Hiroshima Math. J.
9 (1979), 547–554

On a Functional of Distribution Functions having Maximum at Gaussian Distribution Function

Hiroshi Murata

(Received January 19, 1979)

§ 1. Introduction

The entropy functional $H[f]$ is defined by

$$H[f] = -\int_{-\infty}^{\infty} f(x)\log f(x)\,dx, \qquad f\in\mathscr{D},$$

where $\mathscr{D}$ is the set of probability density functions f on R^1 with $\int f(x)|\log f(x)|dx < \infty$. Let $\mathscr{D}_1$ be the subset of $\mathscr{D}$ with $\int x^2 f(x)dx = 1$, and $g\in\mathscr{D}_1$ be the Gaussian density function with mean 0. Then Gibbs' lemma states that

$$H[f] \leqq H[g], \qquad f\in\mathscr{D}_1. \tag{1.1}$$

Consider a class of functionals $\tilde{H}[f]$ of the form

$$\tilde{H}[f] = \int_{-\infty}^{\infty} h(f(x))\,dx, \qquad f\in\mathscr{D}_1.$$

Under some regularity conditions on h, McKean[3] proved that if the inequality (1.1) holds with $H=\tilde{H}$, then $h(x)=c_1x+c_2x\log x$ ($c_2\leqq 0$).

Let $\mathscr{P}_1$ be the set of probability distribution functions with mean 0 and variance 1, and G be the Gaussian distribution function belonging to $\mathscr{P}_1$. Tanaka [6] considered the functional $\mathrm{e}[F]$ defined by

$$\mathrm{e}[F] = \inf\int_{R^2}|x-y|^2 dM(x, y), \qquad F\in\mathscr{P}_1,$$

where the infimum is taken over all two-dimensional distribution functions $M(x, y)$ whose marginals are F and G. It is known (see [6] or [4]) that

$$\begin{aligned}\mathrm{e}[F] &= \int_0^1 |F^{-1}(\alpha)-G^{-1}(\alpha)|^2 d\alpha \\ &= 2-2\Phi_0[F], \qquad \Phi_0[F] = \int_{-\infty}^{\infty} xG^{-1}(F(x))dF(x),\end{aligned}$$

where $F^{-1}(\alpha)$ is the right continuous inverse function of $F(x)$. It can be proved (in § 2) that

$$\Phi_0[F] = \int_{-\infty}^{\infty} g(G^{-1}(F(x)))dx. \tag{1.2}$$

On the other hand, it is obvious that $\mathfrak{e}[F]$ has the minimum at $F=G$, and therefore $\Phi_0[F]$ defined by (1.2) has the maximum at $F=G$. The main purpose of this paper is to prove that, along the same line as McKean [3], the functional Φ_0 is the only one which has the maximum at G among those functionals Φ of the form

$$\Phi[F] = \int_{-\infty}^{\infty} \varphi(F(x))dx, \qquad F \in \mathscr{P}_1. \tag{1.3}$$

Some regularity conditions on φ must be assumed, and the precise statement is as follows.

THEOREM A. *Let φ be a function on* $[0, 1]$, *and assume that*

$$\varphi \in C[0, 1] \cap C^1(0, 1) \quad \text{and} \quad \varphi(0) = \varphi(1) = 0, \tag{1.4a}$$

$$\varphi'(\alpha) = \begin{cases} O(\alpha^{-\delta}), & \alpha \downarrow 0 \\ O((1-\alpha)^{-\delta}), & \alpha \uparrow 1 \end{cases} \quad \text{for any} \quad \delta \in (0, 1). \tag{1.4b}$$

If the functional Φ defined by (1.3) *satisfies*

$$\Phi[F] \leqq \Phi[G], \qquad F \in \mathscr{P}_1, \tag{1.5}$$

and is normalized so that $\Phi[G]=1$, then $\Phi=\Phi_0$.

We also consider Boltzmann's problem for Kac's model of a Maxwellian gas. In this model the probability distribution function $F(t, x)$ of molecular speeds at time t is determined by

$$\frac{\partial F(t, x)}{\partial t} = \frac{1}{2\pi}\int_0^{2\pi} d\theta \int_{R^2} \mathbf{1}_{(-\infty,x]}(y\cos\theta + z\sin\theta)dF(t, y)dF(t, z) - F(t, x), \tag{1.6}$$

where $\mathbf{1}_{(-\infty,x]}$ is the indicator function of $(-\infty, x]$ and $dF(t,\cdot)$ is the probability measure corresponding to $F(t,\cdot)$, t being fixed. It was proved in [6] that the functional $\mathfrak{e}$ decreases along the solutions of (1.6), and therefore the functional Φ_0 increases along the solutions of (1.6). As the converse statement of this, we can prove the following theorem.

THEOREM B. *Let φ be a function on* $[0, 1]$ *satisfying* (1.4*a*) *and* (1.4*b*). *If the functional Φ defined by* (1.3) *increases with time along the solutions of* (1.6) *with initial distribution functions belonging to $\mathscr{P}_1$, then $\Phi=c\Phi_0$, $c \geqq 0$.*

The author wishes to express his thanks to Professor H. Tanaka for valuable advices and encouragements.

§ 2. Proof of Theorem A

Let $F\in\mathscr{P}_1$. Because the assumption (1.4) with $\delta=1/3$ and the estimates

$$\lim_{x\to\infty} x^2(1-F(x)) = \lim_{x\to-\infty} x^2F(x) = 0 \tag{2.1}$$

imply $\varphi(F(x))=o(|x|^{-4/3})$ as $|x|\to\infty$, the integral defining $\Phi[F]$ is absolutely convergent, that is, $\Phi[F]$ is well-defined.

First we prove that (1.2) holds. Because of the well-known estimates (for example, see [1; p175])

$$\begin{cases} 1-G(x)\sim \dfrac{g(x)}{x}, & x\to\infty, \\ G(x)\sim \dfrac{g(x)}{|x|}, & x\to-\infty, \end{cases} \tag{2.2}$$

we have

$$g(G^{-1}(\alpha))\sim\begin{cases} \alpha\sqrt{2\log\dfrac{1}{\alpha}}, & \alpha\downarrow 0, \\ (1-\alpha)\sqrt{2\log\dfrac{1}{1-\alpha}}, & \alpha\uparrow 1, \end{cases}$$

which combined with (2.1) implies that

$$\lim_{|x|\to\infty}|x|g(G^{-1}(F(x)))=0, \qquad F\in\mathscr{P}_1.$$

Integrating $\Phi_0[F]$ by parts and using $d[g(G^{-1}(F(x)))]=-G^{-1}(F(x))dF(x)$, we obtain

$$\int_{-\infty}^{\infty} xG^{-1}(F(x))dF(x)=\int_{-\infty}^{\infty} g(G^{-1}(F(x)))dx,$$

as was to be proved.

Now we proceed to the proof of Theorem A. In order to clarify our method, we perform some formal calculations; rigorous justifications of these will be given later.

Put

$$g_{m,t}(x)=\frac{1}{\sqrt{2\pi t}}\exp\left[-\frac{(x-m)^2}{2t}\right],$$

$$f_\varepsilon(x) = \sigma[(1-\varepsilon)g(\sigma x) + \varepsilon g_{m,t}(\sigma x)], \quad \sigma = \sqrt{1-\varepsilon+\varepsilon(t+m^2)-\varepsilon^2 m^2},$$

for $t \in (0, 1)$ and $m \in R^1$; let F_ε be the distribution function corresponding to the density function f_ε. Since F_ε has mean $\varepsilon m/\sigma$ and variance 1, we have

$$\text{(2.3)} \quad \Phi[F_\varepsilon] \equiv \int_{-\infty}^{\infty} \varphi(F_\varepsilon(x))dx = \int_{-\infty}^{\infty} \varphi\left((1-\varepsilon)G(\sigma x) + \varepsilon G\left(\frac{\sigma x - m}{\sqrt{t}}\right)\right)dx$$
$$= \Phi[\tilde{F}_\varepsilon] \leqq \Phi[G],$$

where $\tilde{F}_\varepsilon(x) = F_\varepsilon(x + \varepsilon m/\sigma) \in \mathscr{P}_1$. Therefore we obtain

$$\text{(2.4)} \quad 0 \leqq \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \{\Phi[G] - \Phi[F_\varepsilon]\}$$
$$= \int_{-\infty}^{\infty} \varphi'(G(x)) \left\{G(x) - \frac{t+m^2-1}{2} xg(x) - G\left(\frac{x-m}{\sqrt{t}}\right)\right\} dx.$$

Letting $t \downarrow 0$ in the above, we have

$$\text{(2.5)} \qquad 0 \leqq \int_{-\infty}^{\infty} \varphi'(G(x))\left\{G(x) - \frac{m^2-1}{2} xg(x) - \mathbf{1}_{[m,\infty)}(x)\right\} dx;$$

this must be the equality, because the integration of the right hand side of the above with respect to $g(m)dm$ yields

$$\int_{-\infty}^{\infty} dx\varphi'(G(x))\int_{-\infty}^{\infty} \left\{G(x) - \frac{m^2-1}{2} xg(x) - \mathbf{1}_{[m,\infty)}(x)\right\} g(m)dm = 0.$$

Differentiating this equality (2.5) in m, we have

$$\text{(2.6)} \qquad \varphi'(G(m)) = m \cdot \int_{-\infty}^{\infty} \varphi'(G(x))xg(x)dx,$$

and therefore $\varphi'(G(m)) = cm$. Since $\Phi[G] = 1$, $c = -1$ and hence $\Phi = \Phi_0$.

Proof of (2.4): Let $\delta \in (0, 1)$ be fixed, and put

$$A_\varepsilon(x) \equiv \varphi(G(x)) - \varphi\left((1-\varepsilon)G(\sigma x) + \varepsilon G\left(\frac{\sigma x - m}{\sqrt{t}}\right)\right).$$

Writing down $\frac{\partial}{\partial \varepsilon} A_\varepsilon(x)$ explicitly and then using the assumption (1.4) on φ, we see that there exists a positive constant c_1 depending upon δ such that the following estimate holds for all sufficiently large x:

$$\text{(2.7)} \quad \left|\frac{\partial}{\partial \varepsilon} A_\varepsilon(x)\right|$$
$$\leqq c_1 \left|1 - (1-\varepsilon)G(\sigma x) - \varepsilon G\left(\frac{\sigma x - m}{\sqrt{t}}\right)\right|^{-\delta} \times \left[\left|G(\sigma x) - G\left(\frac{\sigma x - m}{\sqrt{t}}\right)\right|\right.$$

$$+\frac{|t+m^2-1-2\varepsilon m^2|}{2\sigma}\left\{xg(\sigma x)+\frac{x}{\sqrt{t}}g\left(\frac{\sigma x-m}{\sqrt{t}}\right)\right\}\Bigg].$$

For each fixed $t\in(0, 1)$ and m, there exist positive ε_1 and N_1 such that

$$\max\left\{\sigma x, \frac{\sigma x-m}{\sqrt{t}}\right\}\leqq\frac{x}{t},\quad \min\left\{\sigma x, \frac{\sigma x-m}{\sqrt{t}}\right\}\geqq\frac{x}{\sqrt{2}}$$

for $0<\varepsilon<\varepsilon_1$ and $x>N_1$; and therefore

$$\left|1-(1-\varepsilon)G(\sigma x)-\varepsilon G\left(\frac{\sigma x-m}{\sqrt{t}}\right)\right|\geqq 1-G\left(\frac{x}{t}\right),$$

$$\left|G(\sigma x)-G\left(\frac{\sigma x-m}{\sqrt{t}}\right)\right|\leqq 1-G\left(\frac{x}{\sqrt{2}}\right).$$

Inserting these estimates into (2.7) and then using (2.2), we have

$$\text{(2.8)}\quad \left|\frac{\partial}{\partial\varepsilon}A_\varepsilon(x)\right|\leqq c_1\left(1-G\left(\frac{x}{t}\right)\right)^{-\delta}\left\{\left(1-G\left(\frac{x}{\sqrt{2}}\right)\right)+c_2xg\left(\frac{x}{\sqrt{2}}\right)\right\}$$

$$\leqq c_3\left[\frac{t}{x}\exp\left(-\frac{x^2}{2t^2}\right)\right]^{-\delta}x\exp\left(-\frac{x^2}{4}\right)$$

for $\varepsilon<\varepsilon_1$ and $x>N_1$, where c_2 and c_3 are some positive constants. An estimate similar to (2.8) for $\varepsilon<\varepsilon_2$ and $x<-N_2$ can be obtained for some $\varepsilon_2>0$ and $N_2>0$. Therefore, taking $\delta>0$ small enough, we see that there exists an integrable function $\psi(x)$ (independent of ε) such that

$$\left|\frac{\partial}{\partial\varepsilon}A_\varepsilon(x)\right|\leqq\psi(x),\quad 0<\varepsilon<\varepsilon_0,\quad |x|>N_0,$$

where $\varepsilon_0=\min(\varepsilon_1, \varepsilon_2)$ and $N_0=\max(N_1, N_2)$. On the other hand, from the explicit form of $\frac{\partial}{\partial\varepsilon}A_\varepsilon(x)$, it is clear that, for each fixed t and m, $\frac{\partial}{\partial\varepsilon}A_\varepsilon(x)$ is uniformly bounded on $\{|x|\leqq N_0\}$ for all sufficiently small $\varepsilon>0$. Therefore

$$\left|\frac{A_\varepsilon(x)}{\varepsilon}\right|=\left|\frac{1}{\varepsilon}\int_0^\varepsilon\frac{\partial}{\partial\varepsilon}A_\varepsilon(x)d\varepsilon\right|$$

is bounded by some integrable function for small $\varepsilon>0$, and hence by Lebesgue's convergence theorem we have

$$0\leqq\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\{\Phi[G]-\Phi[F_\varepsilon]\}$$

$$=\lim_{\varepsilon\downarrow 0}\int_{-\infty}^{\infty}\frac{A_\varepsilon(x)}{\varepsilon}dx=\int_{-\infty}^{\infty}\lim_{\varepsilon\downarrow 0}\frac{A_\varepsilon(x)-A_0(x)}{\varepsilon}dx$$

$$= \int_{-\infty}^{\infty} \varphi'(G(x)) \left\{ G(x) - \frac{t+m^2-1}{2} xg(x) - G\left(\frac{x-m}{\sqrt{t}}\right) \right\} dx,$$

which proves (2.4).

Proof of (2.5): Put

$$B_t(x) \equiv \varphi'(G(x)) \left\{ G(x) - \frac{t+m^2-1}{2} xg(x) - G\left(\frac{x-m}{\sqrt{t}}\right) \right\},$$

and evaluate the absolute value of $B_t(x)$. For each m there exists $t_0 \in (0, 1)$ such that

$$|x| \leqq \left| \frac{x-m}{\sqrt{t}} \right|, \qquad 0 < t < t_0,$$

for sufficiently large $|x|$, and we have

$$|B_t(x)| \leqq |\varphi'(G(x))| \left\{ \left| G(x) - G\left(\frac{x-m}{\sqrt{t}}\right) \right| + \frac{|t+m^2-1|}{2} |x| g(x) \right\}$$

$$\leqq (1+m^2)|\varphi'(G(x))|\,|x|g(x), \qquad |x| \to \infty,$$

for $0<t<t_0$. Since the last term in the above is integrable by (1.4), we obtain (2.5) by letting $t \downarrow 0$ in (2.4) and then applying Lebesgue's dominated convergence theorem.

Proof of (2.6): Take $N>|m|$ and write (2.5) with equality sign as

$$0 = \int_{|x|>N} C_m(x)dx + \int_{|x| \leqq N} C_m(x)dx = I_1 + I_2, \tag{2.9}$$

where $C_m(x) \equiv \varphi'(G(x)) \left\{ G(x) - \frac{m^2-1}{2} xg(x) - \mathbf{1}_{[m,\infty)}(x) \right\}$. Then, for $|x|>N$

$$\left| \frac{\partial}{\partial m} C_m(x) \right| = |\varphi'(G(x)) mxg(x)| \leqq |\varphi'(G(x))| x^2 g(x).$$

Since the last term in the above is integrable, we have

$$\frac{d}{dm} I_1 = \int_{|x|>N} \frac{\partial}{\partial m} C_m(x)dx = -m \int_{|x|>N} \varphi'(G(x)) xg(x)dx. \tag{2.10}$$

On the other hand

$$\frac{d}{dm} I_2 = -\frac{d}{dm} \int_{|x| \leqq N} \varphi'(G(x)) \frac{m^2-1}{2} xg(x)dx - \frac{d}{dm} \int_m^N \varphi'(G(x))dx$$

$$= - m\int_{|x| \leqq N} \varphi'(G(x))xg(x)dx + \varphi'(G(m)),$$

which combined with (2.9) and (2.10) proves (2.6).

§ 3. Proof of Theorem B and some remarks

1. Proof of Theorem B: It is enough to prove the following lemma.

Lemma. *Let φ be a function on* [0, 1] *satisfying* (1.4), *and assume that the functional Φ defined by* (1.3) *increases along the solutions of* (1.6) *with initial distribution functions belonging to $\mathscr{P}_1$. Then Φ satisfies* (1.5).

Proof. Let $F(t, x)$ be the solution of (1.6) with initial distribution function $F(x)$ belonging to $\mathscr{P}_1$. $F(t, x)$ can be expressed as Wild's sum (see [3] or [6]), and $F(t, \cdot) \in \mathscr{P}_1$ for each $t \geqq 0$. It was proved in [6] that $\mathfrak{e}[F(t)]$ decreases to 0 as $t \uparrow \infty$ (in [6] it was assumed that $F(x)$ has the finite fourth moment, but it is easy to remove this restriction), and hence $F(t, x)$ converges to $G(x)$ uniformly on compacts as $t \uparrow \infty$. Therefore for each $N > 0$

$$\lim_{t \to \infty} \int_{-N}^{N} \varphi(F(t, x))dx = \int_{-N}^{N} \varphi(G(x))dx. \tag{3.1}$$

On the other hand, since $F(t, \cdot) \in \mathscr{P}_1$ for each $t \geqq 0$, we have

$$F(t, -x) \text{ and } 1 - F(t, x) \leqq \frac{1}{x^2}, \qquad x \geqq 0. \tag{3.2}$$

Making use of (3.2) and the assumption (1.4) on φ, we can prove that

$$\lim_{N \to \infty} \sup_{t > 0} \left| \int_{|x| \geqq N} \varphi(F(t, x))dx \right| = 0,$$

which combined with (3.1) implies $\lim_{t \to \infty} \Phi[F(t)] = \Phi[G]$. Since the convergence is monotone by the assumption, we obtain (1.5).

2. *Inequality of convolution type*: When $F \in \mathscr{P}_1$ is the distribution function of a random variable X, we also write $\mathfrak{e}[X]$ ($\Phi_0[X]$) for $\mathfrak{e}[F]$ ($\Phi_0[F]$). Then the functional $\mathfrak{e}$ satisfies the following inequality of convolution type (see [6]):

(3.3) *Let X_1 and X_2 be independent random variables with distribution functions belonging to $\mathscr{P}_1$. Then, for $a, b > 0$ with $a^2 + b^2 = 1$,*

$$\mathfrak{e}[aX_1 + bX_2] < a^2\mathfrak{e}[X_1] + b^2\mathfrak{e}[X_2]$$

unless both X_1 and X_2 are Gaussian.

This fact was extended to multidimensional case by Murata and Tanaka [5], and to the case of real Hilbert spaces by Kondô and Negoro [2]. It follows immediately that the functional Φ_0 also has the following property:

(3.4) $\left\{\begin{array}{l}\textit{Let } X_1 \textit{ and } X_2 \textit{ be independent random variables with distribution} \\ \textit{functions belonging to } \mathscr{P}_1. \textit{ Then, for } a,\ b>0 \textit{ with } a^2+b^2=1, \\ \qquad \Phi_0[aX_1 + bX_2] > a^2\Phi_0[X_1] + b^2\Phi_0[X_2] \\ \textit{unless both } X_1 \textit{ and } X_2 \textit{ are Gaussian.}\end{array}\right.$

A remarkable application of (3.3) and (3.4) is that one can give a simple proof of the central limit theorem for sums of independent random variables (see [6]); for example, the following assertion can easily be proved by making use of (3.4): If $\{X_n\}_{n\geqq 1}$ is a sequence of independent random variables with a common distribution function belonging to $\mathscr{P}_1$, then $\lim_{n\to\infty}\Phi_0[n^{-1/2}\sum_{k=1}^n X_k]=1$.

References

[1] W. Feller, An Introduction to Probability Theory and Its Applications Vol. 1, 3rd ed., John Wiley and Sons, Inc., New York·London·Sydney.

[2] R. Kondô and A. Negoro, Certain functional of probability measures on Hilbert spaces, Hiroshima Math. J., **6** (1976), 421–428.

[3] H. P. McKean, Entropy is the only increasing functional of Kac's one-dimensional caricature of a Maxwellian gas, Z. Wahrscheinlichkeitstheorie verw. Geb., **2** (1963), 167–172.

[4] H. Murata, Propagation of chaos for Boltzmann-like equation of non-cutoff type in the plane, Hiroshima Math. J., **7** (1977), 479–515.

[5] H. Murata and H. Tanaka, An inequality for certain functional of multidimensional probability distributions, Hiroshima Math. J., **4** (1974), 75–81.

[6] H. Tanaka, An inequality for a functional of probability distributions and its application to Kac's one-dimensional model of a Maxwellian gas, Z. Wahrscheinlichkeitstheorie verw. Geb., **27** (1973), 47–52.

Department of Mathematics,
Faculty of Science,
Hiroshima University