Physica A 391 (2012) 544–551

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

H-theorems for the Brownian motion on the hyperbolic plane

C. Vignat [a,b,*], P.W. Lamberti [c]

[a] L.S.S., Supelec, Université d'Orsay, France
[b] L.T.H.I., E.P.F.L., Lausanne, Switzerland
[c] Facultad de Matematica, Astronomia y Fisica, Universidad Nacional de Cordoba and CONICET, Argentina

ARTICLE INFO

Article history:
Received 16 September 2010
Received in revised form 26 August 2011
Available online 7 September 2011

Keywords:
H-theorem
Brownian motion
Hyperbolic plane

ABSTRACT

We study H-theorems associated with the Brownian motion with constant drift on the hyperbolic plane. Since this random process satisfies a linear Fokker–Planck equation, it is easy to show that, up to a proper scaling, its Shannon entropy is increasing over time. As a consequence, its distribution is converging to a maximum Shannon entropy distribution which is also shown to be related to the non-extensive statistics. In a second part, relying on a theorem by Shiino, we extend this result to the case of Tsallis entropies: we show that under a variance-like constraint, the Tsallis entropy of the Brownian motion on the hyperbolic plane is increasing provided that the non-extensivity parameter of this entropy is properly chosen in terms of the drift of the Brownian motion.

1. Introduction

The existence of an H-theorem for a statistical system subjected to some constraints ensures that the Shannon entropy of its probability density function (p.d.f.) f_t at time t,

$$h\left(f_t\right) = -\int_{\mathbb{R}} f_t\left(x\right) \log f_t\left(x\right) \mathrm{d}x,$$

is increasing with time. As a consequence, the asymptotic (stationary) p.d.f. $f_\infty\left(x\right)$ of the system is the maximum Shannon entropy p.d.f. that satisfies the given constraints.

The Shannon entropy is a particular member of a family of information measures called Tsallis entropies that were introduced [1] in 1988 by Tsallis in the context of statistical physics; they are defined as

$$h_q\left(f_t\right) = \frac{1}{1-q}\left(\int_{\mathbb{R}} f_t^q\left(x\right) \mathrm{d}x - 1\right),$$

where $q > 0$ is the non-extensivity parameter. It can be checked using the l'Hôpital rule that the Shannon entropy is the limit case $\lim_{q\to 1} h_q\left(f_t\right) = h\left(f_t\right)$.

A natural question then arises: under what conditions does an H-theorem extend to an H_q-theorem, where the Shannon entropy is replaced by a Tsallis entropy? Several studies have been devoted to this problem in recent years: for example, Plastino and Plastino [2] study the conditions of existence of an H_q-theorem for the following non-linear Fokker–Planck equation

$$\frac{\partial f_t}{\partial t} = -\frac{\partial}{\partial x}\left(K\left(x\right) f_t\left(x\right)\right) + \frac{1}{2} Q \frac{\partial^2}{\partial x^2} f_t^{2-q}\left(x\right)$$

* Corresponding author at: L.S.S., Supelec, Université d'Orsay, France.
E-mail addresses: christophe.vignat@epfl.ch, vignat@univ-mlv.fr (C. Vignat).

0378-4371/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2011.08.069

for some parameter q, while Tsallis and Bukman [3] apply the same approach to the equation

$$\frac{\partial f_t^{\mu}(x)}{\partial t} = -\frac{\partial}{\partial x}\left(F(x) f_t^{\mu}(x)\right) + D\frac{\partial^2}{\partial x^2} f_t^{\nu}(x)$$

for some positive parameters μ and ν.

Our aim in this paper is to show the existence of both an H- and an H_q-theorem for the Brownian motion – more precisely its x-component – with constant drift on the hyperbolic plane. This study is simplified by the fact that this component of the Brownian motion satisfies a linear Fokker–Planck equation, for which the conditions of existence of H-theorems are well-known, as we will see below. However, even in this simple case, this study reveals an interesting link between the constant drift parameter of the Brownian motion and the non-extensivity parameter q associated with the entropy h_q such that an H_q-theorem exists.

2. The classical entropic approach to the linear Fokker–Planck equation

2.1. General approach

In the case of a system described by a univariate p.d.f. $f_t(x)$ that satisfies the linear Fokker–Planck equation

$$\frac{\partial f_t(x)}{\partial t} = -\frac{\partial}{\partial x}(K(x) f_t(x)) + \frac{\partial^2}{\partial x^2}(Q(x) f_t(x)), \tag{2.1}$$

where $K(x)$ is the drift function and $Q(x)$ the diffusion function, it can be shown (see for example Ref. [4]) that the relative entropy (or Kullback–Leibler divergence)

$$h(f_t \parallel g_t) = \int_{\mathbb{R}} f_t(x) \log \frac{f_t(x)}{g_t(x)} \mathrm{d}x \tag{2.2}$$

between two any solutions of (2.1) decreases to 0 with time. More precisely, it holds that

$$\frac{\partial}{\partial t} h(f_t \parallel g_t) = -\int_{\mathbb{R}} Q(x) f_t(x) \left(\frac{\partial}{\partial x} \log \frac{f_t(x)}{g_t(x)}\right)^2 \mathrm{d}x \le 0 \tag{2.3}$$

since the diffusion function is assumed positive. Thus the relative entropy is decreasing and bounded, so that the limit distributions f_∞ and g_∞ satisfy

$$\int_{\mathbb{R}} Q(x) f_\infty(x) \left(\frac{\partial}{\partial x} \log \frac{f_\infty(x)}{g_\infty(x)}\right)^2 \mathrm{d}x = 0,$$

which implies that f_∞ and g_∞ coincide. This proves the unicity of a stationary solution of the Fokker–Planck equation (2.1).

We note that the quantity

$$I(f \parallel g) = \int_{\mathbb{R}} f(x) \left(\frac{\partial}{\partial x} \log \frac{f(x)}{g(x)}\right)^2 \mathrm{d}x \tag{2.4}$$

is nothing but the relative Fisher information [5, eq. (174)] between the p.d.f.s f and g; the integral in (2.3) is a weighted version of this relative Fisher information, with the diffusion function $Q(x)$ as the positive weighting function.

In order to deduce an H-theorem from this result, we denote as g_∞ the stationary solution to (2.1), assuming that it exists. We then remark that the relative entropy between any solution f_t of (2.1) and the stationary solution g_∞ is related to the Shannon entropy of f_t and to the cross-entropy between f_t and g_∞ as

$$h(f_t \parallel g_\infty) = -h(f_t) - \int_{\mathbb{R}} f_t(x) \log g_\infty(x)\, \mathrm{d}x. \tag{2.5}$$

Thus, provided that the cross-entropy between the solution f_t and the asymptotic solution g_∞

$$\int_{\mathbb{R}} f_t(x) \log g_\infty(x)\, \mathrm{d}x \tag{2.6}$$

is constant over time, we deduce from identity (2.5) that, since the relative entropy decreases to 0, the Shannon entropy of the solution f_t increases with time to its maximum value.[1]

[1] We remark that, in statistical physics, the relative entropy $h(f_t \parallel g_\infty)$ coincides with the free energy.

Let us denote by X_t a random process following the p.d.f. f_t, solution of the Fokker–Planck equation (2.1): there is no reason why its cross-entropy (2.6) should be constant in time. However, let us show that there exists a proper scaling factor $a_t > 0$ such that the rescaled process $\tilde{X}_t = a_t X_t$ satisfies an H-theorem. We use the following steps:

Step 1: *Invariance of the relative entropy by invertible transformation*

If f and g are the distributions of two random variables X and Y respectively, and if T is any *invertible* mapping from $\mathbb{R}$ to $\mathbb{R}$ then the distributions $\tilde{f}$ and $\tilde{g}$ of the transformed random variables $\tilde{X} = T(X)$ and $\tilde{Y} = T(Y)$ respectively satisfy

$$h\left(\tilde{f} \parallel \tilde{g}\right) = h\left(f \parallel g\right). \tag{2.7}$$

This remarkable identity is in fact the case of equality of the *data processing inequality* (see for example Ref. [6])

$$h\left(\tilde{f} \parallel \tilde{g}\right) \leq h\left(f \parallel g\right)$$

that holds for *any* – possibly non-invertible – transformation T: equality is reached if and only if the transformation T is invertible.

Step 2: *Computation of the scaling factor*

By the scaling transformation,

$$\tilde{f}_t(y) = \frac{1}{|a_t|} f_t\left(\frac{y}{|a_t|}\right), \qquad \tilde{g}_\infty(y) = \frac{1}{|a_t|} g_\infty\left(\frac{y}{|a_t|}\right)$$

so that the cross-entropy of the scaled variables can be computed as

$$\int \tilde{f}_t(y) \log \tilde{g}_\infty(y)\, \mathrm{d}y = -\log|a_t| + \int f_t(x) \log g_\infty(x)\, \mathrm{d}x.$$

Thus this cross-entropy can be fixed to a constant value, say $\log K$ with $K > 0$, by choosing

$$a_t = K e^{\int_{\mathbb{R}} f_t(x) \log g_\infty(x) \mathrm{d}x}. \tag{2.8}$$

Step 3: *An H-theorem for the scaled random process*

From step 1, we deduce that the relative entropy of the scaled processes is decreasing with time. From step 2, we deduce that choice (2.8) ensures a constant cross-entropy. We conclude as follows.

Theorem 1. *If the relative entropy $h(f_t \parallel g_\infty)$ is decreasing with time and if the condition $\int_{\mathbb{R}} f_t(x) \log g_\infty(x)\, \mathrm{d}x > -\infty$ holds $\forall t > 0$ then the entropy of the scaled process $\tilde{X}_t = a_t X_t$ is increasing in time provided that the scaling factor a_t is chosen as in* (2.8)*.*

We note that the result of Theorem 1 can be extended to the multivariate context: assuming that we start from a multivariate Fokker–Planck equation in $\mathbb{R}^n$ with a stationary solution $g_\infty(\mathbf{x})$, the relative entropy

$$h\left(f_t \parallel g_\infty\right)$$

is decreasing over time. Then we can consider a matrix scaling

$$\tilde{X}_t = A_t X_t$$

of the process with distribution f_t. Interestingly, the scaling of the cross-entropy reads

$$\int_{\mathbb{R}^n} \tilde{f}_t(\mathbf{y}) \log \tilde{g}_\infty(\mathbf{y})\, \mathrm{d}\mathbf{y} = -\log|\det A_t| + \int_{\mathbb{R}^n} f_t(\mathbf{x}) \log g_\infty(\mathbf{x})\, \mathrm{d}\mathbf{x}$$

so that the cross-entropy can be fixed to a constant value assuming a condition on the determinant only of the matrix A_t, namely

$$\det A_t = K e^{\int_{\mathbb{R}^n} f_t(\mathbf{x}) \log g_\infty(\mathbf{x}) \mathrm{d}\mathbf{x}}.$$

This means that the matrix A_t can be chosen as

$$A_t = K e^{\int_{\mathbb{R}^n} f_t(\mathbf{x}) \log g_\infty(\mathbf{x}) \mathrm{d}\mathbf{x}} I_n$$

where I_n is the identity matrix in $\mathbb{R}^n$, so that only a scalar scaling is required.

2.2. The x-component of the Brownian motion in the Poincaré half-upper plane

In Ref. [7], Comtet and Monthus derived the differential equation satisfied by the x-component X_t of the Brownian motion with constant drift μ and constant diffusion constant D in the Poincaré half-upper plane representation of the hyperbolic plane,

$$\frac{\partial}{\partial t} f_t(x) = D\frac{\partial}{\partial x}\left[\left(1+x^2\right)\frac{\partial}{\partial x} f_t(x) + (2\mu+1)\, x f_t(x)\right].$$

This is a linear Fokker–Planck equation; in the notations of (2.1), the diffusion function is quadratic and positive, $Q(x) = D\left(1+x^2\right)$ and the drift function is linear, $K(x) = D(1-2\mu)x$. Moreover, for a positive drift μ, the asymptotic solution reads

$$g_\infty(x) = A_\mu \left(1+x^2\right)^{-\mu-\frac{1}{2}} \tag{2.9}$$

with normalization constant $A_\mu = \frac{\Gamma\left(\mu+\frac{1}{2}\right)}{\Gamma(\mu)\Gamma\left(\frac{1}{2}\right)}$.

In order to deduce an H-theorem for this random process, we simply need to check the condition $\int f_t(x)\log g_\infty(x)\,\mathrm{d}x > -\infty$ i.e. $\int f_t(x)\log\left(1+x^2\right)\mathrm{d}x > -\infty$, which is an immediate consequence of the fact that $\int f_t(x)\log\left(1+x^2\right)\mathrm{d}x \geq 0$. We deduce the following result.

Theorem 2. *The Shannon entropy of the x-component, scaled according to (2.8), of the Brownian motion on the Poincaré half-upper plane increases over time.*

3. A generalization to the Tsallis entropies

3.1. General approach

The monotone behavior of the relative Shannon entropy between any two solutions f_t and g_t of the linear Fokker–Planck equation (2.1) has been extended by Shiino [8] and simultaneously by Borland et al. [9,10] to the case of the relative Tsallis entropy, defined as

$$h_q(f_t \parallel g_t) = \frac{1}{q-1}\left(\int_{\mathbb{R}} f_t^q(x)\, g_t^{1-q}(x)\,\mathrm{d}x - 1\right).$$

More precisely, the derivative with respect to time of this relative entropy satisfies[2]

$$\frac{\partial}{\partial t} h_q(f_t \parallel g_t) = -q\int_{\mathbb{R}} Q(x) f_t(x)\left(\frac{f_t(x)}{g_t(x)}\right)^{q-1}\left(\frac{\partial}{\partial x}\log\frac{f_t(x)}{g_t(x)}\right)^2 \mathrm{d}x \leq 0, \quad \forall q > 0. \tag{3.1}$$

We note that this inequality holds for any positive value of q and simplifies to (2.3) as $q \to 1$. Moreover, the right-hand side integral in (3.1) can be considered as a q-version of the relative Fisher information that appears in (2.3).

In the non-extensive context, an identity such as (2.5), that relates linearly the relative entropy to the entropy, does not hold in general; as a consequence, a generalization of Theorem 1 can not be derived. However, in the particular case where the asymptotic distribution g_∞ is a q-Gaussian distribution – which is exactly the case for the Brownian motion on the hyperbolic plane – an extension of Theorem 1 to the Tsallis entropy can be provided.

3.2. The x-component of the Brownian motion in the Poincaré half-upper plane

In the special case of the x-component X_t of the Brownian motion in the Poincaré half-upper plane, the stationary solution (2.9) is the q-Gaussian

$$g_\infty(x) = C_{q_*}\left(1+x^2\right)^{\frac{1}{1-q_*}}$$

for some specific value $q = q_*$ of the non-extensivity parameter such that $\frac{1}{1-q_*} = -\mu - \frac{1}{2}$ or

$$q_* = \frac{2\mu+3}{2\mu+1}, \quad 1 < q_* < 3. \tag{3.2}$$

[2] The proof of this result is omitted in Ref. [8]; we provide it in the annex for the interested reader.

The normalization constant C_{q_*} is equal to

$$C_{q_*} = \frac{\Gamma\left(\frac{1}{q_*-1}\right)}{\Gamma\left(\frac{1}{q_*-1}-\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}.$$

It can be easily shown that the identity analogous to (2.5) reads

$$h_{q_*}\left(f_t \parallel g_\infty\right) = -C_{q_*}^{1-q_*} h_{q_*}\left(f_t\right) - \frac{C_{q_*}^{1-q_*}}{1-q_*}\int x^2 f_t^{q_*}(x)\,\mathrm{d}x + \beta_{q_*}\left(C_{q_*}\right) \tag{3.3}$$

with

$$\beta_{q_*}(C) = \frac{C^{1-q_*}-1}{q_*-1}.$$

It is important to note that identity (3.3) holds only for the choice $q = q_*$.

As in Part II, we show now that an H_q theorem holds for a properly scaled version of the process X_t. We follow the three steps as above:

- Step 1: It can be checked that the q-relative entropy is also invariant by any invertible transformation of both processes
- Step 2: The scaling factor a_t is computed by remarking that the integral in (3.3) scales as

  $$\int y^2 \tilde{f}_t^{q_*}(y)\,\mathrm{d}y = a_t^{3-q_*}\int x^2 \tilde{f}_t^{q_*}(x)\,\mathrm{d}x$$

  so that it can be assigned a fixed value $K > 0$ over time by choosing the scaling factor according to

  $$a_t^{3-q_*} = \frac{K}{\int x^2 \tilde{f}_t^{q_*}(x)\,\mathrm{d}x} \tag{3.4}$$

  which defines uniquely the scaling constant a_t for $1 < q_* < 3$.
- Step 3: As in step 3 of Part II, we conclude as follows.

Theorem 3. *With $q = q_*$ as in (3.2), the q-entropy of the scaled Brownian motion, normalized according to (3.4), with constant drift μ in the Poincaré half-upper plane is increasing with time.*

4. Numerical illustration

The following figures depict ten realizations of a discretized version of this Brownian motion, with parameters $D = 0.01$, $m = 0$ and $\mu = 1$ in Fig. 1 while $D = 0.01$, $m = 3$ and $\mu = 7$ in Fig. 2. In both cases, the process starts from the point $x = 0, y = 1$ and the superimposed thick curve is the asymptotic probability density of its x-component. Without external drift (Fig. 1), the random process wanders for a long time far away from the real axis before "falling" on it as $t \to +\infty$. The distribution of the asymptotic "landing points" on the real axis is thus very wide, in fact a Lorentz distribution with infinite variance. With an external drift (Fig. 2), the process is forced to walk in the direction of the real axis, so that the landing points are more concentrated around 0, which is reflected by their narrow distribution, a q-Gaussian distribution with variance $\sigma^2 = 0.2$.

5. The Brownian motion in the unit disk

Another representation of the hyperbolic space is the unit disk $\mathbb{D}$ with the metric in polar coordinates

$$ds^2 = \frac{4}{\left(1-r^2\right)^2}\left(dr^2 + r^2 d\theta^2\right).$$

There is a conformal mapping between the Poincaré upper half-plane $\mathbb{H} = \{z = x + iy,\ y > 0\}$ and the unit disk $\mathbb{D} = \{w \in \mathbb{C},\ |w| = 1\}$ defined as

$$w = \frac{iz+1}{z+i}.$$

Comtet and Monthus show in Ref. [7] that the density of the radial component of the Brownian motion in the unit disk representation of the hyperbolic space converges to $\delta(r-1)$ as $t \to +\infty$. Having no explicit Fokker–Planck for the radial part θ_t of this process, we were unable to prove a corresponding H-theorem. However, we show here that we can use the maximum entropy approach to derive the asymptotic distribution of the angular part θ_t, using the following result:

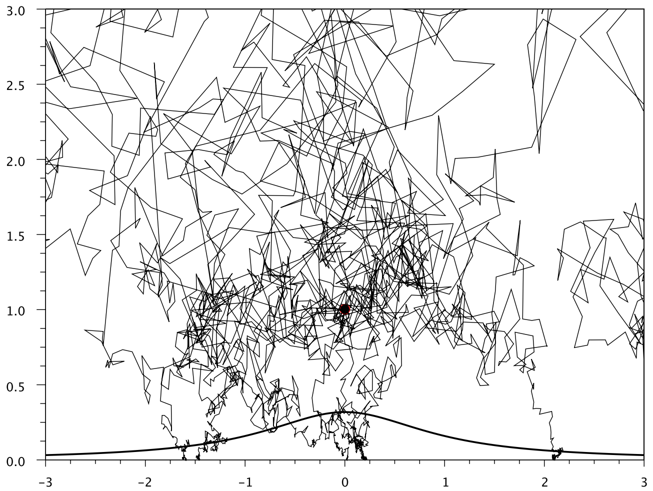

Fig. 1. Ten realizations of a discretized version of the Brownian motion on the hyperbolic plane, with parameters $D = 0.01$, $m = 0$ and $\mu = 1$.

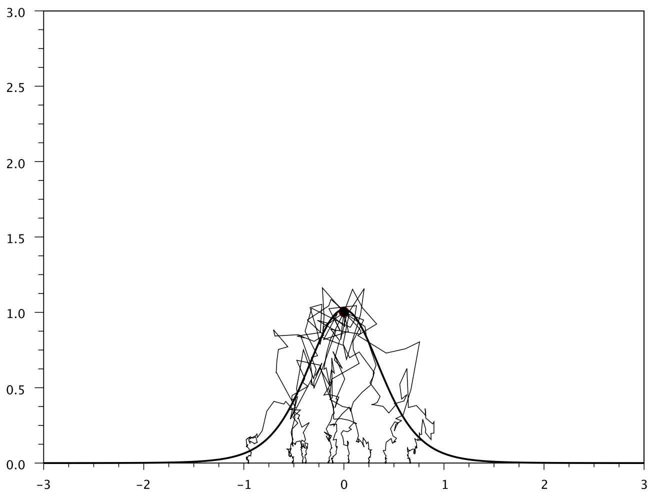

Fig. 2. Ten realizations of a discretized version of the Brownian motion on the hyperbolic plane, with parameters $D = 0.01$, $m = 3$ and $\mu = 7$.

Theorem. *If the random variable X has maximum entropy under the logarithmic constraint* $E\log\left(1+X^2\right) = \gamma$, *then the random variable*

$$\tilde{X} = \frac{X}{\sqrt{1+X^2}}$$

has maximum entropy under the constraint $E\log\left(1-\tilde{X}^2\right) = -\gamma$.

More precisely, if the p.d.f. of X *reads*

$$f_X\left(x\right) = \frac{\Gamma\left(\mu+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\mu\right)}\left(1+x^2\right)^{-\mu-\frac{1}{2}}, \quad x \in \mathbb{R}$$

then $E\log\left(1+X^2\right) = \psi\left(\mu+\frac{1}{2}\right) - \psi\left(\mu\right)$ *and the p.d.f. of* $\tilde{X}$ *reads*

$$f_{\tilde{X}}\left(\tilde{x}\right) = \begin{cases} \dfrac{\Gamma\left(\mu+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\mu\right)}\left(1-\tilde{x}^2\right)^{\mu-1} & -1 \le \tilde{x} \le +1 \\ 0 & \textit{else} \end{cases}$$

with $E\log\left(1-\tilde{X}^2\right) = -E\log\left(1+X^2\right) = \psi\left(\mu\right) - \psi\left(\mu+\frac{1}{2}\right)$ *and where* Ψ *is the digamma function.*

Since in the asymptotic regime, $f_r(r) = \delta(r-1)$, the conformal mapping becomes

$$\cos\theta = \frac{2X}{1+X^2}, \qquad \sin\theta = \frac{X^2-1}{X^2+1},$$

it can be easily verified that

$$\cos\left(\frac{\theta}{2} - \frac{\pi}{4}\right) = \frac{X}{\sqrt{1+X^2}} = \tilde{X}$$

so that a simple change of variable yields the probability density of θ

$$\frac{1}{2^{\mu+\frac{1}{2}}} \frac{\Gamma\left(\mu+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma(\mu)} (1-\sin\theta)^{\mu-\frac{1}{2}}$$

as obtained in Ref. [7].

The asymptotic distribution of the angular part has thus maximum Tsallis entropy with parameter $\tilde{q}$ such that

$$\tilde{q} = \frac{2\mu-3}{2\mu-1} < 1.$$

6. Conclusion

We have shown how to use the monotonicity of the Shannon or Tsallis relative entropies to deduce an H-theorem for the Shannon and Tsallis entropy, first in the general case and then in the case of the Brownian motion on the hyperbolic plane. Three remarkable results have been observed in this study: first, the natural drift induced by the negative constant curvature of the hyperbolic plane transforms the asymptotically Gaussian distribution of the usual Brownian motion on the plane to the Cauchy distribution, which belongs to the extended family of Tsallis distributions. Secondly, the addition of a constant positive external drift transforms this Cauchy behavior to a q-Gaussian behavior with non-extensivity parameter directly related to the value of this drift. This appearance of non-extensive distributions in the context of an underlying curved space remains to be linked to physically relevant experiments and data. Finally, the Tsallis distributions with $q > 1$ on the Poincaré half upper-plane realization of the hyperbolic plane transform, via the conformal mapping, into Tsallis distributions with $q < 1$ on the unit disk realization of the hyperbolic plane.

7. Annex: proof of Shiino's result

Following Shiino's notations, we consider

$$D_q(f_t \parallel g_t) = \int_{\mathbb{R}} f_t(x)^q g_t(x)^{1-q} \mathrm{d}x$$

and omit the time and space variables for readability. The time derivative reads

$$\frac{\partial}{\partial t} D_q(f \parallel g) = q \int f^{q-1} \left(\frac{\partial}{\partial t} f\right) g^{1-q} + (1-q) \int f^q g^{-q} \left(\frac{\partial}{\partial t} g\right).$$

Since f and g are both solutions of the Fokker–Planck equation (2.1), we deduce

$$\frac{\partial}{\partial t} D_q(f \parallel g) = q \int \left(\frac{f}{g}\right)^{q-1} \left(-\frac{\partial}{\partial x}(Kf) + \frac{\partial^2}{\partial x^2}(Qf)\right) + (1-q) \int \left(\frac{f}{g}\right)^q \left(-\frac{\partial}{\partial x}(Kg) + \frac{\partial^2}{\partial x^2}(Qg)\right).$$

The integrals with the drift function $K(x)$ can be integrated by parts, yielding respectively

$$-q \int \left(\frac{f}{g}\right)^{q-1} \frac{\partial}{\partial x}(Kf) = +q(q-1) \int \left(\frac{f}{g}\right)^{q-2} \frac{\partial}{\partial x}\left(\frac{f}{g}\right) Kf$$

and

$$\begin{aligned} -(1-q) \int \left(\frac{f}{g}\right)^q \frac{\partial}{\partial x}(Kg) &= q(1-q) \int \left(\frac{f}{g}\right)^{q-1} \frac{\partial}{\partial x}\left(\frac{f}{g}\right) Kg \\ &= q(1-q) \int \left(\frac{f}{g}\right)^{q-2} \frac{\partial}{\partial x}\left(\frac{f}{g}\right) Kf \end{aligned}$$

so that their sum vanishes.

The integrals with the diffusion function $Q(x)$ are also integrated by parts according respectively to

$$q\int\left(\frac{f}{g}\right)^{q-1}\frac{\partial^2}{\partial x^2}(Qf)=-q\int\frac{\partial}{\partial x}\left(\frac{f}{g}\right)^{q-1}\frac{\partial}{\partial x}(Qf)$$
$$=-q(q-1)\int\left(\frac{f}{g}\right)^{q-2}\frac{\partial}{\partial x}\left(\frac{f}{g}\right)\frac{\partial}{\partial x}(Qf)=-q(q-1)\int\left(\frac{f}{g}\right)^{q-1}\frac{\partial}{\partial x}\left(\log\frac{f}{g}\right)\frac{\partial}{\partial x}(Qf)$$
$$=-q(q-1)\int f\left(\frac{f}{g}\right)^{q-1}\frac{\partial}{\partial x}\left(\log\frac{f}{g}\right)\frac{\partial Q}{\partial x}-q(q-1)\int Q\left(\frac{f}{g}\right)^{q-1}\frac{\partial}{\partial x}\left(\log\frac{f}{g}\right)\frac{\partial f}{\partial x}$$

and

$$(1-q)\int\left(\frac{f}{g}\right)^{q}\frac{\partial^2}{\partial x^2}(Qg)=-(1-q)\int\frac{\partial}{\partial x}\left(\frac{f}{g}\right)^{q}\frac{\partial}{\partial x}(Qg)$$
$$=-(1-q)q\int\left(\frac{f}{g}\right)^{q-1}\frac{\partial}{\partial x}\left(\frac{f}{g}\right)\frac{\partial}{\partial x}(Qg)$$
$$=-(1-q)q\int g\left(\frac{f}{g}\right)^{q-1}\frac{\partial}{\partial x}\left(\frac{f}{g}\right)\frac{\partial Q}{\partial x}-(1-q)q\int Q\left(\frac{f}{g}\right)^{q-1}\frac{\partial}{\partial x}\left(\frac{f}{g}\right)\frac{\partial g}{\partial x}.$$

Their sum consists in an integral with the diffusion function Q

$$(q-1)q\int Q\left(\frac{f}{g}\right)^{q-1}\left(\frac{\partial}{\partial x}\left(\frac{f}{g}\right)\frac{\partial g}{\partial x}-\frac{\partial}{\partial x}\left(\log\frac{f}{g}\right)\frac{\partial f}{\partial x}\right)$$

and an integral with its derivative

$$q(1-q)\int\left(\frac{f}{g}\right)^{q-1}\left(f\frac{\partial}{\partial x}\left(\log\frac{f}{g}\right)-g\frac{\partial}{\partial x}\left(\frac{f}{g}\right)\right)\frac{\partial Q}{\partial x}.$$

The second integral is easily seen to vanish, while the first one can be simplified to

$$(q-1)q\int Q\left(\frac{f}{g}\right)^{q-1}\left(\frac{\partial}{\partial x}\left(\frac{f}{g}\right)\frac{\partial g}{\partial x}-\frac{\partial}{\partial x}\left(\log\frac{f}{g}\right)\frac{\partial f}{\partial x}\right)=-q(q-1)\int Qf\left(\frac{f}{g}\right)^{q-1}\left(\frac{\partial}{\partial x}\log\frac{f}{g}\right)^2.$$

Thus the Tsallis divergence $h_q(f\parallel g)=\frac{1}{q-1}D_q(f\parallel g)$ verifies the stated equality.

Acknowledgment

The authors would like to thank Pr. A.R. Plastino for his invaluable help on this paper.

References

[1] C. Tsallis, Possible generalization of Boltzmann–Gibbs statistics, J. Stat. Phys. 52 (1–2) (1988) 479.
[2] A.R. Plastino, A. Plastino, Non-extensive statistical mechanics and generalized Fokker–Planck equation, Physica A Stat. Theoretical. Phys. 222 (1–4) (1995) 347–354.
[3] C. Tsallis, D.J. Bukman, Anomalous diffusion in the presence of external forces: exact time-dependent solutions and their thermostatistical basis, Phys. Rev. E 54 (3) (1996) R2197–R2200.
[4] H. Risken, The Fokker–Planck Equation: Methods of Solution and Applications, Springer-Verlag, 1989.
[5] C. Villani, A Review of Mathematical Topics in Collisional Kinetic Theory, in: Handbook of Mathematical Fluid Dynamics, 1, North-Holland, 2002, pp. 71–74. 2.
[6] T.M. Cover, J.A. Thomas, Elements of Information Theory, second ed. Wiley-Interscience, Hoboken, NJ, 2006.
[7] A. Comtet, C. Monthus, Diffusion in a one-dimensional random medium and hyperbolic Brownian motion, J. Phys. A Math. Gen. 29 (7) (1996) 1331.
[8] M. Shiino, H-theorem with generalized relative entropies and the Tsallis statistics, J. Phys. Soc. Japan 67 (11) (1998) 3658–3660.
[9] L. Borland, A.R. Plastino, C. Tsallis, Information gain within nonextensive thermostatistics, J. Math. Phys. 39 (1998) 6490.
[10] L. Borland, A.R. Plastino, C. Tsallis, Erratum: information gain within generalized thermostatistics, J. Math. Phys. 39 (1998) 6490; J. Math. Phys. 40 (1999) 2196.