The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific
and Cultural Organization

International Atomic
Energy Agency

SMR.1763- 23

SCHOOL and CONFERENCE
on
COMPLEX SYSTEMS
and
NONEXTENSIVE STATISTICAL MECHANICS

31 July - 8 August 2006

Transition records of stationary Markov chains

J. Naudts

Departement Fysica
Universiteit Antwerpen
Groenenborgerlaan 171
2020 Antwerpen
Belgium

Transition records of stationary Markov chains

Jan Naudts and Erik Van der Straeten
*Departement Fysica, Universiteit Antwerpen, Groenenborgerlaan 171, 2020 Antwerpen, Belgium**

In any Markov chain with finite state space the distribution of transition records always belongs to the exponential family. This observation is used to prove a fluctuation theorem, and to show that the dynamical entropy of a stationary Markov chain is linear in the number of steps. Three applications are discussed. A known result about entropy production is reproduced. A thermodynamic relation is derived for equilibrium systems with Glauber dynamics. Finally, a link is made with recent results concerning a one-dimensional polymer model.

Introduction The fluctuation theorem, discovered in 1993 by Evans, Cohen and Morriss [1], was the start of a series of new developments in the study of stationary non-equilibrium systems (see for example [2] for an overview). A formulation of the fluctuation theorem in stochastic context was given in [3, 4]. These authors showed that the theorem holds in full generality for any stationary probability distribution of any Markov process. The proof given here is restricted to Markov chains with finite state space. In this way technicalities can be avoided.

The method of proof of the present paper is based on the observation that a certain probability distribution, hereafter called *distribution of transition records*, always belongs to the exponential family. From this result the fluctuation theorem follows immediately. Indeed, Lebowitz and Spohn [3] and Maes [4] already observed that a Gibbs distributed random variable is all what is needed to derive a fluctuation theorem of the Gallavotti-Cohen type [5] in a stochastic context. This point has been further elaborated in [6].

The motivation for studying the distribution of transition records was twofold. At one hand the present authors had noticed [7, 8] that, in the Markov chain of increments of some model of random walk describing a polymer, this distribution belongs automatically to the exponential family. At the other hand, it is a straightforward generalization of a distribution introduced in recent work of Carati [9, 10]. This author assumed a Poisson distribution based on the assumption that subsequent visits to macrocells in phase space of a system of classical mechanics are mutually independent.

Paths of Markov chains Throughout the paper we consider a fixed state space Γ, containing a finite number N of states. A Markov chain with state space Γ is determined by initial probabilities $p(x)$, with x in Γ, and by transition probabilities $w(x,y)$, with x and y in Γ.

The probability distribution $p(x)$ is *stationary* if

$$p(x) = \sum_{y\in\Gamma} p(y)w(y,x) \tag{1}$$

for all $x \in \Gamma$. A stronger condition is *detailed balance*

$$p(x)w(x,y) = p(y)w(y,x) \tag{2}$$

for any pair of states $x, y \in \Gamma$. Summing this equality over y yields (1).

The $w(x,y)$ are parameters of the Markovian model. However, they are not independent because of the normalization condition $\sum_y w(x,y) = 1$. Introduce therefore parameters $\theta_{x,y}$ and η_x such that

$$\begin{aligned} w(x,y) &= w_0(x,y)e^{-\theta_{x,y}} \qquad \text{if } x \neq y \\ w(x,x) &= e^{-\eta_x}, \end{aligned} \tag{3}$$

with *a priori* transition rates $w_0(x,y)$ satisfying $w_0(x,y) = w_0(y,x)$, and $w_0(x,y) = 1$ or $w_0(x,y) = 0$. The parameters η_x are considered to be functions of the $\theta_{x,y}$. For convenience, let $\theta_{x,x} = 0$ and $w_0(x,x) = 1$ for all x.

A Gibbs distribution for transition records The record of transitions k is a sequence of numbers $k_{x,y}$, one for each pair of states x, y, counting how many times the transition from x to y is contained in a given path of the Markov chain. Our key quantity is the probability $d_x(k)$ that a path starting in x results in a given record of transitions k.

It is not difficult to see that one can write

$$d_x(k) = c_n(x,k) \prod_{y,z} w(y,z)^{k_{y,z}}$$

*E-mail jan.naudts@ua.ac.be, erik.vanderstraeten@ua.ac.be

$$= c_n(x,k)\exp\left(-\Psi_\theta(k)\right). \tag{4}$$

In this expression, $c_n(x,k)$ counts the number of paths that have the same record of transitions k and that are allowed by the a priori transition rates $w_0(x,y)$. The quantity $\Psi_\theta(k)$ is given by

$$\Psi_\theta(k) = \sum_y \eta_y k_{y,y} + \sum_{y,z} \theta_{y,z} k_{y,z}. \tag{5}$$

The special form of (4) makes it easy to calculate the following identities

$$\begin{aligned} 0 &= \frac{\partial}{\partial\theta(u,v)} \sum_k d_x(k) \\ &= -\langle k_{u,v}\rangle_x + \frac{w(u,v)}{w(u,u)}\langle k_{u,u}\rangle_x, \end{aligned} \tag{6}$$

with

$$\langle k_{u,v}\rangle_x = \sum_k d_x(k) k_{u,v}. \tag{7}$$

This can be written as

$$w(u,u)\langle k_{u,v}\rangle_x = \langle k_{u,u}\rangle_x w(u,v). \tag{8}$$

The latter result is clearly also valid for $u = v$. This means that $\langle k_{u,v}\rangle_x$, the average number of transitions from u to v, is proportional to $w(u,v)$, the probability to go from u to v.

Linear production of dynamical entropy A path $\gamma = (x_0, x_1, \cdots x_n)$ of the Markov chain has probability $p(x_0)w(\gamma)$ with

$$w(\gamma) = w(x_0,x_1)\cdots w(x_{n-1},x_n). \tag{9}$$

Let $\gamma_i = x_0$ denote the initial and $\gamma_f = x_n$ the final state of γ. The *dynamical entropy* $S_\theta^{(n)}$ of the Markov chain is defined by

$$S_\theta^{(n)} = -\sum_\gamma p(\gamma_i) w(\gamma) \ln w(\gamma). \tag{10}$$

The time-reversed dynamical entropy [11] is defined by

$$\begin{aligned} \overline{S}_\theta^{(n)} &= -\sum_\gamma p(\gamma_i) w(\gamma) \ln w(\overline{\gamma}), \\ &= -\sum_\gamma p(\gamma_f) w(\overline{\gamma}) \ln w(\gamma), \end{aligned} \tag{11}$$

where $\overline{\gamma}$ is the reversed path $\overline{\gamma} = (x_n, x_{n-1}, \cdots x_0)$.

Observe that the dynamical entropy $S_\theta^{(n)}$ can be expressed in terms of the distributions $d_x(k)$

$$S_\theta^{(n)} = \sum_x p(x)\langle\Psi_\theta\rangle_x. \tag{12}$$

If $p(x)$ is stationary then this expression can be further simplified. One expects intuitively that each occurrence of some state y contributes to the dynamic entropy production with an amount

$$I_y = -\sum_u w(y,u)\ln w(y,u). \tag{13}$$

Indeed, this is the result of Theorem 1 below.

The corresponding relations for the time-reversed paths are

$$\overline{S}_\theta^{(n)} = \sum_x p(x)\langle\overline{\Psi}_\theta\rangle_x. \tag{14}$$

with $\overline{\Psi}_\theta(k) = \Psi_\theta(\overline{k})$, $\overline{k}_{x,y} = k_{y,x}$, and

$$\overline{I}_y = -\sum_u w(y,u) \ln w(u,y). \tag{15}$$

Given a path with record of transitions k, the number of occurrences of the state y (neglecting the final state) is $\sum_z k_{y,z}$. The probability distribution of the latter quantity is the distribution studied by Carati [9, 10]. Using (8), the average number of occurrences of state y is

$$\sum_z \langle k_{y,z} \rangle_x = \frac{1}{w(y,y)} \langle k_{y,y} \rangle_x. \tag{16}$$

The average entropy produced when leaving y is I_y. Hence

$$\sum_y \frac{1}{w(y,y)} \langle k_{y,y} \rangle_x I_y \tag{17}$$

is the average production of dynamical entropy per step. One therefore expects the following result.

Lemma 1 *Assume $w(x,x) > 0$ for all x. Then one has*

$$\langle \Psi_\theta \rangle_x = \sum_y \frac{1}{w(y,y)} \langle k_{y,y} \rangle_x I_y, \tag{18}$$

$$\langle \overline{\Psi}_\theta \rangle_x = \sum_y \frac{1}{w(y,y)} \langle k_{y,y} \rangle_x \overline{I}_y. \tag{19}$$

Next, one proves by full induction that

Lemma 2 *Let p be a stationary probability distribution. Then*

$$\sum_x p(x) \langle k_{y,y} \rangle_x = n w(y,y) p(y). \tag{20}$$

Combining the different pieces gives the result

Theorem 1 *Assume p is stationary. Then*

$$S_\theta^{(n)} = n \sum_x p(x) I_x, \tag{21}$$

$$\overline{S}_\theta^{(n)} = n \sum_x p(x) \overline{I}_x. \tag{22}$$

Fluctuation theorem Without any assumption about the Markov chain, the probability distribution $d_x(k)$ belongs automatically to the exponential family. As noted in the introduction, this observation suffices to derive a fluctuation theorem.

Let $C(x,k)$ denote the class of all paths that start in state x and have the same record of transitions k. The probability of this class is $d_x(k)$. It is easy to see that all paths of this class have the same final state, denoted $f(x,k)$. This implies that there is a one-to one-correspondence between paths of $C(x,k)$ and reversed paths belonging to $C(f(x,k),\overline{k})$. Introduce the quantity

$$W_x(k) = \ln \frac{d_x(k)}{d_{f(x,k)}(\overline{k})} \tag{23}$$

(called *action functional* in [3]). Then (4) implies that

$$W_x(k) = -\Psi_\theta(k) + \Psi_\theta(\overline{k}). \tag{24}$$

It is now straightforward to derive the fluctuation theorem

Theorem 2

$$\frac{\mathrm{Prob}(W_x(k)=K)}{\mathrm{Prob}(W_x(k)=-K)} = e^K. \tag{25}$$

Proof
One has

$$\begin{aligned}
\mathrm{Prob}(W_x(k)=K) &= \sum_k d_x(k)\delta_{-\Psi_\theta(k)+\Psi_\theta(\overline{k}),K} \\
&= e^{K/2}\sum_k c_n(x,k)w_0(k)\exp\left(-\frac{1}{2}\Psi_\theta(k)-\frac{1}{2}\Psi_\theta(\overline{k})\right)\delta_{-\Psi_\theta(k)+\Psi_\theta(\overline{k}),K} \\
&= e^{K/2}\sum_k c_n(f(x,k),\overline{k})w_0(k)\exp\left(-\frac{1}{2}\Psi_\theta(k)-\frac{1}{2}\Psi_\theta(\overline{k})\right)\delta_{-\Psi_\theta(\overline{k})+\Psi_\theta(k),K} \\
&= e^K\mathrm{Prob}(W_x(k)=-K).
\end{aligned} \tag{26}$$

□

Note that the condition of Lemma 1 that $w(x,x)>0$ for all x is not essential for the theorem to hold. It can be removed by a limiting procedure.

Entropy production The standard and the time-reversed dynamical entropies are related by

$$\overline{S}_\theta^{(n)} - S_\theta^{(n)} = \sum_x p(x)\langle W_x\rangle_x. \tag{27}$$

Application of the fluctuation theorem then gives

$$\begin{aligned}
\overline{S}_\theta^{(n)} - S_\theta^{(n)} &= \sum_x p(x)\sum_K K\mathrm{Prob}(W_x(k)=K) \\
&= \frac{1}{2}\sum_x p(x)\sum_K \mathrm{Prob}(W_x(k)=K) \\
&\quad \times K(1-e^{-K}) \\
&\geq 0.
\end{aligned} \tag{28}$$

But using Theorem 1 and assuming stationarity, it is possible to write

$$\overline{S}_\theta^{(n)} - S_\theta^{(n)} = n\sum_x p(x)(\overline{I}_x - I_x) = n\Delta S, \tag{29}$$

with

$$\begin{aligned}
\Delta S &= \frac{1}{2}\sum_{x,y}(p(x)w(x,y)-p(y)w(y,x)) \\
&\quad \times \ln\frac{p(x)w(x,y)}{p(y)w(y,x)}.
\end{aligned} \tag{30}$$

ΔS is the entropy production [3, 11, 12]. It vanishes for distributions satisfying the detailed balance condition.

A one-parameter model of equilibrium states We now restrict ourselves to equilibrium, which means that probability distributions satisfy the detailed balance condition. We start from a symmetric matrix $a_{x,y}$ labelled by x, y in state space. We fix a parameter ξ, which is the analogue of the inverse temperature β. In fact, in the case of Glauber dynamics the matrix $a_{x,y}$ is given by

$$a_{x,y} = \max\{H(x), H(y)\}, \tag{31}$$

where $H(x)$ is the Hamiltonian, and ξ equals β.

Introduce a partition function $\Xi(\xi)$, a probability distribution $p(x)$, and transition probabilities $w(x,y)$, given by

$$\Xi(\xi) = \sum_{x,y} e^{-\xi a_{x,y}}, \tag{32}$$

$$p(x) = \frac{1}{\Xi(\xi)} \sum_{y} e^{-\xi a_{x,y}}, \tag{33}$$

$$w(x,y) = \frac{1}{\Xi(\xi)p(x)} e^{-\xi a_{x,y}}. \tag{34}$$

Clearly, $a_{x,y} = a_{y,x}$ implies that the detailed balance condition is satisfied. It is now straightforward to calculate the distributions $d_x(k)$. They are given by (4), with

$$\Psi_\theta(k) = \xi \sum_{x,y} k_{x,y} a_{x,y} + \sum_x \left(\sum_y k_{x,y} \right) \ln p(x) + n \ln \Xi(\xi). \tag{35}$$

A short calculation, using (35) and the results of the section on linear production of entropy, gives

$$\ln \Xi(\xi) = S(p) + \frac{1}{n} S_\theta^{(n)} - \xi \langle a \rangle. \tag{36}$$

Here, $S(p)$ is the Boltzmann-Gibbs entropy, and $\langle a \rangle$ is the path average of the matrix $a_{x,y}$. Expression (36) should be compared with the well-known result for the Massieu function $\ln \Xi(\beta) = S - \beta U$, where $U = \langle H \rangle$ is the average energy. This comparison shows that in (36) the thermodynamic entropy is replaced by the sum of two contributions, a static $S(p)$ and a dynamic $\frac{1}{n} S_\theta^{(n)}$. Thermodynamic relations similar to (36) are derived by Carati [9, 10]. Here, we do not need a Legendre transformation to define entropy. In addition, (35) contains a path dependent term, which is absent in [9, 10].

Related results are found in [12, 13], where the parameter s, appearing in the dynamical partition function, is the argument of a Laplace transform. Its role is comparable to that of our parameter ξ, which controls the transition probabilities of the Markov chain.

A model of random walk Let us finally make the connection with the non-Markovian random walk model of [7, 8]. The position after n steps is denoted X_n. The increments $x_n = X_{n+1} - X_n$ are Markovian with transition matrix

$$w = \begin{pmatrix} \epsilon & 1-\epsilon \\ 1-\mu & \mu \end{pmatrix}. \tag{37}$$

The two states are denoted $+$ and $-$. The a priori rates are all equal to 1.

Let $X_0 = 0$. The position X_n after n steps is related to the record of transitions k by $X_n = k_{++} + k_{-+} - k_{--} - k_{+-}$ (number of steps to the right minus number of steps to the left). The energy of the polymer may be assumed to be proportional to $k_{+-} + k_{-+}$, which is the number of changes of direction. Finally, let $\Delta = k_{+-} - k_{-+}$ and note that Δ is either 0 or ± 1. Together with the identity $n = k_{++} + k_{-+} + k_{--} + k_{+-}$, this means that the record of transitions k can be expressed in terms of the physical quantities energy and position of the end point, up to an error Δ. This observation was used in [7, 8] to prove that the joint probability distribution of energy and position of endpoint automatically belongs to the exponential family, up to a small error which is negligible in the limit of large n.

Discussion We have shown that some known results about Markov chains can be formulated in terms of the distributions $d_x(k)$ which give the probability of a transition record k for paths of the Markov chain starting in state x. These results are connected with what is known as the fluctuation theorem. In addition we have pointed out, without going into much detail, that recent work of Carati [9, 10], and of the present authors [7, 8] involves the same or related probability distributions. We expect many more applications of our approach. We believe that the distribution of transition records $d_x(k)$ will be the preferred tool, rather than the fluctuation theorem (Theorem 2).

Like in [14, 15], the fluctuation theorem and the extensivity of the entropy production, as proved here, hold for arbitrary number of steps n because of the Markov assumption. Usually, these results hold only in average. In particular, the assumption of a Gibbsian field in space-time, as studied in [4], is more general than Markovianity. A generalization of our approach to this context is unlikely because it spoils the equivalence of paths that have the same transition record.

Acknowledgments

We thank Christian Maes for helpful comments. EVdS is Research Assistant of the Research Foundation - Flanders (FWO - Vlaanderen).

---

[1] D.J. Evans, E.G.D. Cohen, G.P. Morriss, *Probability of second law violations in shearing steady states,* Phys. Rev. Lett. **71**, 2401-2404 (1993).
[2] D.J. Evans, D.J. Searles, L. Rondonic, *Application of the Gallavotti-Cohen fluctuation relation to thermostatted steady states near equilibrium,* Phys. Rev. E**71**, 056120 (2005).
[3] J. Lebowitz, H. Spohn, *A Gallavotti-Cohen Type Symmetry in the Large Deviation Functional for Stochastic Dynamics,* J. Stat. Phys. **95**, 333- (1999)
[4] Ch. Maes, *The fluctuation theorem as a Gibbs property,* J. Stat. Phys. **95**, 367- (1999).
[5] G. Gallavotti, E.G.D. Cohen, *Dynamical ensembles in nonequilibrium statistical mechanics,* Phys. Rev. Lett. **74**, 2694-2697 (1995).
[6] Da-Quan Jiang, Min Qian, Fu-Xi Zhang, *Entropy production fluctuations of finite Markov chains,* J. Math. Phys. **44**(9), 4176-4188 (2003).
[7] E. Van der Straeten, J. Naudts, *A one-dimensional model for theoretical analysis of single molecule experiments,* J. Phys. A: Math. Gen. 39, 5715-5726 (2006).
[8] E. Van der Straeten, J. Naudts, *A two-parameter random walk with approximate exponential probability distribution,* J. Phys. A: Math. Gen. 39, 7245-7256 (2006).
[9] A. Carati, *Thermodynamics and time averages,* Physica A**348**, 110-120 (2005).
[10] A. Carati, *On the definition of temperature using time-averages,* Physica A**369**, 417-431 (2006).
[11] P. Gaspard, *Time-Reversed Dynamical Entropy and Irreversibility in Markovian Random Processes,* J. Stat. Phys. **117**, 599-615 (2004).
[12] V. Lecomte, C. Appert-Rolland, F. van Wijland, *Chaotic Properties of Systems with Markov Dynamics,* Phys. Rev. Lett. **95**, 010601 (2005).
[13] V. Lecomte, C. Appert-Rolland, F. van Wijland, *Thermodynamic formalism for systems with Markov dynamics,* arXiv:cond-mat/0606211 (2006).
[14] J. Kurchan, *Fluctuation theorem for stochastic dynamics,* J. Phys. A**31**, 3719-3729 (1998).
[15] G.E. Crooks, *Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences,* Phys. Rev. E**60**(3), 2721-2726 (1999).