Electronic Journal of Differential Equations, Vol. 2023 (2023), No. 12, pp. 1–11.
ISSN: 1072-6691. URL: https://ejde.math.txstate.edu or https://ejde.math.unt.edu

DE BRUIJN IDENTITIES IN DIFFERENT MARKOVIAN CHANNELS

HASSAN EMAMIRAD, ARNAUD ROUGIREL

Communicated by Jerome A. Goldstein

Abstract. De Bruijn's identity in information theory states that if u is the solution of the heat equation, then the time derivative of the Shannon entropy for this solution is equal to the amount of Fisher information at u. In this article, we show how this identity changes if we replace the heat channel by the Fokker Planck, or passing from Fokker Planck to Ornstein-Uhlenbeck channels. Through these passages we investigate the different properties of these solutions. We exclusively dissect different properties of Ornstein-Uhlenbeck semigroup given by the Mehler formula expression.

1. Introduction

Let the probability triplet be $(\Omega, \mathscr{F}, \mu)$, where Ω is the sample space, $\mathscr{F}$ a σ-algebra and μ is a probability measure. Let $\varphi : \mathbb{R}^n \to \mathbb{R}_+$ with $\int_{\mathbb{R}^n} \varphi(x)\, d\mu(x) = 1$ be a density function of a random variable X. We extend the function $x \in (0, \infty) \mapsto x \ln x$ by 0 at $x = 0$ and assume that $\varphi \ln \varphi \in L^1(\mathbb{R}^n, d\mu)$. This function defines so called *Shannon's entropy* or *Boltzmann H function*

$$H(\varphi) := -\int_{\mathbb{R}^n} \varphi(x) \ln \varphi(x)\, d\mu(x).$$

If we place ourselves in a dynamical system at time t, this entropy will be written as $H(\varphi(\cdot, t))$. In this context the term of de Bruijn identity was pointed by Stam [5], which was communicated to him by Prof. de Bruijn and indicate that Shannon's entropy decreases in time when u runs through a Gaussian channel with rate equal to Fisher information.

In mathematical information theory, the Fisher information is a way of measuring the amount of information that an observable random variable X carries about the distribution that models φ. For example by taking $X : \Omega \times [0, \infty) \mapsto \mathbb{R}$ a Markovian process and defining the density function $\varphi(\cdot, t)$ of the random variable valued in $\mathbb{R}_+$ can be considered as a probability distribution depending on $t \in \mathbb{R}_+$. Formally, the Fisher information is the variance of the score, which is the gradient of the log-likelihood function which is logarithm of $\varphi(\cdot, t)$. This is the fundamental concept

2020 *Mathematics Subject Classification.* 94A17, 94A40.
Key words and phrases. De Bruijn identity; Gaussian; Ornstein-Uhlenbeck channels; Fokker Planck; relative Fisher information; Kullback-Leibler divergence.

Submitted October 27, 2022. Published February 6, 2023.

in information theory as it is indicated in the seminal book of Cover and Thomas [2].

The Fisher information is defined by the following quantity in $[0,\infty]$

$$\begin{aligned}\mathcal{I}(\varphi) &= \int_{\mathbb{R}^n} \Big(\frac{\partial}{\partial x}\ln\varphi(x,t)\Big)^2 \varphi(x,t)\,d\mu(x)\\ &= \int_{\mathbb{R}^n} \Big(\frac{\partial}{\partial x}\varphi(x,t)\Big)^2 \varphi(x,t)^{-1}\,d\mu(x).\end{aligned}$$

The most well-known example in a dynamical system is the unitary heat equation

$$\frac{\partial u}{\partial t} = \Delta u, \quad t>0,\ x\in\mathbb{R}^n, \tag{1.1}$$

with an initial condition $u_0(\cdot) = u(\cdot,0) \in L^1_+(\mathbb{R}^n,\, d\mu)$ where $d\mu = dx$ is the standard Lebesgue measure with

$$\int_{\mathbb{R}^n} u_0(x)dx = 1. \tag{1.2}$$

As a straightforward consequence of the explicit expression of the solution in term of the Green function $G(t,x,y) := (4\pi t)^{-n/2}\exp\big(-\frac{|x-y|^2}{4t}\big)$,

$$u(x,t) = \int_{\mathbb{R}^n} u_0(y)G(t,x,y)\,dy, \tag{1.3}$$

it follows that $u(x,t)$ is positive for any $(x,t)\in\mathbb{R}^n\times\mathbb{R}_+$. The mass conservation implies also that

$$\int_{\mathbb{R}^n} u(x,t)dx = \int_{\mathbb{R}^n} u_0(x)dx = 1. \tag{1.4}$$

In this case we say that u runs through the Gaussian or heat channel.

The noticeable connection between Fisher information and Shannon's entropy is the so-called De Bruijn relation [2] (see also [1, 4]). That is, if u runs through the Gaussian channel, then

$$\frac{d}{dt}H(u) = \mathcal{I}(u) \tag{1.5}$$

(For completeness, the proof is provided in the Appendix).

Before trying to deduce the De Bruijn relation for Fokker-Planck equation

$$\frac{\partial}{\partial t}v(x,t) = \frac{\partial^2}{\partial x^2}(v(x,t)) - \frac{\partial}{\partial x}(xv(x,t)),$$

in the next section we show how one can derive the Fokker-Planck equation from the heat equation and in the section 3 we establish the De Bruijn relation for Fokker-Planck equation.

In section 4 we show how the Ornstein-Uhlenbeck equation $\frac{\partial w}{\partial t} = \frac{\partial^2}{\partial y^2}w(x,t) - x\frac{\partial}{\partial x}w(x,t)$ can be deduce from Fokker-Planck equation. For this equation the mass conservation takes place in $L^1(\mathbb{R}^n,\, d\mu)$, where $d\mu$ is the Gaussian measure. The section 5 is devoted to Ornstein-Uhlenbeck semigroup in which we prove the hypercontractivity of this semigroup which deduces the Chapman-Kolmogorov relation for its kernel. In the section 6 we recover the De Bruijn relation for this channel.

Finally in section 7 we prove the De Bruijn identity for relative Fisher information and Kullback-Leibler divergence which is already discussed in [7].

2. Relationship between Fokker-Planck and heat equation

In general the Fokker-Planck equation in one dimensional space reads as

$$\frac{\partial}{\partial t}v(y,\tau)=\frac{\partial^2}{\partial y^2}(g^2(y,t)v(y,\tau))-\frac{\partial}{\partial y}(f(y)v(y,\tau)), \tag{2.1}$$

where $f(y,t)$ and $g(y,\tau)$ can be arbitrary positive functions define on $\mathbb{R}_y\times\mathbb{R}_+$. In this section we take $g=1$ and $f(y)=-y$ and the following theorem gives an explicit expression of the solution of (2.1).

Theorem 2.1. *If u is the solution of the heat equation* (1.1) *with initial condition u_0 satisfying* (1.2), *then $v(y,\tau)=e^\tau u(e^\tau y,(e^{2\tau}-1)/2)$, that is*

$$v(y,\tau)=\frac{e^\tau}{\sqrt{2\pi(e^{2\tau}-1)}}\int_{\mathbb{R}}v_0(\xi)e^{-\frac{(e^\tau y-\xi)^2}{2(e^{2\tau}-1)}}\,d\xi, \tag{2.2}$$

satisfies the Fokker-Planck equation

$$\frac{\partial}{\partial\tau}v(y,\tau)=\frac{\partial^2}{\partial y^2}v(y,\tau)+v(y,\tau)+y\frac{\partial}{\partial y}v(y,\tau). \tag{2.3}$$

Proof. First we remark that for $t=(e^{2r}-1)/2=0$, we have $e^{2\tau}-1=0$, so τ should be equal zero. Hence

$$u_0=v(y,0)=v_0. \tag{2.4}$$

If we replace $u(x,t)$ by its explicit expression (1.3) and we find (2.2). Now we have to verify that this function bears out (2.3). Indeed, let us denote

$$A(\tau)=\frac{e^\tau}{\sqrt{2\pi(e^{2\tau}-1)}},\quad B:=B(\tau,y,\xi)=\frac{(e^\tau y-\xi)^2}{2(e^{2\tau}-1)},$$
$$I(\tau,y)=\int_{\mathbb{R}}v_0(\xi)e^{-B(\tau,y,\xi)}\,d\xi,$$

such that (2.2) can be expressed as

$$v(y,\tau)=A(\tau)I(\tau,y).$$

We remark that

$$\frac{\partial}{\partial\tau}A(\tau)=\underbrace{\frac{e^\tau}{\sqrt{2\pi(e^{2\tau}-1)}}}_{=A_1(\tau)}-\underbrace{\frac{e^{3\tau}}{\sqrt{2\pi(e^{2\tau}-1)^3}}}_{=A_2(\tau)},$$
$$\frac{\partial}{\partial\tau}B(\tau,y,\xi)=\underbrace{\frac{ye^\tau(e^\tau y-\xi)}{e^{2\tau}-1}}_{=B_1(\tau,y,\xi)}-\underbrace{\frac{e^{2\tau}(e^\tau y-\xi)^2}{(e^{2\tau}-1)^2}}_{=B_2(\tau,y,\xi)},$$
$$\frac{\partial}{\partial y}B(\tau,y,\xi)=\underbrace{\frac{e^\tau(e^\tau y-\xi)}{e^{2\tau}-1}}_{=B_3(\tau,y,\xi)},$$
$$\frac{\partial}{\partial y}B_3(\tau,y,\xi)=\frac{e^{2\tau}}{e^{2\tau}-1}.$$

Hence,

$$\frac{\partial}{\partial y}v(y,\tau)=-A(\tau)\int_{\mathbb{R}}v_0(\xi)B_3(\tau,y,\xi)e^{-B(\tau,y,\xi)}\,d\xi$$

and

$$\begin{aligned}\frac{\partial^2}{\partial y^2} v(y,\tau) &= A(\tau)\Big(-\int_{\mathbb{R}} v_0(\xi)\frac{e^{2\tau}}{e^{2\tau}-1} e^{-B(\tau,y,\xi)}\, d\xi \\ &\quad + \int_{\mathbb{R}} v_0(\xi)(B_3(\tau,y,\xi))^2 e^{-B(\tau,y,\xi)}\, d\xi\Big) \\ &= -A_2(\tau)I(\tau,y) + A(\tau)\int_{\mathbb{R}} u_0(\xi)(B_3(\tau,y,\xi))^2 e^{-B(\tau,y,\xi)}\, d\xi \\ &= -A_2(\tau)I(\tau,y) + A(\tau)\int_{\mathbb{R}} u_0(\xi)B_2(\tau,y,\xi) e^{-B(\tau,y,\xi)}\, d\xi.\end{aligned}$$

Consequently,

$$\begin{aligned}&\frac{\partial}{\partial \tau} v(y,\tau) \\ &= \underbrace{A_1(\tau)I(\tau,y)}_{=v(y,\tau)} - A_2(\tau)I(\tau,y) \\ &\quad \underbrace{-A(\tau)\Big(\int_{\mathbb{R}} u_0(\xi)B_1(\tau,y,\xi)e^{-B(\tau,y,\xi)}\, d\xi}_{=y\frac{\partial v(y,\tau)}{\partial y}} - \int_{\mathbb{R}} v_0(\xi)B_2(\tau,y,\xi)e^{-B(\tau,y,\xi)}\, d\xi\Big) \\ &= v(y,\tau) + y\frac{\partial v(y,\tau)}{\partial y} + \frac{\partial^2}{\partial y^2} v(y,\tau).\end{aligned}$$

□

3. De Bruijn identity in Fokker-Planck channels

In this section we will use the above Theorem for obtaining an identity similar to (1.5).

Theorem 3.1. *Assume that $t \geq 0$, the Fisher information is defined by the positive quantity*

$$\mathcal{I}(v(\cdot,t)) = \int_{\mathbb{R}} \Big(\frac{\partial}{\partial y}\ln v(y,t)\Big)^2 v(y,t)\, dy\,.$$

and Shannon's entropy is

$$H(v(\cdot,t)) = -\int_{\mathbb{R}} v(y,t)\ln v(y,t) dy. \tag{3.1}$$

Then $v := v(y,\tau)$ the solution of

$$\begin{gathered}\frac{d}{d\tau} v(y,\tau) = \frac{\partial^2}{\partial y^2} v(y,\tau) + v(y,\tau) + y\frac{\partial}{\partial y} v(y,\tau), \\ v(y,0) = v_0(y), \quad \text{with } \int_{\mathbb{R}} v_0(y)\, dy = 1\end{gathered} \tag{3.2}$$

will satisfy the modified De Bruijn identity in Fokker-Planck channels

$$\frac{d}{d\tau} H(v) = \mathcal{I}(v) - 1. \tag{3.3}$$

Proof. According (2.2),

$$\lim_{|y|\to\infty} |yv| = 0, \tag{3.4}$$

and

$$\begin{aligned}\frac{d}{d\tau}H(v(\cdot,\tau)) &= -\int_{\mathbb{R}}\Big(\frac{\partial^2}{\partial y^2}v(y,\tau)+\frac{\partial}{\partial y}(yv(y,\tau))\Big)\ln v(y,\tau)\,dy\\ &\quad -\int_{\mathbb{R}}\Big(\frac{\partial^2}{\partial y^2}v(y,\tau)+\frac{\partial}{\partial y}(yv(y,\tau))\Big)\,dy\\ &= \int_{\mathbb{R}}\Big((\frac{\partial v}{\partial y})^2/v+y\frac{\partial v}{\partial y}(y,\tau)\Big)\,dy\\ &= \mathcal{I}(v)-\int_{\mathbb{R}}v(y,\tau)\,dy=\mathcal{I}(v)-1,\end{aligned}$$

according to (3.4), Fokker-Planck equation has the mass conservation property and for $t=0$ the mass is equal 1, that is

$$\int_{\mathbb{R}}v(y,\tau)\,dy=1\quad\text{for all }\ \tau\in\mathbb{R}_+. \tag{3.5}$$

□

4. Relationship between Fokker-Planck and Ornstein-Uhlenbeck equation

Let $v_\infty(y):=\frac{1}{\sqrt{2\pi}}e^{-y^2/2}$ be the unique stationary solution of Fokker-Planck equation (2.3) with $\int_{\mathbb{R}}v_\infty(y)dy=1$, and denote by $d\mu=v_\infty\,dx$ the Gaussian measure. Now, if we transform the expression (2.2) to the form

$$v(y,\tau)=\frac{1}{\sqrt{2\pi(1-e^{-2\tau})}}\int_{\mathbb{R}}v_0(\xi)e^{-\frac{(y-\xi/e^{\tau})^2}{2(1-e^{-2\tau})}}\,d\xi, \tag{4.1}$$

we notice that $v(y,\tau)\to v_\infty(y)$ as $\tau\to\infty$.

Theorem 4.1. *Assume that v is the solution of Fokker-Planck equation* (2.3)*. Then $w=v/v_\infty$ satisfies the Ornstein-Uhlenbeck equation*

$$\frac{\partial w}{\partial\tau}=\frac{\partial^2}{\partial y^2}w(y,\tau)-y\frac{\partial}{\partial y}w(y,\tau), \tag{4.2}$$

with initial data $w_0(y)=u_0(y)/v_\infty$, $u_0\in L^1_+(\mathbb{R}^n,dy)$ and $\int_{\mathbb{R}^n}u_0(y)\,dy=1$.

Proof. First note that by (2.4), $w_0=w(y,0)=v_0(y)/v_\infty(y)=u_0(y)/v_\infty(y)$. Consequently, by (3.5) we have $\int_{\mathbb{R}}w(y,\tau)\,d\mu=1$.

Now, knowing that v is the solution of (3.1) we can write

$$\frac{\partial w}{\partial\tau}=\Big(\frac{\partial^2 v}{\partial y^2}+v+y\frac{\partial v}{\partial y}\Big)/v_\infty. \tag{4.3}$$

On the other hand, by $\frac{\partial v_\infty}{\partial y}=-yv_\infty$, we have

$$\begin{aligned}\frac{\partial w}{\partial y}&=v_\infty^{-1}(\frac{\partial v}{\partial y}+yv),\\ \frac{\partial^2 w}{\partial y^2}&=v_\infty^{-1}\Big(\frac{\partial^2 v}{\partial y^2}+2y\frac{\partial v}{\partial y}+v+y^2v\Big).\end{aligned}$$

By insert these expressions in (4.3) we obtain (4.2) . □

5. Mehler formula and Ornstein-Uhlenbeck semigroup

On $\mathbb{R}^n$, let μ_n be the canonical Gaussian measure with density $(2\pi)^{-n/2}e^{(-|x|^2/2)}$ with respect to the Lebesgue measure dx. With this measure we consider the Banach space $L^p(\mathbb{R}^n, d\mu_n)$, $0 \le p < \infty$ with the norm $\|f\|_p = \big(\int_{\mathbb{R}^n} |f|^p\, d\mu_n\big)^{1/p}$ on which we can define the Ornstein-Uhlenbeck semigroup P_t by mean of *Mehler formula*

$$P_t f(x) := \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}}y)\, d\mu_n(y), \quad \text{for } f \in L^p(\mathbb{R}^n, d\mu_n). \tag{5.1}$$

By taking $\alpha_t = e^{-t}$, $\beta_t = \sqrt{1-e^{-2t}}$ and making a change of variable $z = \alpha_t x + \beta_t y$ in this formula we obtain

$$P_t f(x) = (2\pi\beta_t^2)^{-n/2} \int_{\mathbb{R}^n} f(z) \exp\big(-|z-\alpha_t x|^2/2(\beta_t^2)\big)\, dz\,. \tag{5.2}$$

This equality implies that the Gaussian measure $d\mu_n$ is invariant for P_t, that is

$$\int_{\mathbb{R}^n} P_t f(x)\, d\mu_n(x) = \int_{\mathbb{R}^n} f(x)\, d\mu_n(x) \quad \text{for all } f \in L^p(\mathbb{R}^n, d\mu_n). \tag{5.3}$$

To show this we need the following lemma.

Lemma 5.1. *For $c_1, c_2 \ge 0$, $c_1 + c_2 \ne 0$ and $a, b \in \mathbb{R}^n$, we have*

$$\int_{\mathbb{R}^n} e^{-c_1|a-z|^2 - c_2|z-b|^2} dz = \Big(\frac{\pi}{c_1+c_2}\Big)^{n/2} \exp\Big(-\frac{c_1c_2}{c_1+c_2}|a-b|^2\Big). \tag{5.4}$$

Proof. This follows from the unity of the Gaussian measure that for any $p \in \mathbb{R}^n$ and $\alpha > 0$,

$$\int_{\mathbb{R}^n} \exp\big(-\alpha|x-p|^2\big)\, dx = \Big(\frac{\pi}{\alpha}\Big)^{n/2},$$

which implies

$$\int_{\mathbb{R}^n} \exp\big(-\alpha|x|^2 + 2\langle \alpha p, x\rangle\big)\, dy = (\frac{\pi}{\alpha})^{n/2} \exp(\alpha|p|^2).$$

Let $\alpha = c_1 + c_2$ and $p = (c_1a + c_2b)/\alpha$, then

$$\begin{aligned} &\int_{\mathbb{R}^n} \exp\big(-(c_1+c_2)|x|^2 + 2\langle (c_1a + c_2b), x\rangle\big)\, dx \\ &= \Big(\frac{\pi}{c_1+c_2}\Big)^{n/2} \exp\Big(\frac{|(c_1a+c_2b)|^2}{c_1+c_2}\Big). \end{aligned}$$

Now since

$$\begin{aligned} &\int_{\mathbb{R}^n} \exp\big(-c_1|a-x|^2 - c_2|x-b|^2\big)\, dx \\ &= \Big(\frac{\pi}{c_1+c_2}\Big)^{n/2} \exp\Big(\frac{|(c_1a+c_2b)|^2}{c_1+c_2} - c_1|a|^2 - c_2|b|^2\Big) \\ &= \Big(\frac{\pi}{c_1+c_2}\Big)^{n/2} \exp\Big(\frac{-c_1c_2|a-b|^2}{c_1+c_2}\Big), \end{aligned}$$

we obtain (5.4). $\square$

For proving (5.3) we write

$$\int_{\mathbb{R}^n} P_t f(x)\, d\mu_n(x) = (4\pi^2\beta_t^2)^{-n/2} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(z)\exp\Big(-\frac{|z-\alpha_t x|^2+|\beta_t x|^2}{2\beta_t^2}\Big)\,dx\,dz.$$

Taking $c_1 = \frac{\alpha_t^2}{2\beta_t^2}, c_2 = \frac{1}{2}, a = \alpha_t^{-1}$ and $b = 0$ in Lemma 5.1, since $\alpha_t^2 + \beta_t^2 = 1$ we obtain

$$\int_{\mathbb{R}^n} P_t f(x)\, d\mu_n(x) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} f(z)\exp\Big(-\frac{|z|^2}{2}\Big)dz.$$

which is (5.3).

Theorem 5.2. *On $X_p := L^p(\mathbb{R}^n, d\mu_n)$ the operator T_t defines a hypercontractive semigroup; that is,*

(i) $\|T_t\|_p \le \|f\|_q$, *for all $p \ge q > 1$ such that $p-1 \le e^{2t}(q-1)$;*
(ii) $\lim_{t\to 0}\|T_t f - f\|_p = 0$, *for all $f \in X_p$;*
(iii) $T_t T_s = T_{t+s}$, *for all $(t,s) \in \mathbb{R}_+^2$.*

Proof. (i) Since the constant function 1is in X_p, and the Mehler formula $P_t 1 = 1$, the equality (5.3) implies that P_t is doubly Markovian in the sense of Nelson (see [3]). In the same paper (Theorem 2), Nelson shows that an operator which is doubly Markovian is hypercontractive in the sense of item (i).

(ii) Since $\alpha_t^2 + \beta_t^2 = 1$, the vectors (α_t, β_t) and $(1,0)$ are both on the unit circle and $(\alpha_t, \beta_t) \to (1,0)$ as $t \to 0$. For any continuous bounded function f (taking e.g. $f \in \mathscr{S}$),

$$f(\alpha_t x + \beta_t y) - f(x) = f\Big((\alpha_t,\beta_t)\begin{pmatrix}x\\ y\end{pmatrix}\Big) - f\Big((1,0)\begin{pmatrix}x\\ y\end{pmatrix}\Big) \to 0$$

as $t \to 0$. Furthermore, if $M = \sup_{x\in\mathbb{R}}|f(x)|$, then $|f(\alpha_t x + \beta_t y) - f(x)| \le 2M$ in $L^1(\mathbb{R}^n, d\mu_n)$. Thus according to Lebesgue's dominated converence theorem

$$\int_{\mathbb{R}^n} |f(\alpha_t x + \beta_t y) - f(x)|\, d\mu_n(y) \to 0.$$

Hence

$$\begin{aligned}\|T_t f - f\|_p^p &= \int_{\mathbb{R}^n} |P_t f(x) - f(x)|^p\, d\mu_n(x)\\ &= \int_{\mathbb{R}^n}\Big|\int_{\mathbb{R}^n}[f(\alpha_t x + \beta_t y) - f(x)]\, d\mu_n(y)\Big|^p d\mu_n(x) \to 0\end{aligned}$$

Since the Schwartz space $\mathscr{S}$ being dense in X_p, this implies item (ii).

(iii) Taking the expression of the Ornstein-Uhlenbeck semigroup P_t (5.2),

$$\begin{aligned}P_t P_s f(x) &= (2\pi\beta_t^2)^{-n/2}\int_{\mathbb{R}^n} P_s f(z)\exp(-|z-\alpha_t x|^2/2(\beta_t^2)dz\\ &= (2\pi\beta_t\beta_s)^{-n}\int_{\mathbb{R}^n} f(y)\underbrace{\int_{\mathbb{R}^n}\exp\Big(-\frac{|y-\alpha_s z|^2}{2\beta_s^2} - \frac{|z-\alpha_t x|^2}{2\beta_t^2}\Big)dz}_{=A}\,dy\,.\end{aligned}$$

To simplify the expression A we will use the Lemma 5.1. Let

$$A = \int_{\mathbb{R}^n}\exp\Big(-\frac{\alpha_s^2}{2\beta_s^2}|\alpha_s^{-1}y - z|^2 - \frac{1}{2\beta_t^2}|z-\alpha_t x|^2\Big)dz.$$

Comparing this with (5.4), we obtain

$$c_1 = \frac{\alpha_s^2}{2\beta_s^2}, \quad c_2 = \frac{1}{2\beta_t^2}, \quad a = \alpha_s^{-1} y, \quad b = \alpha_t x.$$

Hence,

$$\begin{aligned} A &= \Big(\frac{\pi}{\frac{\alpha_s^2}{2\beta_s^2} + \frac{1}{2\beta_t^2}}\Big)^{n/2} \exp\Big(-\frac{\frac{\alpha_s^2}{4\beta_s^2\beta_t^2}}{\frac{\alpha_s^2}{2\beta_s^2} + \frac{1}{2\beta_t^2}} |\alpha_s^{-1} y - \alpha_t x|^2\Big) \\ &= \Big(\frac{2\pi(1-e^{-2t})(1-e^{-2s})}{1-e^{-2(t+s)}}\Big)^{n/2} \exp\Big(-\frac{|y-e^{-(t+s)}x|^2}{2(1-e^{-2(t+s)})}\Big). \end{aligned} \tag{5.5}$$

Replacing the expression of A in (5.5) we find that

$$P_t P_s f(x) = (2\pi(1-e^{-2(t+s)}))^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(y) \exp\Big(-\frac{|y-e^{-(t+s)}x|^2}{2(1-e^{-2(t+s)})}\Big)\, dy = P_{t+s} f(x).$$

□

Remark 5.3. From (iii) of the above Theorem one can deduce the Chapman-Kolmogorov formula

$$\int_{\mathbb{R}^n} K(x,y,t)K(y,z,s)\, dy = K(x,z,t+s) \quad \text{for all } x \in \mathbb{R}^n, t, s > 0, \tag{5.6}$$

where $K(x,y,t)$ is the heat kernel of Ornstein-Uhlenbeck semigroup, that is

$$P_t f(x) := \int_{\mathbb{R}^n} K(x,y,t) f(y)\, dy \quad \text{for all } f \in L^p(\mathbb{R}^n,\, d\mu_n).$$

From (5.2) it follows that

$$K(x,y,t) = (2\pi(1-t^{-2t}))^{-n/2} \exp\Big(-\frac{|e^{-t}x-y|^2}{2(1-t^{-2t})}\Big).$$

Hence

$$\begin{aligned} P_t P_s f(x) &= \int_{\mathbb{R}^n} K(x,y,t) P_s f(y)\, dy \\ &= \iint_{\mathbb{R}^{2n}} K(x,y,t)K(y,z,s) f(z) dz dy \\ &= P_{t+s} f(x) = \int_{\mathbb{R}^n} K(x,z,t+s) f(z) dz \end{aligned}$$

Since this identity holds for all $f \in L^p(\mathbb{R}^n,\, d\mu_n)$, we deduce formula (5.6) for μ_n-a.e. $z \in \mathbb{R}^n$. The equality holds on $\mathbb{R}^n$ by continuity of the left and right hand side with respect to z.

6. De Bruijn identity in Ornstein-Uhlenbeck channels

In this section we work in $L^1(\mathbb{R}, \mu)$ which is a Lebesgue space with the Gaussian measure $\mu := \mu_1$. In this space $\int_{\mathbb{R}} w(\cdot,\tau)\, d\mu = 1$. If we define the entropy by

$$H_\mu(w(\cdot,\tau)) := -\int_{\mathbb{R}} w(y,\tau) \ln w(y,\tau)\, d\mu(y) \tag{6.1}$$

and the Fisher information by

$$\mathcal{I}_\mu(w(\cdot,\tau)) := \int_{\mathbb{R}} w(y,\tau)\Big(\frac{\partial}{\partial y} \ln w(y,\tau)\Big)^2\, d\mu(y), \tag{6.2}$$

then the De Bruijn identity in Ornstein-Uhlenbeck channels reads as follows.

Theorem 6.1. *Assume that $t \geq 0$, then*

$$\frac{d}{d\tau} H_\mu(w(\cdot,\tau)) = \mathcal{I}_\mu(w(\cdot,\tau)), \tag{6.3}$$

where w is the solution of Ornstein-Uhlenbeck equation (4.2).

Proof. For the proof we will use

$$\frac{d}{dy}\mu(y) = \frac{-y}{\sqrt{2\pi}} e^{-y^2/2} = -y v_\infty \,. \tag{6.4}$$

The derivative of (6.1) with respect to τ reads

$$\begin{aligned}
&\frac{d}{d\tau} H_\mu(w(y,\tau)) \\
&= -\int_{\mathbb{R}} \Big(\frac{d}{d\tau} w(y,\tau) \ln w(y,\tau)\Big)\, d\mu \\
&= -\int_{\mathbb{R}} \Big(\frac{\partial^2}{\partial y^2} w(y,\tau) - y\frac{\partial}{\partial y} w(y,\tau)\Big) \ln w(y,\tau)\, d\mu \underbrace{- \frac{d}{d\tau}\int_{\mathbb{R}} w(y,\tau)\, d\mu}_{=0 \text{ by } (3.5)} \\
&= \int_{\mathbb{R}} w(y,\tau)^{-1}\Big(\frac{\partial}{\partial y} w(y,\tau)\Big)^2 d\mu(y) - \int_{\mathbb{R}} \frac{\partial}{\partial y} w(y,\tau)(\ln w(y,\tau)) y v_\infty dy \\
&\quad + \int_{\mathbb{R}} y\frac{\partial}{\partial y} w(y,\tau)(\ln w(y,\tau))\, d\mu = \mathcal{I}_\mu(w(y,\tau)).
\end{aligned}$$

□

7. De Bruijn identity for relative Fisher information and Kullback-Leibler divergence

Let φ and ψ, be two distribution functions for two random variables X and Y. The relative Fisher information with respect to a is defined by

$$\mathcal{I}_a(\varphi||\psi) := \int_{\mathbb{R}} \varphi(x)\Big(\frac{\partial}{\partial x} \ln \frac{\varphi(x)}{\psi(x)}\Big)^2 a(x) dx.$$

We define the Kullback-Leibler divergence which can be interpreted as the relative entropy between φ and ψ by

$$D_{KL}(\varphi||\psi) := \int_{\mathbb{R}} \varphi(x) \ln \frac{\varphi(x)}{\psi(x)} dx.$$

(see [6]).

The following result establishes that the relative entropy between any two solutions of (2.3) is always decreasing, with a rate given by the relative Fisher information:

Theorem 7.1. *Assume that $\varphi(x,t)$ and $\psi(x,t)$ two distinct solutions of the Fokker-Planck equation in its general form:*

$$\frac{\partial \phi}{\partial t}(x,t) = \frac{\partial^2}{\partial x^2}(a(x,t)\phi(x,t)) - \frac{\partial}{\partial x}(b(x,t)\phi(x,t)) \tag{7.1}$$

Then

$$\frac{d}{dt} D_{KL}(\varphi||\psi) = -\mathcal{I}_b(\varphi||\psi). \tag{7.2}$$

Proof. Let $(x,t) \in \mathbb{R} \times \mathbb{R}_+$ and assume $\varphi(\cdot,t), \psi(\cdot,t) \in H^2(\mathbb{R})$. By differentiating under the integral sign and using the chain rule, we have (For simplicity we write $\varphi(x)$ instead of $\varphi(x,t)$)

$$\begin{aligned}
&\frac{d}{dt} D_{KL}(\varphi||\psi) \\
&= \frac{d}{dt} \int_{\mathbb{R}} \varphi(x) \ln \frac{\varphi(x)}{\psi(x)} dx \\
&= \int_{\mathbb{R}} \frac{\partial}{\partial t}\varphi(x) \ln \frac{\varphi(x)}{\psi(x)} dx + \int_{\mathbb{R}} \varphi(x) \frac{\partial}{\partial t} \ln \varphi(x) dx - \int_{\mathbb{R}} \varphi(x) \frac{\partial}{\partial t} \ln \psi(x) dx \\
&= \int_{\mathbb{R}} \frac{\partial}{\partial t}\varphi(x) \ln \frac{\varphi(x)}{\psi(x)} dx + 0 - \int_{\mathbb{R}} \frac{\varphi(x)}{\psi(x)} \frac{\partial}{\partial t} \psi(x) dx.
\end{aligned} \tag{7.3}$$

By replacing $\frac{\partial}{\partial t}\varphi(x)$ in the Fokker-Planck equation (7.1) and using integration by parts we can write the first integral in (7.3) as

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{\partial}{\partial t}\varphi(x) \ln \frac{\varphi(x)}{\psi(x)} dx \\
&= \int_{\mathbb{R}} \Big(\frac{\partial^2}{\partial x^2} a(x)\varphi(x) - \frac{\partial}{\partial x} b(x)\varphi(x) \Big) \ln \frac{\varphi(x)}{\psi(x)} dx, \\
&= \int_{\mathbb{R}} \Big(a(x)\varphi(x) \frac{\partial^2}{\partial x^2} \ln \frac{\varphi(x)}{\psi(x)} + b(x)\varphi(x) \frac{\partial}{\partial x} \ln \frac{\varphi(x)}{\psi(x)} \Big) dx.
\end{aligned} \tag{7.4}$$

Since

$$\frac{\partial}{\partial x} \ln \frac{\varphi(x)}{\psi(x)} = \Big(\frac{\psi(x)}{\varphi(x)} \Big) \Big(\frac{\partial}{\partial x} \frac{\varphi(x)}{\psi(x)} \Big),$$

and

$$\frac{\partial^2}{\partial x^2} \ln \frac{\varphi(x)}{\psi(x)} = \Big(\frac{\psi(x)}{\varphi(x)} \Big) \Big(\frac{\partial^2}{\partial x^2} \frac{\varphi(x)}{\psi(x)} \Big) - \Big(\frac{\partial}{\partial x} \ln \frac{\varphi(x)}{\psi(x)} \Big)^2,$$

by replacing these relations in (7.4) and using integration by parts and the Fokker-Planck equation (7.1) for $\psi(x,t)$ we find that

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{\partial}{\partial t}\varphi(x) \ln \frac{\varphi(x)}{\psi(x)} dx \\
&= \int_{\mathbb{R}} a(x)\varphi(x) \Big(\frac{\psi(x)}{\varphi(x)} \frac{\partial^2}{\partial x^2} \frac{\varphi(x)}{\psi(x)} - \Big(\frac{\partial}{\partial x} \ln \frac{\varphi(x)}{\psi(x)} \Big)^2 \Big) dx + \int_{\mathbb{R}} b(x)\psi(x) \frac{\partial}{\partial x} \frac{\varphi(x)}{\psi(x)} dx.
\end{aligned}$$

Thus,

$$\int_{\mathbb{R}} \frac{\partial}{\partial t}\varphi(x) \ln \frac{\varphi(x)}{\psi(x)} dx = -\mathcal{I}_a(\varphi||\psi) + \int_{\mathbb{R}} a(x)\psi(x) \frac{\partial^2}{\partial x^2} \frac{\varphi(x)}{\psi(x)} + b(x)\psi(x) \frac{\partial}{\partial x} \frac{\varphi(x)}{\psi(x)} dx$$

Plugging this relation into equation (7.3) and take into account that $\psi(x)$ is also the solution of the Fokker Planck equation (7.1) we conclude that

$$\frac{d}{dt} D_{KL}(\varphi||\psi) = -\mathcal{I}_a(\varphi||\psi),$$

as desired. □

8. Appendix

Here we give the proof of de Bruijn identity for a function running through a Gaussian channel.

Let $H(u) = \int_{\mathbb{R}} u(x,\cdot)\ln u(x,\cdot)dx$ be Shannon's entropy, then

$$\begin{aligned}\frac{d}{dt}H(u) &= \frac{d}{dt}\int_{\mathbb{R}} u(x,t)\ln u(x,t)dx\\ &= \int_{\mathbb{R}}(\frac{d}{dt}u)(\ln u) + u\frac{d}{dt}\ln u(x,t))dx\\ &= \int_{\mathbb{R}}\Delta u \ln u dx + \int_{\mathbb{R}}\frac{d}{dt}u(x,t)dx\\ &= -\int_{\mathbb{R}}\left((\nabla u)^2/u\right)dx + \frac{d}{dt}\int_{\mathbb{R}} u(x,t)dx.\end{aligned}$$

Since $\int_{\mathbb{R}} u(x,t)dx = 1$, we obtain $\frac{d}{dt}H(u) = \mathcal{I}(u)$.

References

[1] J.-Ph. Bartier, A. Blanchet, J. Dolbeault, M. Escobedo; Improved intermediate asymptotics for the heat equation. *Appl. Math. Lett.*, *24* (2011), 76–81.

[2] T. M. Cover, J. A. Thomas. Elements of information theory. Second edition. Wiley-Interscience , Hoboken, NJ, 2006

[3] E. Nelson; The free Markoff field. *J. Funct. Anal.*, *12* (1973), 211–227.

[4] G. Toscani; The fractional Fisher information and the central limit theorem for stable laws. *Ric. Mat.*, *65*(2016), 71–91.

[5] A. J. Stam; Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Contr.*, *2* (1959), 101–112.

[6] J. Voigt; Stochastic Operators, Information, and Entropy *Commun. Math. Phys.*, *81* (1981) 31–38.

[7] A. Wibisono, V. Jog, P-L. Loh; Information and estimation in Fokker-Plank channels. 2017 IEEE Inter. Symp. on Infor. Theory, arXiv 1702.03656v1.

Hassan Emamirad

Laboratoire de Mathématiques, Université de Poitiers, teleport 2, BP 179, 86960 Chasseneuil du Poitou, Cedex, France

Email address: emamirad@math.univ-poitiers.fr

Arnaud Rougirel

Laboratoire de Mathématiques, Université de Poitiers, teleport 2, BP 179, 86960 Chasseneuil du Poitou, Cedex, France

Email address: rougirel@math.univ-poitiers.fr