Similarly, we derive

$$\mathrm{DIF}_0 = p(p-1) \prod_{j=1}^{(n-2)/4} p^{p^{j-1}} \prod_{j=1}^{(n-2)/4} (p^m)^{p^{j-1}}$$

for an odd $n/2$, since $\sharp H_{n/2}^a = p(p-1)$ from (7). Therefore, we obtain DIF_k for an odd $n/2$. Q.E.D.

Example 3: We are interested in the difficulty DIF_0 for 2^m-phase bent sequences with period $2^n - 1$, since signals with 2^m-phase have often been used in some communications. We compute DIF_0 from Theorem 3 as shown in Table I. In order to get large difficulty, we need a bent sequence set with long period or many phases.

V. Conclusion

We have tried to construct a DS-SSMA system, in which the difficulty of wiretapping from interception increases and the worst case error probability decreases as much as possible. Bent sequences with optimal periodic correlation properties and high linear span have been applied as spreading sequences corresponding to secret keys.

The difficulty of wiretapping has been defined, and derived for a p^m-phase bent sequence set. If a bent sequence set possesses long period or many phases, the DS-SSMA system seems to be able to protect information data from interception, as shown in Example 3.

If we allow the bent functions to take real values, the difficulty seems infinite.

References

[1] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications*. Rockville, MD: Computer Science, 1985, vol. 1, ch 5.

[2] H. Imai, R. Kohno, and T. Matsumoto, “Information security of spread spectrum system,” *IEICE Trans. Commun.*, vol. E74, no. 3, Mar. 1991.

[3] T. Kohda and A. Tsuneda, “Pseudonoize sequences by chaotic nonlinear maps and their correlation properties,” *IEICE Trans. Commu.*, vol. E76-B, no. 8, Aug. 1993.

[4] G. H. Bateni and C. D. McGillem, “A chaotic direct-sequence spread-spectrum communication system,” *IEEE Trans. Commum.*, vol. 42, Feb./Mar./Apr. 1994.

[5] T. Kasami, “Weight distribution formula for some classses of cyclic codes,” Coordinated Sci. Lab., Univ. Illinois, Urbana, Tech. Rep. R-285, Apr. 1966.

[6] D. V. Sarwate and M. B. Pursley, “Crosscorrelation properties of pseudorandom and related sequences,” *Proc. IEEE*, vol. 68, pp. 593–619, May 1980.

[7] E. L. Key, “An analysis of the structure and complexity of nonlinear binary sequence generator,” *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 732–736, Nov. 1976.

[8] R. Gold, “Optimum binary sequences for spread spectrum multiplexing,” *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 619–321, Oct. 1967.

[9] M. B. Pursley, “Performance evaluation for phase-coded spread-spectrum multiple-access communication—Part 1,” *IEEE Trans. Commun.*, vol. COM-24, pp. 795–799, Aug. 1977.

[10] K. H. A. Karkkainen, “Mean-square cross-correlation as a performance measue for spreading code families,” in *Proc. ISSSTA’92*, 1992, pp. 147–150.

[11] L. R. Welch, “Lower bound on the maximum cross correlation of signals,” *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 397–399, May 1974.

[12] J. D. Olsen, “Nonlinear Binary Sequences with Asymptotically Optimum Periodic Cross-Correlation,” Ph.D. dissertation, Elec. Eng. Dept., Univ. Southern Calif., Los Angeles, 1977.

[13] P. V. Kumar, “On Bent Sequences and Generalized Bent Functions,” Ph.D. dissertation, Elec. Eng. Dept., Univ. Southern Calif., Los Angeles, Aug. 1983.

[14] S. Matsufuji and K. Imamura, “Real-valued bent function and its application to the design of balanced quadriphase sequences with optimal correlation properties,” in *Lecture Notes in Computer Science*. Berlin, Germany: Springer Verlag, Mar. 1991, vol. 508, pp. 113–121.

[15] ——, “Balanced quadriphase sequences with optimal correlation properties constructed by real-valued bent functions,” *IEEE Trans. Inform. Theory*, vol. 39, pp. 305–310, Jan. 1993.

A Short Proof of the “Concavity of Entropy Power”

Cédric Villani

***Abstract*—We give a simple proof of the “concavity of entropy power.”**

***Index Terms*—Entropy power, Fisher information, heat semigroup.**

I. Introduction

Let f be a probability measure on $\mathbb{R}^n$. We define the action of the heat semigroup $(P_t)_{t\geq 0}$ on f, by the solution of the partial differential equation

$$\frac{\partial}{\partial t} P_t f = \Delta(P_t f).$$

Equivalently, $P_t f$ is the convolution of f with the n-dimensional Gaussian density having mean vector 0 and covariance matrix $2tI_n$, where I_n is the identity matrix. The “concavity of entropy power” theorem states that

$$\frac{d^2}{dt^2} N(P_t f) \leq 0. \tag{1}$$

Here

$$N(f) = \frac{e^{\frac{2H(f)}{n}}}{2\pi e} \qquad H(f) = -\int_{\mathbb{R}^n} f \log f.$$

The functional $N(f)$ is the so-called “entropy power” of f, as introduced by Shannon, while $H(f)$ is Shannon’s entropy functional (which coincides with Boltzmann’s entropy up to a change of sign). The normalizing factor $2\pi e$ is nonessential and we mention it only to stick to the conventions of Shannon.

Inequality (1) is due to Costa [4]. Later, Dembo [5], [6] simplified the proof, by an argument based on the Blachman–Stam inequality [3]

$$\frac{1}{I(f*g)} \leq \frac{1}{I(f)} + \frac{1}{I(g)}.$$

Here f and g are two arbitrary probability densities, and

$$I(f) = \int_{\mathbb{R}^n} \frac{\left|\nabla f\right|^2}{f}$$

stands for the Fisher information of f. Actually, Dembo proved inequality (1) in the equivalent form

$$J(f) \geq \frac{I(f)^2}{n} \qquad J(f) = -\frac{1}{2} \left.\frac{d}{dt}\right|_{t=0} I(P_t f). \tag{2}$$

Manuscript received September 22, 1999; revised January 13, 2000.

The author is with DMA, Ecole Normale Supérieure, 75230 Paris, Cedex 05, France (e-mail: villani@dma.ens.fr).

Communicated by I. Csiszár, Associate Editor for Shannon Theory.
Publisher Item Identifier S 0018-9448(00)04639-3.

Using basic considerations on the heat equation, like the continuity of $H(P_t f)$ with respect to f (when f varies in a class s.t. $H(f)$ stays bounded), it is sufficient to prove (1), or equivalently (2), for a very smooth initial datum f, with fast decay at infinity. In order not to worry about logarithms, we may also impose that $|\log f(x)| \leq C(1+|x|^2)$ for some constant C. The general case will follow by density.

Our goal here is to give a direct proof of (2), in a strengthened form, with an exact error term. Our proof relies on the following lemma, well-known in certain circles.

Lemma: Let f be a smooth, rapidly decaying probability density, such that $\log f$ has growth at most polynomial at infinity. Then

$$J(f) = \sum_{ij} \int_{\mathbf{R}^n} f[\partial_{ij}(\log f)]^2 = \sum_{ij} \int_{\mathbf{R}^n} f\left[\frac{\partial_{ij} f}{f} - \frac{\partial_i f \partial_j f}{f^2}\right]^2.$$

Here the summation is taken over all indices $1 \leq i \leq n, 1 \leq j \leq n$. This computation (or actually a variant of it) was performed by McKean [7] in one dimension of space, and easily generalized by Toscani [8] to the n-dimensional case. But this lemma is also a particular case of the identities of Bakry and Emery [2], established through the so-called Γ_2 calculus as part of their famous work on logarithmic Sobolev inequalities and hypercontractive diffusions. For the sake of completeness, we give here a simple proof which is inspired from Bakry and Emery.

Proof of Lemma: Write the Fisher information in the form

$$I(f) = \int f\left|\nabla(\log f)\right|^2$$

so that, by differentiation under the integral sign,

$$\left.\frac{d}{dt}\right|_{t=0} I(P_t f) = \int \Delta f\left|\nabla(\log f)\right|^2 + 2\int f\nabla(\log f)\cdot\nabla\left(\frac{\Delta f}{f}\right). \tag{3}$$

We express $\Delta f/f$ in terms of $\log f$, thanks to the elementary identity

$$\frac{\Delta f}{f} = \Delta(\log f) + \left|\nabla(\log f)\right|^2$$

so that (3) becomes

$$\int \Delta f\left|\nabla(\log f)\right|^2 + 2\int f\nabla(\log f)\cdot\nabla\Delta(\log f) + 2\int f\nabla(\log f)\cdot\nabla\left|\nabla(\log f)\right|^2. \tag{4}$$

The first integral in (4) can, of course, be rewritten as

$$\int f\Delta\left|\nabla(\log f)\right|^2$$

while the third one is

$$2\int \nabla f\cdot\nabla\left|\nabla\log f\right|^2 = -2\int f\Delta\left|\nabla(\log f)\right|^2.$$

On the whole, (3) is equal to

$$\int f\left[2\nabla(\log f)\cdot\nabla\Delta(\log f) - \Delta|\nabla(\log f)|^2\right].$$

We conclude by the elementary identity (in which the reader may recognize a trivial particular case of Bochner's formula)

$$2\nabla u\cdot\nabla\Delta u - \Delta\left|\nabla u\right|^2 = -2\sum_{ij}(\partial_{ij}u)^2. \qquad \square$$

With this lemma at hand, the proof of (2) is almost immediate.

Proposition: Let f be a smooth, rapidly decaying probability density, such that $\log f$ has growth at most polynomial at infinity. Then

$$J(f) \geq \frac{I(f)^2}{n}.$$

Proof: Consider the nonnegative quantity

$$A(\lambda) = \sum_{ij}\int\left(\frac{\partial_{ij}f}{f} - \frac{\partial_i f\partial_j f}{f^2} + \lambda\delta_{ij}\right)^2 f$$

and expand this expression as a trinom in λ. Since

$$\int \partial_{ii} f = 0, \quad \sum_i \int \frac{(\partial_i f)^2}{f} = I(f), \qquad \int f = 1$$

we obtain

$$A(\lambda) = \sum_{ij}\int\left(\frac{\partial_{ij}f}{f} - \frac{\partial_i f\partial_j f}{f^2}\right)^2 f - 2\lambda I(f) + \lambda^2 n.$$

Now, the choice $\lambda = I(f)/n$ yields the *equality*

$$J(f) - \frac{I(f)^2}{n} = \sum_{ij}\int\left(\frac{\partial_{ij}f}{f} - \frac{\partial_i f\partial_j f}{f^2} + \frac{I(f)}{n}\delta_{ij}\right)^2 f \geq 0. \tag{5}$$

$\square$

Remarks:

1) It is easy to check that, at least under suitable smoothness assumptions, equality in (2) occurs if and only if f is an isotropic Gaussian.
2) As one of the referees pointed out, a proof of the Proposition in the same spirit as the above argument is implicit in the notes by D. Bakry [1, p. 103, remarks following the proof of Proposition 6.7]. Namely, applying the Cauchy–Schwarz inequality twice

$$\sum_{ij}\left[\partial_{ij}(\log f)\right]^2 \geq \frac{1}{n}\left[\Delta(\log f)\right]^2$$

$$\int f\left[\Delta(\log f)\right]^2 \geq \left(\int f\Delta(\log f)\right)^2 = I(f)^2$$

(where we used again $\int f = 1$). While this proof does not give any simple remainder term, one advantage is that—as again pointed out by the referee—it also works for Riemannian manifolds with nonnegative Ricci curvature.

References

[1] D. Bakry, "L'hypercontractivité et son utilization en théorie des semigroupes," in *Ecole d'été de Probabilités de Saint-Flour (Lecture Notes in Mathematics)*. Berlin, Germany: Springer-Verlag, 1994.

[2] D. Bakry and M. Emery, "Diffusions hypercontractives," in *Sém. Proba. XIX (Lecture Notes in Mathematics)*. Berlin, Germany: Springer-Verlag, 1985, pp. 177–206.

[3] N. M. Blachman, "The convolution inequality for entropy powers," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 267–271, Apr. 1965.

[4] M. Costa, "A new entropy power inequality," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 751–760, Nov. 1985.

[5] A. Dembo, "A simple proof of the concavity of the entropy power with respect to the variance of additive normal noise," *IEEE Trans. Inform. Theory*, vol. 35, pp. 887–888, July 1989.

[6] A. Dembo, T. Cover, and J. Thomas, "Information theoretic inequalities," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1501–1518, Nov. 1991.

[7] H. P. McJKean, Jr., "Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas," *Arch. Rat. Mech. Anal.*, vol. 21, pp. 343–367, 1966.

[8] G. Toscani, "Entropy production and the rate of convergence to equilibrium fr the Fokker-Planck equation," *Quart. Appl. Math.*, vol. 57, no. 3, pp. 521–541, 1999.