Relative Entropy in Biological Systems

John C. Baez [1,2,*] **and Blake S. Pollard** [3,*]

1   Department of Mathematics, University of California, Riverside, CA 92521, USA
2   Centre for Quantum Technologies, National University of Singapore, Singapore 117543, Singapore
3   Department of Physics and Astronomy, University of California, Riverside, CA 92521, USA
*   Correspondence: baez@math.ucr.edu (J.C.B.); bpoll002@ucr.edu (B.S.P.)

Abstract: In this paper we review various information-theoretic characterizations of the approach to equilibrium in biological systems. The replicator equation, evolutionary game theory, Markov processes and chemical reaction networks all describe the dynamics of a population or probability distribution. Under suitable assumptions, the distribution will approach an equilibrium with the passage of time. Relative entropy—that is, the Kullback–Leibler divergence, or various generalizations of this—provides a quantitative measure of how far from equilibrium the system is. We explain various theorems that give conditions under which relative entropy is nonincreasing. In biochemical applications these results can be seen as versions of the Second Law of Thermodynamics, stating that free energy can never increase with the passage of time. In ecological applications, they make precise the notion that a population gains information from its environment as it approaches equilibrium.

---

1. Introduction

Life relies on nonequilibrium thermodynamics, since in thermal equilibrium there are no flows of free energy. Biological systems are also open systems, in the sense that both matter and energy flow in and out of them. Nonetheless, it is important in biology that systems can sometimes be treated as approximately closed, and sometimes approach equilibrium before being disrupted in one way or another. This can occur on a wide range of scales, from large ecosystems to within a single cell or organelle. Examples include:

- a population approaching an evolutionarily stable state;
- random processes such as mutation, genetic drift, the diffusion of organisms in an environment or the diffusion of molecules in a liquid;
- a chemical reaction approaching equilibrium.

A common feature of these processes is that as they occur, quantities mathematically akin to entropy tend to increase. Closely related quantities such as free energy tend to decrease. In this review we explain some mathematical results that explain why this occurs.

Most of these results involve a quantity that is variously known as "relative information", "relative entropy", "information gain" or the "Kullback–Leibler divergence". We shall use the first term.

Given two probability distributions p and q on a finite set X, their **relative information**, or more precisely the **information of p relative to q**, is

$$I(p,q) = \sum_{i \in X} p_i \ln\left(\frac{p_i}{q_i}\right). \tag{1}$$

We use the word "information" instead of "entropy" because one expects entropy to increase with time, and the theorems we present will say that $I(p,q)$ decreases with time under various conditions. The reason is that the Shannon entropy

$$S(p) = -\sum_{i \in X} p_i \ln p_i \tag{2}$$

contains a minus sign that is missing from the definition of relative information.

Intuitively, $I(p,q)$ is the amount of information gained when we start with a hypothesis given by some probability distribution q and then change our hypothesis, perhaps on the basis of some evidence, to some other distribution p. For example, if we start with the hypothesis that a coin is fair and then are told that it landed heads up, the relative information is $\ln 2$, so we have gained 1 bit of information. If however we started with the hypothesis that the coin always lands heads up, we would have gained no information.

Mathematically, relative information is a **divergence**: it obeys

$$I(p,q) \geq 0$$

and

$$I(p,q) = 0 \iff p = q,$$

but not necessarily the other axioms for a distance function, symmetry and the triangle inequality, which indeed fail for relative information. There are many other divergences besides relative information [1,2]. However, relative information can be singled out by a number of characterizations [3], including one based on ideas from Bayesian inference [4]. The relative information is also close to the expected number of extra bits required to code messages distributed according to the probability measure p using a code optimized for messages distributed according to q (Theorem 5.4.3 in [5]).

In this review we describe various ways in which a population or probability distribution evolves continuously according to some differential equation. For all these differential equations, we describe conditions under which relative information decreases. Briefly, the results are as follows. We hasten to reassure the reader that our paper explains all the terminology involved, and the proofs of the claims are given in full:

- In Section 2 we consider a very general form of the Lotka–Volterra equations, which are a commonly used model of population dynamics. Starting from the population P_i of each type of replicating entity, we can define a probability distribution

$$p_i = \frac{P_i}{\sum_{i \in X} P_i},$$

  which evolves according to a nonlinear equation called the replicator equation. We describe a necessary and sufficient condition under which $I(q, p(t))$ is nonincreasing when $p(t)$ evolves according to the replicator equation while q is held fixed.
- In Section 3 we consider a special case of the replicator equation that is widely studied in evolutionary game theory. In this case we can think of probability distributions as mixed strategies in a two-player game. When q is a dominant strategy, $I(q, p(t))$ can never increase when $p(t)$ evolves according to the replicator equation. We can think of $I(q, p(t))$ as the information that the

population has left to learn. Thus, evolution is analogous to a learning process—an analogy that in the field of artificial intelligence is exploited by evolutionary algorithms.

- In Section 4 we consider continuous-time, finite-state Markov processes. Here we have probability distributions on a finite set X evolving according to a linear equation called the master equation. In this case $I(p(t), q(t))$ can never increase. Thus, if q is a steady state solution of the master equation, both $I(p(t), q)$ and $I(q, p(t))$ are nonincreasing. We can always write q as the Boltzmann distribution for some energy function $E \colon X \to \mathbb{R}$, meaning that

$$q_i = \frac{\exp(-E_i/kT)}{\sum\limits_{j \in X} \exp(-E_j/kT)},$$

where T is temperature and k is Boltzmann's constant. In this case, $I(p(t), q)$ is proportional to a difference of free energies:

$$I(p(t), q) = \frac{F(p) - F(q)}{kT}.$$

Thus, the nonincreasing nature of $I(p(t), q)$ is a version of the Second Law of Thermodynamics. In a companion paper [6], we examine how this result generalizes to non-equilibrium steady states of "open Markov processes", in which probability can flow in or out of the set X.

- Finally, in Section 5 we consider chemical reactions and other processes described by reaction networks. In this context we have nonnegative real populations P_i of entities of various kinds $i \in X$, and these population distributions evolve according to a nonlinear equation called the rate equation. We can generalize relative information from probability distributions to populations by setting

$$I(P, Q) = \sum_{i \in X} P_i \ln\left(\frac{P_i}{Q_i}\right) - (P_i - Q_i).$$

The extra terms cancel when P and Q are both probability distributions, but they ensure that $I(P, Q) \geq 0$ for arbitrary populations. If Q is a special sort of steady state solution of the rate equation, called a complex balanced equilibrium, $I(P(t), Q)$ can never increase when $P(t)$ evolves according to the rate equation.

2. The Replicator Equation

The replicator equation is a simplified model of how populations change with time. Suppose we have n different types of self-replicating entity. We will call these entities **replicators**. We will call the types of replicators **species**, but they do not need to be species in the biological sense. For example, the replicators could be genes, and the types could be alleles. Or the replicators could be restaurants, and the types could be restaurant chains.

Let $P_i(t)$, or just P_i for short, be the population of the i-th species at time t. Then a very general form of the **Lotka–Volterra equations** says that

$$\frac{dP_i}{dt} = f_i(P_1, \dots, P_n) \, P_i. \tag{3}$$

Thus the population P_i changes at a rate proportional to P_i, but the "constant of proportionality" need not be constant: it can be any smooth function f_i of the populations of all the species. We call f_i the **fitness function** of the i-th species. We can create a vector whose components are all the populations:

$$P = (P_1, \dots, P_n).$$

This lets us write the Lotka–Volterra equations more concisely as

$$\dot{P}_i = f_i(P)P_i,$$

where the dot stands for a time derivative.

Instead of considering the population P_i of the i-th species, one often considers the probability p_i that a randomly chosen replicator will belong to the i-th species. More precisely, this is the fraction of replicators belonging to that species:

$$p_i = \frac{P_i}{\sum_j P_j}.$$

As a mnemonic, remember that the **P**opulation P_i is being normalized to give a **p**robability p_i.

How do these probabilities p_i change with time? The quotient rule gives

$$\dot{p}_i = \frac{\dot{P}_i \left(\sum_j P_j\right) - P_i \left(\sum_j \dot{P}_j\right)}{\left(\sum_j P_j\right)^2},$$

so the replicator equation gives

$$\dot{p}_i = \frac{f_i(P)P_i \left(\sum_j P_j\right) - P_i \left(\sum_j f_j(P)P_j\right)}{\left(\sum_j P_j\right)^2}.$$

Using the definition of p_i, this simplifies to:

$$\dot{p}_i = f_i(P)p_i - \left(\sum_j f_j(P)p_j\right) p_i.$$

The expression in parentheses here has a nice meaning: it is the **mean fitness**. In other words, it is the average, or expected, fitness of a replicator chosen at random from the whole population. Let us write it thus:

$$\langle f(P) \rangle = \sum_j f_j(P)p_j.$$

This gives the **replicator equation** in its classic form:

$$\dot{p}_i = \left(f_i(P) - \langle f(P) \rangle \right) p_i, \tag{4}$$

where the dot stands for a time derivative. Thus, for the fraction of replicators of the i-th species to increase, their fitness must exceed the mean fitness.

So far, all this is classic material from population dynamics. At this point, Marc Harper considers what information theory has to say [7,8]. For example, consider the relative information $I(q, p)$ where q is some fixed probability distribution. How does this change with time? First, recall that

$$I(q, p) = \sum_i q_i \ln \left(\frac{q_i}{p_i} \right),$$

and we are assuming only p_i depends on time, not q_i, so

$$\frac{d}{dt} I(q, p(t)) = -\sum_i \frac{\dot{p}_i}{p_i} q_i.$$

By the replicator equation we obtain

$$\frac{d}{dt} I(q, p(t)) = -\sum_i \left(f_i(P) - \langle f(P) \rangle \right) q_i.$$

This is nice, but we can massage this expression to get something more enlightening. Remember, the numbers q_i sum to one. So:

$$\frac{d}{dt}I(q,p(t)) \;=\; \langle f(P)\rangle - \sum_i f_i(P)q_i$$

$$=\; \sum_i f_i(P)(p_i - q_i).$$

This result looks even better if we treat the numbers $f_i(P)$ as the components of a vector $f(P)$, and similarly for the numbers p_i and q_i. Then we can use the dot product of vectors to write

$$\frac{d}{dt}I(q,p(t)) = f(P(t)) \cdot (p(t) - q), \tag{5}$$

whenever p evolves according to the replicator equation while q is fixed. It follows that the relative information $I(q,p(t))$ will be nonincreasing if and only if

$$f(P(t)) \cdot (p(t) - q) \leq 0.$$

This nice result can be found in Marc Harper's 2009 paper relating the replicator equation to Bayesian inference (Theorem 1 in [8]). He traces its origins to much earlier work by Akin [9,10], and also Hofbauer, Schuster and Sigmund [11], who worked with a certain function of $I(q,p(t))$ rather than this function itself.

Next we turn to the question: how can we interpret the above inequality, and when does it hold?

3. Evolutionary Game Theory

To go further, evolutionary game theorists sometimes assume the fitness functions are linear in the probabilities p_j. Then

$$f_i(P) = \sum_{j=1}^{n} A_{ij}p_j$$

for some matrix A, called the **fitness matrix**.

In this situation the mathematics is connected to the usual von Neumann–Morgenstern theory of two-player games. In this approach to game theory, each player has the same finite set X of **pure strategies**. The **payoff matrix** A_{ij} specifies the first player's winnings if the first player chooses the pure strategy i and the second player chooses the pure strategy j. A probability distribution on the set of pure strategies is called a **mixed strategy**. The first player's expected winnings will be $p \cdot Aq$ if they use the mixed strategy p and the second player uses the mixed strategy q.

To apply this analogy to game theory, we assume that we have a well-mixed population of replicators. Each one randomly roams around, "plays games" with each other replicator it meets, and reproduces at a rate proportional to its expected winnings. A pure strategy is just what we have been calling a species. A mixed strategy is a probability distribution of species. The payoff matrix is the fitness matrix.

In this context, the vector of fitness functions is given by

$$f(P) = Ap.$$

Then, if $p(t)$ evolves according to the replicator equation while q is fixed, the time derivative of relative information, given in Equation (5), becomes

$$\frac{d}{dt}I(q,p(t)) = (p(t) - q) \cdot Ap(t). \tag{6}$$

Thus, we define q to be a **dominant mixed strategy** if

$$q \cdot Ap \geq p \cdot Ap \tag{7}$$

for all mixed strategies p. If q is dominant, we have

$$\frac{d}{dt} I(q, p(t)) \leq 0, \tag{8}$$

whenever $p(t)$ obeys the replicator equation. Conversely, if the information of $p(t)$ relative to q is nonincreasing whenever $p(t)$ obeys the replicator equation, the mixed strategy q must be dominant.

The question is then: what is the meaning of dominance? First of all, if q is dominant then it is a **steady state** solution of the replicator equation, meaning one that does not depend on time. To see this, let $r(t)$ be the solution of the replicator equation with $r(0) = q$. Then $I(q, r(t))$ is nonincreasing because q is dominant. Furthermore $I(q, r(t)) = 0$ at $t = 0$, since for any probability distribution we have $I(q, q) = 0$. Thus we have $I(q, r(t)) \leq 0$ for all $t \geq 0$. However, relative information is always non-negative, so we must have $I(q, r(t)) = 0$ for all $t \geq 0$. This forces $r(t) = q$, since the relative information of two probability distributions can only vanish if they are equal.

Thus, a dominant mixed strategy is a special kind of steady state solution of the replicator equation. But what is special about it? We can understand this if we think in game-theoretic terms. The inner product $q \cdot Ap$ will be my expected winnings if I use the mixed strategy q and you use the mixed strategy p. Similarly, $p \cdot Ap$ will be my expected winnings if we both use the mixed strategy p. So, the dominance of the mixed strategy q says that my expected winnings can never increase if I switch from q to whatever mixed strategy you are using.

It helps to set these ideas into the context of evolutionary game theory [12]. In 1975, John Maynard Smith, the founder of evolutionary game theory, defined a mixed strategy q to be an "evolutionarily stable state" if when we add a small population of "invaders" distributed according to any other probability distribution p, the original population is more fit than the invaders [13]. He later wrote [14]: "A population is said to be in an evolutionarily stable state if its genetic composition is restored by selection after a disturbance, provided the disturbance is not too large."

More precisely, Maynard Smith defined q to be an **evolutionarily stable state** if

$$q \cdot A((1-\epsilon)q + \epsilon p) > p \cdot A((1-\epsilon)q + \epsilon p)$$

for all mixed strategies $p \neq q$ and all sufficiently small $\epsilon > 0$. Here

$$(1-\epsilon)q + \epsilon p$$

is the population we get by replacing an ϵ-sized portion of our original population by invaders.

Taking this inequality and separating out the terms of order ϵ, one can easily check that q is an evolutionarily stable state if and only if two conditions hold for all probability distributions $p \neq q$:

$$q \cdot Aq \geq p \cdot Aq \tag{9}$$

and

$$q \cdot Aq = p \cdot Aq \;\Rightarrow\; q \cdot Ap > p \cdot Ap. \tag{10}$$

The first condition says that q is a **symmetric Nash equilibrium**. In other words, the invaders can't on average do better playing against the original population than members of the original population are. The second says that if the invaders are just as good at playing against the original population, they must be worse at playing against each other! The combination of these conditions means the invaders won't take over.

Note, however, that the dominance condition (7), which guarantees nonincreasing relative information, is different from either Equation (9) or (10). Indeed, after Maynard Smith came up with his definition of "evolutionarily stable state", Bernard Thomas [15] came up with a different definition. For him, q is an **evolutionarily stable strategy** if Maynard Smith's condition (9) holds along with

$$q \cdot Ap \geq p \cdot Ap. \tag{11}$$

This condition is stronger than Equation (10), so he renamed Maynard Smith's evolutionarily stable states **weakly evolutionarily stable strategies**.

More importantly for us, Equation (11) is precisely the same as the condition we are calling "dominance", which implies that the relative information $I(q, p(t))$ can never increase as $p(t)$ evolves according to the replicator equation. We can interpret $I(q, p(t))$ as the amount of information "left to learn" as the population approaches the dominant strategy.

This idea of evolution as a learning process is exploited by genetic algorithms in artificial intelligence [16]. Conversely, some neuroscientists have argued that individual organisms act to minimize "surprise"—that is, relative information: the information of perceptions relative to predictions [17]. As we shall see in the next section, relative information also has the physical interpretation of free energy. Thus, this hypothesis is known as the "free energy principle". Another hypothesis, that neurons develop in a manner governed by natural selection, is known as "neural Darwinism" [18]. The connection between relative information decrease and evolutionary game theory shows that these two hypotheses are connected.

4. Markov Processes

One limitation of replicator equations is that in these models, when the population of some species is initially zero, it must remain so for all times. Thus, they cannot model mutation, horizontal gene transfer, or other sources of novelty.

The simplest model of mutation is a discrete-time Markov chain, where there is a fixed probability per time for a genome to change from one genotype to another each time it is copied [19]. The information-theoretic aspects of Markov models in genetics have been discussed by Sober and Steel [20]. To stay within our overall framework, here we instead consider continuous-time Markov chains, which we shall simply call Markov processes. These are a very general framework that can be used to describe any system with finite set X of states where the probability that the system is in its i-th state obeys a differential equation of this form:

$$\frac{dp_i}{dt} = \sum_{j \in X} H_{ij} p_j(t),$$

with the matrix H chosen so that total probability is conserved.

In what follows we shall explain a very general result saying that for any Markov process, relative information is nonincreasing [21–23]. It is a form of the Second Law of Thermodynamics. Some call this result the "H-theorem", but this name goes back to Boltzmann, and strictly this name should be reserved for arguments like Boltzmann's which seek to derive the Second Law from time-symmetric dynamics together with time-asymmetric initial conditions [24,25]. The above equation is not time-symmetric, and the relative information decrease holds for all initial conditions.

We can describe a Markov process starting with a directed graph whose nodes correspond to states of some system, and whose edges correspond to transitions between these states. The transitions are labelled by "rate constants", like this:
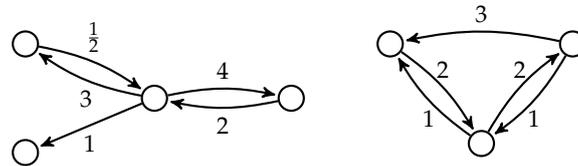
The rate constant of a transition from $i \in X$ to $j \in X$ represents the probability per time that an item hops from the i-th state to the j-th state.

More precisely, we say a **Markov process** M consists of:

- a finite set X of **states**,
- a finite set T of **transitions**,
- maps $s, t: T \to X$ assigning to each transition its **source** and **target**,
- a map $r: T \to (0, \infty)$ assigning a **rate constant** $r(\tau)$ to each transition $\tau \in T$.

If $\tau \in T$ has source i and target j, we write $\tau: i \to j$.

From a Markov process we can construct a square matrix, or more precisely a function $H: X \times X \to \mathbb{R}$, called its **Hamiltonian**. If $i \neq j$ we define

$$H_{ij} = \sum_{\tau: j \to i} r_\tau$$

to be the sum of the rate constants of all transitions from j to i. We choose the diagonal entries in a more subtle way:

$$H_{ii} = - \sum_{\substack{\tau: i \to j \\ j \neq i}} r(\tau).$$

Given a Markov process, the **master equation** for a time-dependent probability distribution on X is:

$$\frac{d}{dt} p(t) = H p(t), \tag{12}$$

where H is the Hamiltonian. Thus, given a probability distribution p on X, for $i \neq j$ we interpret $H_{ij} p_j$ as the rate at which population flows from state j to state i, while the quantity $H_{ii} p_i$ is the outflow of population from state i. The diagonal entries H_{ii} are chosen in a way that ensures total population is conserved.

More precisely, H is **infinitesimal stochastic**, meaning that its off-diagonal entries are non-negative and the entries in each column sum to zero:

$$H_{ij} \geq 0 \;\; \text{if } i \neq j \quad \text{and} \quad \sum_i H_{ij} = 0.$$

This guarantees that if $p(t)$ obeys the master equation and if it is initially a probability distribution, it remains a probability distribution for all times $t \geq 0$.

Markov processes are an extremely general formalism for dealing with randomly evolving systems, and they are presented in many different ways in the literature. For example, besides the master equation one often sees the **Kolmogorov forward equation**:

$$\frac{d}{dt} G(t, s) = H G(t, s)$$

where $G(t, s)$ is a square matrix depending on two times $s, t \in \mathbb{R}$ with $s \leq t$. The idea here is that the matrix element $G_{ij}(t, s)$ is the probability that if the system is in the j-th state at time s, it will be in the i-th state at some later time t. We thus demand that $G(t, s)$ is the identity matrix when $s = t$, and we can show that

$$G(t, s) = \exp((t - s)H),$$

whenever $s \leq t$. From this it is easy to see that $G(t, s)$ also obeys the **Kolmogorov backward equation**:

$$\frac{d}{ds} G(t, s) = -G(t, s)H.$$

We should warn the reader that conventions differ and many, perhaps even most, authors multiply these matrices in the reverse order.

The master equation and Kolmogorov forward equation are related as follows. If $p(t)$ obeys the master equation and $G(t, s)$ solves the Kolmogorov forward equation, then

$$p(t) = G(t, s)p(s),$$

whenever $s \leq t$. Thus, knowledge of $G(t, s)$ immediately tells us all solutions of the master equation.

Most of our discussion so far, and the results to follow, can be generalized to the case where X is an arbitrary measure space, for example $\mathbb{R}^n$. The Kolmogorov forward equation is often studied in this more general context, sometimes in the guise of the "Fokker–Planck equation". This formulation is often used to study Brownian motion and other random walk processes in the continuum. A careful treatment of this generalization involves more analysis: sums become integrals, and one needs to worry about convergence and passing derivatives through integrals [26–29]. To keep things simple and focus on basic concepts, we only treat the case where X is a finite set.

As one evolves any two probability distributions p and q according to a Markov process, their relative information is nonincreasing:

$$\frac{d}{dt} I(p(t), q(t)) \leq 0.$$

This is a very nice result, because it applies regardless of the Markov process. It even applies to a master equation where the Hamiltonian depends on time, as long as it is always infinitesimal stochastic.

To prove this result, we start by computing the derivative:

$$
\begin{aligned}
\frac{d}{dt} I(p(t), q(t)) &= \frac{d}{dt} \sum_{i \in X} p_i \ln\left(\frac{p_i}{q_i}\right) \\
&= \sum_i \dot{p}_i \ln\left(\frac{p_i}{q_i}\right) + p_i \left(\frac{\dot{p}_i}{p_i} - \frac{\dot{q}_i}{q_i}\right) \\
&= \sum_{i,j} H_{ij} p_j \ln\left(\frac{p_i}{q_i}\right) + p_i \left(\frac{H_{ij} p_j}{p_i} - \frac{H_{ij} q_j}{q_i}\right),
\end{aligned}
$$

where in the second line we used the master equation. We can rewrite this as

$$\frac{d}{dt} I(p(t), q(t)) = \sum_{i,j} H_{ij} p_j \left(\ln\left(\frac{p_i}{q_i}\right) + 1 - \frac{p_i q_j}{p_j q_i}\right).$$

Note that the last two terms cancel when $i = j$. Thus, if we break the sum into an $i \neq j$ part and an $i = j$ part, we obtain

$$\frac{d}{dt} I(p(t), q(t)) = \sum_{i \neq j} H_{ij} p_j \left(\ln(\frac{p_i}{q_i}) + 1 - \frac{p_i q_j}{p_j q_i} \right) + \sum_j H_{jj} p_j \ln(\frac{p_j}{q_j}).$$

Next we use the infinitesimal stochastic property of H to write H_{jj} as the sum of $-H_{ij}$ over i not equal to j:

$$\frac{d}{dt} I(p(t), q(t)) = \sum_{i \neq j} H_{ij} p_j \left(\ln(\frac{p_i}{q_i}) + 1 - \frac{p_i q_j}{p_j q_i} \right) - \sum_{i \neq j} H_{ij} p_j \ln(\frac{p_j}{q_j})$$

$$= \sum_{i \neq j} H_{ij} p_j \left(\ln(\frac{p_i q_j}{p_j q_i}) + 1 - \frac{p_i q_j}{p_j q_i} \right).$$

Since $H_{ij} \geq 0$ when $i \neq j$ and $\ln(s) + 1 - s \leq 0$ for all $s > 0$, we conclude that

$$\frac{d}{dt} I(p(t), q(t)) \leq 0 \tag{13}$$

as desired. To be precise, this derivation only applies when q_i is nonzero for all $i \in X$. If this is true at any time, it will be true for all later times. If some probability q_i vanishes, the relative entropy $I(p, q)$ can be infinite. As we evolve p and q in time according to the master equation, the relative entropy can drop from infinity to a finite value, but never increase.

　　One of the nice features of working with a finite state space X is that in this case every Markov process admits one or more **steady states**: probability distributions q that obey

$$Hq = 0,$$

and thus give solutions of the master equation that are constant in time [30]. If we fix any one of these, we can conclude

$$\frac{d}{dt} I(q, p(t)) \leq 0 \tag{14}$$

for any solution of the master equation. This is the same inequality we have already seen for the replicator Equation (8), when q is a dominant mixed strategy. But for a Markov process, we also have

$$\frac{d}{dt} I(p(t), q) \leq 0, \tag{15}$$

and this, it turns out, has a nice meaning in terms of statistical mechanics.

　　In statistical mechanics we want to assign an **energy** E_i to each state such that the steady state probabilities q_i are given by the so-called **Boltzmann distribution**:

$$q_i = \frac{e^{-\beta E_i}}{Z(\beta)}, \tag{16}$$

here β is a parameter which in physics is defined in terms of the temperature T by $\beta = 1/kT$, where k is Boltzmann's constant. The quantity $Z(\beta)$ is a normalizing constant called the **partition function**, defined by

$$Z(\beta) = \sum_{i \in X} e^{-\beta E_i} \tag{17}$$

to ensure that the probabilities q_i sum to one.

However, whenever we have a probability distribution q on a finite set X, we can turn this process on its head. We start by arbitrarily choosing $\beta > 0$. Then we define energy differences by

$$E_i - E_j = -\beta^{-1} \ln\left(\frac{q_i}{q_j}\right). \tag{18}$$

This determines the energies up to an additive constant. If we make a choice for these energies, we can define the partition function by Equation (17), and Boltzmann's law, Equation (16), will follow.

We can thus apply ideas from statistical mechanics to any Markov process, for example the process of genetic drift. The concepts of "energy" and "temperature" play only a metaphorical role here; they are not the ordinary physical energy and temperature. However, the metaphor is a useful one.

So, let us fix a Markov process on a set X together with a steady state probability distribution q. Let us choose a value of β, choose energies obeying Equation (18), and define the partition function $Z(\beta)$ by Equation (17). To help the reader's intuition we define a **temperature** $T = 1/\beta$, setting Boltzmann's constant to 1. Then, for any probability distribution p on X we can define the **expected energy**:

$$\langle E \rangle_p = \sum_{i \in X} p_i E_i$$

and the **entropy**:

$$S(p) = - \sum_{i \in X} p_i \ln(p_i).$$

From these, we can construct the all-important **free energy**

$$F(p) = \langle E \rangle_p - TS(p).$$

In applications to physics and chemistry this is, roughly speaking, the amount of "useful" energy, meaning energy not in the form of random "heat", which gives the term $TS(p)$.

We can prove that this free energy can never increase with time if we evolve p in time according to the master equation. This is a version of the Second Law of Thermodynamics. To prove this, note that

$$
\begin{aligned}
F(p) &= \langle E \rangle_p - TS(p) \\
&= \sum_{i \in X} p_i E_i + T\, p_i \ln(p_i),
\end{aligned}
$$

but by Boltzmann's law, Equation (16), we have

$$E_i = -T \left(\ln(q_i) + \ln(Z) \right),$$

so we obtain

$$F(p) = -T \sum_{i \in X} \left(p_i \ln(q_i) - p_i \ln(p_i) + p_i \ln(Z) \right),$$

or, using the definition of relative information and the fact that the p_i sum to one:

$$F(p) = I(p, q) - T \ln(Z).$$

In the special case where $p = q$ the relative information vanishes and we obtain

$$F(q) = -T \ln(Z).$$

Substituting this into the previous equation, we reach an important result:

$$\frac{F(p) - F(q)}{T} = I(p, q). \tag{19}$$

Relative information is proportional to a difference in free energies! Since relative entropy is nonnegative, we immediately see that any probability distribution p has at least as much free energy as the steady state q:

$$F(p) \geq F(q).$$

Moreover, if we evolve $p(t)$ according to the master equation, the decrease of relative entropy given by Equation (15) implies that

$$\frac{d}{dt}F(p(t)) \leq 0.$$

These two facts suggest, but do not imply, that $p(t) \to q$ as $t \to \infty$. This is in fact true when there is a unique steady state, but not necessarily otherwise. One can determine the number of linearly independent steady states from the topology of the graph associated to the Markov process (Section 22.2 in [30]).

5. Reaction Networks

Reaction networks are commonly used in chemistry, where they are called "chemical reaction networks". An example is the Krebs cycle, important in the metabolism of aerobic organisms. A simpler example is the Michaelis–Menten model of an enzyme E binding to a substrate S to form an intermediate I, which in turn can break apart into a product P and the original enzyme:

$$E + S \underset{\beta}{\overset{\alpha}{\rightleftarrows}} I \overset{\gamma}{\longrightarrow} E + P.$$
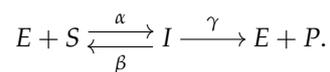
Mathematically, this is a directed graph. The nodes of this graph, namely $E + S$, I, and $E + P$, are called "complexes". Complexes are finite sums of "species", which in this example are E, S, I, and P. The edges of this graph are called "reactions". Each reaction has a name, which may also serve as the "rate constant" of that reaction. In real-world chemistry, every reaction has a reverse reaction going the other way, but if the rate constant for the reverse reaction is low enough, we may simplify our model by omitting it. This is why the Michaelis–Menten model has no reaction going from $E + P$ back to I.

From a reaction network we can extract a differential equation called its "rate equation", which describes how the population of each species changes with time. We treat these populations as functions of time, taking values in $[0, \infty)$. If we use P_E as the name for the population of the species E, and so on, the rate equation for the above reaction network is:

$$\begin{aligned}
\frac{d}{dt}P_E &= -\alpha P_E P_S + \beta P_I + \gamma P_I \\
\frac{d}{dt}P_S &= -\alpha P_E P_S + \beta P_I \\
\frac{d}{dt}P_I &= \alpha P_E P_S - \gamma P_I \\
\frac{d}{dt}P_P &= \gamma P_I
\end{aligned}$$

We will give the general rules for extracting the rate equation from a reaction network, but the reader may enjoy guessing them from this example. It is worth noting that chemists usually deal with "concentrations" rather than populations: a concentration is a population per unit volume. This changes the meaning and the values of the rate constants, but the mathematical formalism is the same.

More precisely, a **reaction network** consists of:

- a finite set S of **species**,
- a finite set X of **complexes** with $X \subseteq \mathbb{N}^S$,
- a finite set T of **reactions** or **transitions**,
- maps $s, t \colon T \to X$ assigning to each reaction its **source** and **target**,

- a map $r\colon T \to (0, \infty)$ assigning to each reaction a **rate constant**.

The reader will note that this looks very much like our description of a Markov process in Section 4. As before, we have a graph with edges labelled by rate constants. However, now instead of the nodes of our graphs being abstract "states", they are complexes: finite linear combinations of species with natural number coefficients, which we can write as elements of $\mathbb{N}^S$.

For convenience we shall often write k for the number of species present in a reaction network, and identify the set S with the set $\{1, \ldots, k\}$. This lets us write any complex as a k-tuple of natural numbers. In particular, we write the source and target of any reaction τ as

$$s(\tau) = (s_1(\tau), \ldots, s_k(\tau)) \in \mathbb{N}^k,$$

$$t(\tau) = (t_1(\tau), \ldots, t_k(\tau)) \in \mathbb{N}^k.$$

The rate equation involves the **population** $P_i \in [0, \infty)$ of each species i. We can summarize these in a population vector

$$P = (P_1, \ldots, P_k).$$

The rate equation says how this vector change with time. It says that each reaction τ contributes to the time derivative of P via the product of

- the vector $t(\tau) - s(\tau)$ whose i-th component is the change in the number of items of the i-th species due to the reaction τ;
- the concentration of each input species i of τ raised to the power given by the number of times it appears as an input, namely $s_i(\tau)$;
- the rate constant $r(\tau)$ of τ.

The **rate equations** are

$$\frac{d}{dt}P(t) = \sum_{\tau \in T} r(\tau)(t(\tau) - s(\tau))P(t)^{s(\tau)}, \tag{20}$$

where $P\colon \mathbb{R} \to [0, \infty)^k$ and we have used multi-index notation to define

$$P^{s(\tau)} = P_1^{s_1(\tau)} \cdots P_k^{s_k(\tau)}.$$

Alternatively, in components, we can write the rate equation as

$$\dot{P}_i = \sum_{\tau \in T} r(\tau)(t_i(\tau) - s_i(\tau))P^{s(\tau)}.$$

The reader can check that this rule gives the rate equations for the Michaelis–Menten model.

Reaction networks include Markov processes as a special case. A reaction network where every complex is just a single species—that is, a vector in $\mathbb{N}^S$ with one component being 1 and all the rest 0—can be viewed as a Markov process. For a reaction network that corresponds to a Markov process in this way, the rate equation is linear, and it matches the master equation for the corresponding Markov process. The goal of this section is to generalize results on relative information from Markov processes to other reaction networks. However, the nonlinearity of the rate equation introduces some subtleties.

The applications of reaction networks are not limited to chemistry. Here is an example that arose in work on HIV, the human immunodeficiency virus [31]:

$$0 \xrightarrow{\alpha} H \xrightarrow{\beta} 0 \qquad H + V \xrightarrow{\gamma} I \xrightarrow{\delta} I + V$$

$$I \xrightarrow{\epsilon} 0 \qquad V \xrightarrow{\zeta} 0.$$
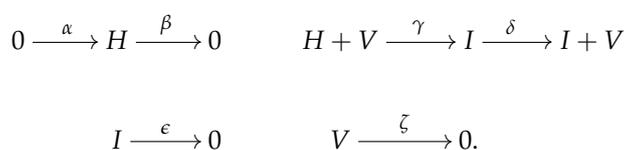
Here we have three species:

- *H*: healthy white blood cells,
- *I*: infected white blood cells,
- *V*: virions (that is, individual virus particles).

The complex 0 above is short for $0H + 0I + 0V$: that is, "nothing". We also have six reactions:

- α: the birth of one healthy cell, which has no input and one *H* as output.
- β: the death of a healthy cell, which has one *H* as input and no output.
- γ: the infection of a healthy cell, which has one *H* and one *V* as input, and one *I* as output.
- δ: the reproduction of the virus in an infected cell, which has one *I* as input, and one *I* and one *V* as output.
- ϵ: the death of an infected cell, which has one *I* as input and no output.
- ζ: the death of a virion, which has one *V* as input and no output.

For this reaction network, if we use the Greek letter names of the reactions as names for their rate constants, we get these rate equations:

$$
\begin{aligned}
\frac{d}{dt}P_H &= \alpha - \beta P_H - \gamma P_H P_V \\
\frac{d}{dt}P_I &= \gamma P_H P_V - \epsilon P_I \\
\frac{d}{dt}P_V &= -\gamma P_H P_V + \delta P_I - \zeta P_V.
\end{aligned}
$$

The equations above are not of the Lotka–Volterra type shown in Equation (3), because the time derivative of P_H contains a term with no factor of P_H, and similarly for P_I and P_V. Thus, even when the population of one of these three species is initially zero, it can become nonzero. However, many examples of Lotka–Volterra equations do arise from reaction networks. For example, we could take two species:

- *R*: rabbits,
- *W*: wolves,

and form this reaction network:

$$
R \xrightarrow{\alpha} 2R \qquad R + W \xrightarrow{\beta} 2W \qquad W \xrightarrow{\gamma} 0.
$$

Taken literally, this seems like a ludicrous model: rabbits reproduce asexually, a wolf can eat a rabbit and instantly give birth to another wolf, and wolves can also die. However, the resulting rate equations are a fairly respectable special case of the famous Lotka–Volterra predator-prey model:

$$
\begin{aligned}
\frac{d}{dt}P_R &= \alpha P_R - \beta P_R P_W \\
\frac{d}{dt}P_W &= \beta P_R P_W - \gamma P_W.
\end{aligned}
$$

It is probably best to think of this as saying no more than this: general results about reaction networks will also apply to Lotka–Volterra equations that can arise from this framework.

In our discussion of the replicator equation, we converted populations to probability distributions by normalizing them, and defined relative information only for the resulting probability distributions. We can, however, define relative information for populations, and this is important for work on reaction networks. Given two populations $P, Q \colon X \to [0, \infty)$, we define

$$
I(P, Q) = \sum_{i \in X} P_i \ln\left(\frac{P_i}{Q_i}\right) - (P_i - Q_i). \tag{21}
$$

When P and Q are probability distributions on X this reduces to the relative information defined before in Equation (1). As before, one can prove that

$$I(P, Q) \geq 0.$$

To see this, note that a differentiable function $f : \mathbb{R} \to \mathbb{R}$ is convex precisely when its graph lies above any of its tangent lines:

$$f(y) \geq f(x) + f'(x)(y - x).$$

This is true for the exponential function, so

$$e^y \geq e^x + e^x(y - x),$$

and thus for any $p, q > 0$ we have

$$q \geq p + p(\ln(q) - \ln(p)),$$

or

$$p \ln(\frac{p}{q}) - (p - q) \geq 0.$$

Thus, each term in the sum Equation (21) is greater than or equal to zero, so $I(P, Q) \geq 0$. Furthermore since we have equalities above only when $x = y$, or in other words $p = q$, we also obtain

$$I(P, Q) = 0 \iff P = Q.$$

So, relative information has the properties of a divergence, but for arbitrary populations $P, Q \colon X \to [0, \infty)$. A function very similar to $I(P, Q)$ was used by Friedrich Horn and Roy Jackson in their important early paper on reaction networks [32]. They showed that this function is nonincreasing when P evolves according to the rate equation and Q is a steady state of a special sort, called a "complex balanced equilibrium". Later Martin Feinberg, another of the pioneers of reaction network theory, gave a shorter proof of this fact [33]. Our goal here is to explain this result and present Feinberg's proof.

We say that a population Q is **complex balanced** if for each complex $\kappa \in K$ we have

$$\sum_{\tau : s(\tau) = \kappa} r(\tau) Q^{s(\tau)} = \sum_{\tau : t(\tau) = \kappa} r(\tau) Q^{s(\tau)}. \tag{22}$$

This says that each *complex* is being produced at the same rate at which it is being destroyed. This is stronger than saying Q is a steady state solution of the rate equation. On the other hand, it is weaker than the "detailed balance" condition saying that each reaction occurs at the same rate as the reverse reaction. The founders of chemical reaction theory discovered that many results about detailed balanced equilibria can just as easily be shown in the complex balanced case. The calculation below is an example.

We have

$$\frac{d}{dt} I(P(t), Q) = \sum_{i \in X} \dot{P}_i \ln(\frac{P_i}{Q_i}),$$

so, using the rate Equation (20), we obtain:

$$
\begin{aligned}
\frac{d}{dt} I(P(t), Q) &= \sum_{i \in X} \sum_{\tau \in T} r(\tau) (t_i(\tau) - s_i(\tau)) \ln(\frac{P_i}{Q_i}) \, P^{s(\tau)} \\
&= \sum_{i \in X} \sum_{\tau \in T} r(\tau) \left[\ln\left(\left(\frac{P_i}{Q_i}\right)^{t_i(\tau)} \right) - \ln\left(\left(\frac{P_i}{Q_i}\right)^{s_i(\tau)} \right) \right] P^{s(\tau)}.
\end{aligned}
$$

We can convert each sum over i of the logarithms into a logarithm of a product, and if we define a vector

$$\frac{P}{Q} = (\frac{P_1}{Q_1}, \dots, \frac{P_k}{Q_k}),$$

we can use multi-index notation to write these products very concisely, obtaining

$$
\begin{aligned}
\frac{d}{dt} I(P(t), Q) &= \sum_{\tau \in T} r(\tau) \left[\ln \left(\left(\frac{P}{Q} \right)^{t(\tau)} \right) - \ln \left(\left(\frac{P}{Q} \right)^{s(\tau)} \right) \right] P^{s(\tau)} \\
&= \sum_{\tau \in T} r(\tau) \left[\ln \left(\left(\frac{P}{Q} \right)^{t(\tau)} \right) - \ln \left(\left(\frac{P}{Q} \right)^{s(\tau)} \right) \right] \left(\frac{P}{Q} \right)^{s(\tau)} Q^{s(\tau)}.
\end{aligned}
$$

Then, using the fact that $(\ln x - \ln y)y \le x - y$, we obtain

$$\frac{d}{dt} I(P(t), Q) \le \sum_{\tau \in T} r(\tau) \left[\left(\frac{P}{Q} \right)^{t(\tau)} - \left(\frac{P}{Q} \right)^{s(\tau)} \right] Q^{s(\tau)}.$$

Next, we write the sum over reactions as a sum over complexes κ and then a sum over reactions having κ is their target (for the first term) or target (for the second):

$$\frac{d}{dt} I(P(t), Q) \le \sum_{\kappa \in X} \left[\sum_{\tau : t(\tau) = \kappa} r(\tau) \left(\frac{P}{Q} \right)^{t(\tau)} Q^{s(\tau)} - \sum_{\tau : s(\tau) = \kappa} r(\tau) \left(\frac{P}{Q} \right)^{s(\tau)} Q^{s(\tau)} \right].$$

We can pull out the factors involving $\frac{P}{Q}$:

$$\frac{d}{dt} I(P(t), Q) \le \sum_{\kappa \in X} \left(\frac{P}{Q} \right)^{\kappa} \left[\sum_{\tau : t(\tau) = \kappa} r(\tau) Q^{s(\tau)} - \sum_{\tau : s(\tau) = \kappa} r(\tau) Q^{s(\tau)} \right],$$

but now the right side is zero by the complex balanced condition, Equation (22). Thus, we have

$$\frac{d}{dt} I(P(t), Q) \le 0, \tag{23}$$

whenever $P(t)$ evolves according to the rate equation and Q is a complex balanced equilibrium.

As noted above, a reaction network where every complex consists of a single species gives a linear rate equation. In this special case we can strengthen the above result: we have

$$\frac{d}{dt} I(P(t), Q(t)) \le 0, \tag{24}$$

whenever $P(t)$ and $Q(t)$ evolve according to the rate equation. The reason is in this case, the rate equation is also the master equation for a Markov process. Thus, we can reuse the argument leading up to inequality (13) for Markov processes, since nothing in this argument used the fact that the probability distributions were normalized.

6. Conclusions

We have seen theorems guaranteeing that relative information cannot increase in three different situations: evolutionary games described by the replicator equation, Markov processes, and reaction networks. In all cases, the decrease of relative entropy is closely connected to the approach to equilibrium as $t \to \infty$. For the replicator equation, this equilibrium is a dominant mixed strategy. For a Markov process, whenever there is a unique steady state, all probability distributions approach this

steady state as $t \to \infty$. For reaction networks, the appropriate notion of equilibrium is a complex balanced equilibrium, generalizing the more familiar concept of detailed balanced equilibrium.

It is natural to inquire about the mathematical relation between these results. Inequality (24) for Markov processes resembles inequality (23) for reaction networks. However, neither result subsumes the other. The master equation for a Markov process is a special case of the rate equation for a reaction network. However, the result for reaction networks says only that

$$\frac{d}{dt} I(P(t), Q) \leq 0,$$

when Q is a complex balanced equilibrium and $P(t)$ obeys the rate equation, while the result for Markov processes says that

$$\frac{d}{dt} I(P(t), Q(t)) \leq 0,$$

whenever $P(t)$ and $Q(t)$ obey the master equation. Furthermore, neither of these inequalities subsume or are subsumed by the result for the replicator equation, inequality (8). Indeed, this result applies only to the probability distributions obtained by normalizing population distributions, not populations. Furthermore it is "turned around", in that sense that q appears first:

$$\frac{d}{dt} I(q, p(t)) \leq 0,$$

whenever $p(t)$ obeys the replicator equation and q is a dominant strategy. We know of no results showing that $I(p(t), q)$ is nonincreasing when $p(t)$ obeys the replicator equations, nor results showing that $I(P(t), Q)$ or $I(Q, P(t))$ is nonincreasing when $P(t)$ obeys the Lotka–Volterra equation.

In short, while relative entropy is nonincreasing in the approach to equilibrium in all three situations considered here, the details differ in significant ways. A challenging open problem is thus to find some "super-theorem" that has all three of these results as special cases. The work of Gorban [21] is especially interesting in this regard, since he tackles the challenge of finding new nonincreasing functions for reaction networks.

Author Contributions: John C. Baez and Blake S. Pollard did the research, carried out the computations and wrote the paper. Both authors have read and approved the final manuscript.

References

1. Crooks, G.E. On measures of entropy and information. Available online: http://threeplusone.com/info (accessed on 27 January 2016).
2. Gorban, A.N.; Gorban P.A.; Judge, G. Entropy: The Markov ordering approach. *Entropy* **2010**, *12*, 1145–1193.
3. Hobson, A. *Concepts in Statistical Mechanics*; Gordon and Breach: New York, NY, USA, 1971.
4. Baez, J.C.; Fritz, T. A Bayesian characterization of relative entropy. *Theory Appl. Categories* **2014**, *29*, 422–456.
5. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2006.
6. Pollard, B. Open Markov processes: a compositional perspective on non-equilibrium steady states in biology. **2016**, arXiv:1601.00711.
7. Harper, M. Information geometry and evolutionary game Theory. **2009**, arXiv:0911.1383.
8. Harper, M. The replicator equation as an inference dynamic. **2009**, arXiv:0911.1763.
9. Akin, E. *The Geometry of Population Genetics*; Springer: Berlin/Heidelberg, Germany, 1979.

10. Akin, E. The differential geometry of population genetics and evolutionary games. In *Mathematical and Statistical Developments of Evolutionary Theory*; Lessard, S., Ed.; Springer: Berlin/Heidelberg, Germany, 1990; pp. 1–93.

11. Hofbauer, J.; Schuster, P.; Sigmund, K. A note on evolutionarily stable strategies and game dynamics. *J. Theor. Biol.* **1979**, *81*, 609–612.

12. Sandholm, W.H. Evolutionary game theory. Available online: http://www.ssc.wisc.edu/~whs/research/egt.pdf (accessed on 27 January 2016).

13. Smith, J.M. Game theory and the evolution of fighting. In *On Evolution*; Edinburgh University Press: Edinburgh, UK, 1972; pp. 8–28.

14. Smith, J.M. *Evolution and the Theory of Games*; Cambridge University Press: Cambridge, UK, 1982.

15. Thomas, B. On evolutionarily stable sets. *J. Math. Biol.* **1985**, *22*, 105–115.

16. Mitchell, M. *An Introduction to Genetic Algorithms*; MIT Press: Cambridge, MA, USA, 1998.

17. Friston, K.; Ao, P. Free energy, value, and attractors. *Comput. Math. Methods Med.* **2012**, *2012*, 937860.

18. Edelman, G.M. *Neural Darwinism: The Theory of Neuronal Group Selection*; Basic Books: New York, NY, USA, 1987.

19. Nielsen, R. *Statistical Methods in Molecular Evolution*; Springer: Berlin/Heidelberg, Germany, 2005.

20. Sober, E.; Steel, M. Entropy increase and information loss in Markov models of evolution. *Biol. Philos.* **2011**, *26*, 223–250.

21. Gorban, A.N. General *H*-theorem and entropies that violate the Second Law. *Entropy* **2014**, *16*, 2408–2432.

22. Liese, F.; Vajda, I. *Convex Statistical Distances*; Teubner: Leipzig, Germany, 1987.

23. Moran, P.A.P. Entropy, Markov processes and Boltzmann's *H*-theorem. *Math. Proc. Camb. Philos. Soc.* **1961**, *57*, 833–842.

24. Price, H. *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time*; Oxford University Press: Oxford, UK, 1997.

25. Zeh, H.D. *The Physical Basis of the Direction of Time*; Springer: Berlin/Heidelberg, Germany, 2001.

26. Norris, J.R. *Markov Processes*; Cambridge University Press: Cambridge, UK, 1997.

27. Rogers, L.C.G.; Williams, D. *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2000.

28. Rogers, L.C.G.; Williams, D. *Diffusions, Markov Processes, and Martingales: Volume 2, Itô Calculus*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2000.

29. Ethier, S.N.; Kurtz, T.G. *Markov Processes: Characterization and Convergence*; Wiley: Hoboken, NJ, USA, 2005.

30. Baez, J.C.; Biamonte, J. Quantum Techniques for Stochastic Mechanics. **2015**, arXiv:1209.3632.

31. Korobeinikov, A. Global properties of basic virus dynamics models. *Bull. Math. Biol.* **2004**, *66*, 879–883.

32. Horn, F.; Jackson, R. General mass action kinetics. *Arch. Ration. Mech. Anal.* **1972**, *49*, 81–116.

33. Feinberg, M. Lectures on chemical reaction networks. Available online: https://crnt.osu.edu/LecturesOnReactionNetworks (accessed on 27 January 2016).

34. Gopalkrishnan, M. Lyapunov functions for complex-balanced systems. Available online: https://johncarlosbaez.wordpress.com/2014/01/07/lyapunov-functions-for-complex-balanced-systems/ (accessed on 27 January 2016).

35. Anderson, D. Comment on Azimuth Blog, 2014. Available online: https://johncarlosbaez.wordpress.com/2014/01/07/lyapunov-functions-for-complex-balanced-systems/#comment-35537 (accessed on 27 January 2016).