of uncertainty” phenomenon: we measure uncertainty by entropy but we have uncertainty in the entropy choice [27]. The selection of the appropriate entropy and introduction of new entropies are essentially connected with the class of kinetics. H -theorems in physics are formalizations of the second law of thermodynamics: entropy of isolated systems should increase in relaxation to equilibrium. If we know the class of kinetic equations (for example, the Markov kinetics given by master equations) then the H theorem states that it is possible to use this entropy with the given kinetics. If we know the entropy and are looking for kinetic equations then such a statement turns into the thermodynamic restriction on the thermodynamically admissible kinetic equations. For information processing, the class of kinetic equations describes possible manipulations with data. In this case, the H -theorems mean that under given class of manipulation the information does not increase.

It is not possible to compare different entropies without any relation to kinetics. It is useful to specify the class of kinetic equations, for which they are the Lyapunov functionals. For the GMAL equations, we can introduce the dynamic equivalence between divergences (free entropies or conditional entropies). Two functionals $H(N)$ and $F(N)$ in a convex set U are *dynamically consistent* with respect to the set of stoichiometric vectors $\{\gamma_\rho\}$ ($\rho = 1, \dots, m$) if

(1) F and H are directionally quasiconvex functions in directions $\{\gamma_\rho\}$ ($\rho = 1, \dots, m$)

(2) For all $\rho = 1, \dots, m$ and $N \in U$

$$(\nabla_N F(N), \gamma_\rho)(\nabla_N H(N), \gamma_\rho) \geq 0$$

For the Markov kinetics, the partial equilibria criterion is sufficient for a convex function $H(P)$ to be dynamically consistent with the relative entropy $\sum_i p_i (\ln(p_i/p_i^*) - 1)$ in the unit simplex Δ_n . For GMAL, any convex function $H(N)$ defines a class of kinetic equations. Every reaction mechanism defines a family of kinetic equations from this class and a class of Lyapunov functions F , which are dynamically consistent with H . The main message of this paper is that it is necessary to discuss the choice of the non-classical entropies in the context of kinetic equations.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Rényi, A. On measures of entropy and information. In Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability, Berkeley, CA, USA, 1960; University of California Press: Berkeley, CA, USA, 1961; Volume 1, pp. 547–561.
2. Csiszár, I. Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. *Magyar. Tud. Akad. Mat. Kutato Int. Kozl.* **1963**, *8*, 85–108. (in German)

3. Morimoto, T. Markov processes and the H -theorem. *J. Phys. Soc. Jap.* **1963**, *12*, 328–331.
4. Callen, H.B. *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed.; Wiley: New York, NY, USA, 1985.
5. Burg, J.P. The relationship between maximum entropy spectra and maximum likelihood spectra. *Geophysics* **1972**, *37*, 375–376.
6. Cressie, N.; Read, T. Multinomial Goodness of Fit Tests. *J. R. Stat. Soc. Ser. B* **1984**, *46*, 440–464.
7. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487.
8. Abe, S., Okamoto, Y., Eds. *Nonextensive Statistical Mechanics and its Applications*; Springer: Heidelberg, Germany, 2001.
9. Cichocki, A.; Amari, S.-I. Families of alpha- beta- and gamma- divergences: Flexible and robust measures of similarities. *Entropy* **2010**, *12*, 1532–1568.
10. Esteban, M.D.; Morales, D. A summary of entropy statistics. *Kybernetika* **1995**, *31*, 337–346.
11. Gorban, A.N.; Gorban, P.A.; Judge, G. Entropy: The Markov ordering approach. *Entropy*, **2010** *12*, 1145–1193. arXiv:1003.1377 [physics.data-an].
12. Bregman, L.M. The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. *USSR Comput. Math. Math. Phys.* **1967**, *7*, 200–217.
13. Banerjee, A.; Merugu, S.; Dhillon, I.S.; Ghosh, J. Clustering with Bregman divergences. *J. Mach. Learn. Res.* **2005**, *6*, 1705–1749.
14. Csiszár, I.; Matúš, F. Generalized minimizers of convex integral functionals, Bregman distance, Pythagorean identities. **2012**, arXiv:1202.0666 [math.OC].
15. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–656.
16. Cohen, J.E.; Derriennic, Y.; Zbaganu, G.H. Majorization, monotonicity of relative entropy and stochastic matrices. *Contemp. Math.* **1993**, *149*, 251–259.
17. Cohen, J.E.; Iwasa, Y.; Rautu, G.; Ruskai, M.B.; Seneta, E.; Zbaganu, G. Relative entropy under mappings by stochastic matrices. *Linear Algebra Appl.* **1993**, *179*, 211–235.
18. Gorban, P.A. Monotonically equivalent entropies and solution of additivity equation. *Physica A* **2003**, *328*, 380–390.
19. Amari, S.-I. Divergence, Optimization, Geometry. In Proceedings of the 16th International Conference on Neural Information Processing, Bangkok, Thailand, 1–5 December 2009; Leung, C.S., Lee, M., Chan, J.H., Eds.; Springer: Berlin, Germany, 2009; pp. 185–193.
20. Meyn, S.R. *Control Techniques for Complex Networks*; Cambridge University Press: Cambridge, UK, 2007.
21. Meyn, S.R.; Tweedie, R.L. *Markov Chains and Stochastic Stability*; Cambridge University Press: Cambridge, UK, 2009.
22. Feinberg, M.; Horn, F.J. Dynamics of open chemical systems and the algebraic structure of the underlying reaction network. *Chem. Eng. Sci.* **1974**, *29*, 775–787.
23. Szederkényi, G.; Hangos, K.M.; Tuza, Z. Finding weakly reversible realizations of chemical reaction networks using optimization. *Computer* **2012**, *67*, 193–212.

24. Gorban, A.N. Kinetic path summation, multi-sheeted extension of master equation, and evaluation of ergodicity coefficient. *Physica A* **2011**, *390*, 1009–1025.
25. Van Kampen, N.G. *Stochastic processes in physics and chemistry*; North-Holland: Amsterdam, The Netherlands, 1981.
26. Gorban, A.N. Local equivalence of reversible and general Markov kinetics. *Physica A* **2013**, *392*, 1111–1121.
27. Gorban, A.N. *Maxallent: Maximizers of all entropies and uncertainty of uncertainty*. *Comput. Math. Appl.* **2013**, *65*, 1438–1456.
28. Golub, G.H.; van Loan, C.F. *Matrix Computations*; Johns Hopkins University Press: Baltimore, MD, USA, 1996.
29. Gorban, A.N. *Equilibrium encircling. Equations of Chemical Kinetics and Their Thermodynamic Analysis*; Nauka: Moscow, Russia, 1984. (in Russian)
30. Yablonskii, G.S.; Bykov, V.I.; Gorban, A.N.; Elokhin, V.I. *Kinetic Models of Catalytic Reactions*; Series Comprehensive Chemical Kinetics, Volume 32; Elsevier: Amsterdam, The Netherlands, 1991.
31. Grmela, M.; Öttinger, H.C. Dynamics and thermodynamics of complex fluids. I. Development of a general formalism. *Phys. Rev. E* **1997**, *56*, doi:10.1103/PhysRevE.56.
32. Grmela, M. Fluctuations in extended mass-action-law dynamics. *Physica D* **2012**, *241*, 976–986.
33. Giovangigli, V.; Matuszewski, L. Supercritical fluid thermodynamics from equations of state. *Physica D* **2012**, *241*, 649–670.
34. Gorban, A.N.; Shahzad, M. The Michaelis-Menten-Stueckelberg Theorem. *Entropy* **2011**, *13*, 966–1019.
35. Hangos, K.M. Engineering model reduction and entropy-based Lyapunov functions in chemical reaction kinetics. *Entropy* **2010**, *12*, 772–797.
36. Orlov, N.N.; Rozonoer, L.I. The macrodynamics of open systems and the variational principle of the local potential. *J. Franklin Inst-Eng. Appl. Math.* **1984**, *318*, 283–341.
37. Feinberg, M. On chemical kinetics of a certain class. *Arch. Rat. Mech. Anal.* **1972**, *46*, 1–41.
38. Gorban, A.N.; Mirkes, E.M.; Yablonsky, G.S. Thermodynamics in the limit of irreversible reactions. *Physica A* **2013**, *392*, 1318–1335.
39. Boltzmann, L. Neuer Beweis zweier Sätze über das Wärmegleichgewicht unter mehratomigen Gasmoleküle. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien* **1887**, *95*, 153–164. (in German)
40. Lorentz, H.-A. Über das Gleichgewicht der lebendigen Kraft unter Gasmolekülen. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien* **1887**, *95*, 115–152. (in German)
41. Stueckelberg, E.C.G. Theoreme *H* et unitarite de *S*. *Helv. Phys. Acta* **1952**, *25*, 577–580.
42. Horn, F.; Jackson, R. General mass action kinetics. *Arch. Ration. Mech. Anal.* **1972**, *47*, 81–116.
43. Szederkényi, G.; Hangos, K.M. Finding complex balanced and detailed balanced realizations of chemical reaction networks. *J. Math. Chem.* **2011**, *49*, 1163–1179.
44. Kalpazidou, S.L. *Cycle Representations of Markov Processes*; Book Series: Applications of Mathematics, Volume 28; Springer: New York, NY, USA, 2006.
45. Rockafellar, R.T. *Convex Analysis*; Princeton University Press: Princeton, NJ, USA, 1997.

46. Kolokoltsov, V.N. *Nonlinear Markov Processes and Kinetic Equations*; Book series Cambridge Tracts in Mathematics, Volume 182; Cambridge University Press: Cambridge, UK, 2010.
47. Greenberg, H.J.; Pierskalla, W.P. A review of quasi-convex functions. *Oper. Res.* **1971**, *19*, 1553–1570.
48. Ponstein, J. Seven kinds of convexity. *SIAM Rev.* **1967**, *9*, 115–119.
49. Matoušek, J. On directional convexity. *Discret. Comput. Geom.* **2001**, *25*, 389–403.
50. Hwang, F.K.; Rothblum, U.G. Directional-quasi-convexity, asymmetric Schur-convexity and optimality of consecutive partitions. *Math. Oper. Res.* **1996**, *21*, 540–554.
51. Marshall, A.W.; Olkin, I.; Arnold, B.C. *Inequalities: Theory of Majorization and its Applications*; Springer: New York, NY, USA, 2011.

© 2014 by the author; 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/3.0/>).