Z. Wahrscheinlichkeitstheorie 3, 177–183 (1964)

Zur gemeinsamen Charakterisierung der Entropien α-ter Ordnung und der Shannonschen Entropie bei nicht unbedingt vollständigen Verteilungen

Von

J. Aczél

1.

In einer vorhergehenden Arbeit [3] hat Herr Z. Daróczy in dieser Zeitschrift den folgenden interessanten Satz bewiesen:

Genügen die für beliebige natürliche n und für $p_k > 0$ $(k = 1, 2, \ldots, n)$ bei $p_1 + p_2 + \cdots + p_n \leqq 1$ definierten Funktionen $I(p_1, p_2, \ldots, p_n)$ den Bedingungen

I. $I(p)$ *ist stetig in* $(0,1)$

II. $I(\frac{1}{2}) = 1$

III. $I(p_1 q_1, \ldots, p_1 q_m, \ldots, p_n q_1, \ldots, p_n q_m) = I(p_1, \ldots, p_n) + I(q_1, \ldots, q_m)$
$(p_k > 0, k = 1,2, \ldots, n; p_1 + p_2 + \cdots + p_n \leqq 1; q_j > 0, j = 1,2, \ldots, m;$
$q_1 + q_2 + \cdots + q_m \leqq 1)$

IV. $I(p_1, \ldots, p_n, q_1, \ldots, q_m) =$
$$= g^{-1}\left(\frac{(p_1 + \cdots + p_n)\, g(I(p_1, \ldots, p_n)) + (q_1 + \cdots + q_m)\, g(I(q_1, \ldots, q_m))}{p_1 + \cdots + p_n + q_1 + \cdots + q_m}\right)$$

$(p_1 + \cdots + p_n + q_1 + \cdots + q_m \leqq 1;\ \ p_k > 0,\ \ q_j > 0;\ \ k = 1, 2, \ldots, n;\ j = 1, 2,$ $\ldots, m$; *g stetig und streng monoton), dann und nur dann gilt entweder*

$$I(p_1, \ldots, p_n) = I_\alpha(p_1, \ldots, p_n) = \frac{1}{1-\alpha} \log_2 \left(\sum_{k=1}^{n} p_k^\alpha \Big/ \sum_{k=1}^{n} p_k \right) \qquad (\alpha \neq 1)$$

oder

$$I(p_1, \ldots, p_n) = I_1(p_1, \ldots, p_n) = -\sum_{k=1}^{n} p_k \log p_k \Big/ \sum_{k=1}^{n} p_k .$$

Hier wollen wir zwei wesentlich kürzere Beweise für den obigen Satz geben. Mehrere Grenzprozesse und ähnliche Betrachtungen, darunter auch der Übergang auf Mittelwerte der Wahrscheinlichkeitsverteilungen können dabei beseitigt werden.

Wir bemerken, daß es genügt III und IV für $n = m = 1$ sowie III für $n = 2$, $m = 1$ ausnützen, so daß statt der unendlichen Funktionenfolge

$$I(p_1), \quad I(p_1, p_2), \quad I(p_1, p_2, p_3), \ldots, \quad I(p_1, p_2, \ldots, p_n), \ldots$$

nur die beiden Funktionen

$$I(p) \quad \text{und} \quad I(p, q)$$

in Bedingungen und Behauptung vorkommen (ähnlich können $I(p)$ und $I(p_1, p_2, \ldots, p_{n_0})$ für beliebiges fixes $n_0 \geqq 2$ charakterisiert werden). —

Wir schließen diese kleine Note mit einer Bemerkung, wie die praktisch unbrauchbaren Entropien nicht-positiver Ordnung ausgesondert werden können und mit einer teilweisen Verallgemeinerung eines bei dem einen Beweis benötigten Lemma's, die auch sonst interesant zu sein scheint.

2.

Aus I, II und aus III für $n = m = 1$ folgten sofort (vgl. [*1*] S. 48—51 oder [*3*] oder [*5*] S. 552)

$$I(p) = -\log p \tag{1}$$

(hier und im folgenden *bedeutet* log *den Logarithmus mit Basis 2*) und aus IV für $n = m = 1$ folgt dann

$$I(p, q) = g^{-1}\left(\frac{p\,g(-\log p) + q\,g(-\log q)}{p+q}\right) (p > 0, q > 0, \quad p + q \leqq 1)\,. \tag{2}$$

Endlich ergibt III für $n = 2$, $m = 1$

$$I(pr, qr) = I(p, q) + I(r)$$

d. h. wegen (1) und (2)

$$\begin{aligned} &g^{-1}\left(\frac{p\,g(-\log p - \log r) + q\,g(-\log q - \log r)}{p+q}\right) = \\ &\qquad = g^{-1}\left(\frac{p\,g(-\log p) + q\,g(-\log q)}{p+q}\right) - \log r\,. \end{aligned} \tag{3}$$

Schreiben wir $t = I(r) = -\log r$, dann drückt (3) die „Translativität" von I aus. Wir beweisen also den folgenden

Satz 1. *Die Entropiefunktion*

$$I(p, q) = g^{-1}\left(\frac{p\,g(-\log p) + q\,g(-\log q)}{p+q}\right),$$

wo g stetig und streng monoton ist und $p > 0, q > 0, p + q \leqq 1$ *gilt, ist dann und nur dann translativ:*

$$I(pr, qr) = I(p, q) + I(r) = I(p, q) - \log r\,,$$

d. h.

$$g^{-1}\left(\frac{p\,g(-\log p + t) + q\,g(-\log q + t)}{p+q}\right) = g^{-1}\left(\frac{p\,g(-\log p) + q\,g(-\log q)}{p+q}\right) + t\,, \tag{4}$$

$(0 < r \leqq 1, t \geqq 0)$, *wenn* $g(x) = ax + b$ *oder* $g(x) = a2^{(1-\alpha)x} + b$ $(a \neq 0, \alpha \neq 1)$ *ist, d. h.*

$$I(p, q) = I_1(p, q) = -\frac{p \log p + q \log q}{p+q}$$

(Shannonsche Entropie) oder

$$I(p, q) = I_\alpha(p, q) = \frac{1}{1-\alpha} \log \frac{p^\alpha + q^\alpha}{p+q} \quad (\alpha \neq 1)$$

ist (Entropien α*-ter Ordnung).*

Beweis. Wir setzen in (4) noch $x = -\log p$, $y = -\log q$

$$g^{-1}\left(\frac{2^{-x}g(x+t)+2^{-y}g(y+t)}{2^{-x}+2^{-y}}\right) = g^{-1}\left(\frac{2^{-x}g(x)+2^{-y}g(y)}{2^{-x}+2^{-y}}\right) + t \quad (t \geqq 0, 2^{-x}+2^{-y} \leqq 1).$$

Schreiben wir

$$h_t(x) = g(x+t), \quad h_t^{-1}(z) = g^{-1}(z) - t, \tag{5}$$

so geht dies (wenn wir im Augenblick den Index t weglassen) in

$$h^{-1}\left(\frac{2^{-x}h(x)+2^{-y}h(y)}{2^{-x}+2^{-y}}\right) = g^{-1}\left(\frac{2^{-x}g(x)+2^{-y}g(y)}{2^{-x}+2^{-y}}\right) \quad (2^{-x}+2^{-y} \leqq 1) \tag{6}$$

über. Wir beweisen (vgl. [*3*] Satz 2), daß (6) *dann und nur dann besteht, falls*

$$h(x) = A\,g(x) + B \qquad (A \neq 0) \quad \textit{für alle} \quad x \geqq 0. \tag{7}$$

Jedenfalls erfüllt (7) die Gleichung (6) und mit $h(x)$ wird (6) auch durch

$$H(x) = h(x) - A\,g(x) - B \tag{8}$$

erfüllt:

$$H\left(g^{-1}\left(\frac{2^{-x}g(x)+2^{-y}g(y)}{2^{-x}+2^{-y}}\right)\right) = \frac{2^{-x}H(x)+2^{-y}H(y)}{2^{-x}+2^{-y}} \quad (2^{-x}+2^{-y} \leqq 1). \tag{9}$$

In (8) können A und B so gewählt werden, daß

$$H(1) = H(2) = 0 \tag{10}$$

ist. In (9) führen wir die Bezeichnungen

$$x \circ y = g^{-1}\left(\frac{2^{-x}g(x)+2^{-y}g(y)}{2^{-x}+2^{-y}}\right), \quad x \,\square\, y = \frac{2^{-x}H(x)+2^{-y}H(y)}{2^{-x}+2^{-y}} \tag{11}$$

ein, so daß

$$H(x \circ y) = x \,\square\, y \qquad (2^{-x}+2^{-y} \leqq 1) \tag{12}$$

bleibt, wo wegen (8) $H(x)$ und wegen (11) $x \circ y$ stetig als Funktion von x und y sind, und ebenfalls wegen (11) $x \circ y$ im *Inneren* des Intervalles (x, y) un x p$\square$ y im *Inneren* von $(H(x), H(y))$ liegt $((u, v) \equiv (v, u); (u, u) \equiv \{u\})$:

$$x \circ y \in (x, y), \qquad x \,\square\, y \in (H(x), H(y)). \tag{13}$$

Die Idee unseres ersten Beweises ist, daß dieser Sachverhalt zusammen mit (10) schon genügt, um

$$H(x) \equiv 0 \quad \text{für alle} \quad x \geqq 0$$

beweisen zu können. Dies besagt das folgende

Lemma. *Sind $H(x)$ und $x \circ y$ stetig, gelten*

$$x \circ y \in (x, y), \quad x \,\square\, y \in (H(x), H(y)) \tag{13}$$

und

$$H(x \circ y) = x \,\square\, y \tag{12}$$

für alle x, y, mit $2^{-x} + 2^{-y} \leqq 1$, ist ferner $H(1) = H(2)$, dann ist

$$H(x) \equiv 0 \tag{14}$$

für alle $x \geqq 0$.

Dieses Lemma verallgemeinert den Satz 1 von [*3*].

Beweis des Lemmas. a) Erstens beweisen wir (14) für alle $x \in [1, 2]$. Gäbe es nämlich ein $x_1 \in [1, 2]$ derart, daß $H(x_1) \neq 0$, so würde die Menge $\boldsymbol{E}_1$ aller $x \in [1, x_1)$ mit $H(x) = 0$ eine obere Grenze C_1 und die Menge $\boldsymbol{E}_2$ aller $x \in (x_1, 2]$ mit $H(x) = 0$ eine untere Grenze C_2 besitzen, die wegen der Stetigkeit von H zu $\boldsymbol{E}_1$ bzw. $\boldsymbol{E}_2$ gehören, ($1 \in \boldsymbol{E}_1, 2 \in \boldsymbol{E}_2, 1 \leqq C_1 < C_2 \leqq 2$) so daß $H(x) \neq 0$ für alle $x \in (C_1, C_2)$, dagegen $H(C_1) = H(C_2) = 0$ wäre. Laut (12) gilt aber

$$H(C_1 \circ C_2) = C_1 \square C_2$$

und wegen (13) $C_1 \square C_2 \in (H(C_1), H(C_2)) = (0, 0)$ und $C_1 \circ C_2 \in (C_1, C_2)$, so daß $C_1 \square C_2 = 0$ während $H(C_1 \circ C_2) \neq 0$: ein Widerspruch. ($2^{-C_1} + 2^{-C_2} < 1$, da $C_2 > C_1 \geqq 1$.) So gilt schon

$$H(x) = 0 \quad \text{für alle} \quad x \in [1, 2] \tag{15}$$

b) Zweitens beweisen wir (14) für alle $x \in (2, \infty)$. Gäbe es nämlich eine größte Zahl E, so daß $H(x) = 0$ für alle $x \in [1, E]$, dann würde es für jedes ε ein x geben, so daß $H(x) \neq 0$ für $x \in (E, E + \varepsilon)$ und es gäbe in $(E, E + \varepsilon)$ genügend nahe zu E ein x_2 derart, daß $H(x_1) \neq 0$ aber $1 \circ x_2 \in [1, E]$. Wäre nämlich $1 \circ x > E$ für je ein $x \in (E, E + \varepsilon)$, dann würde — wegen der Stetigkeit von $x \circ y$ — $1 \circ E \geqq E$ ($E > 1$) folgen, im Gegensatz zu (13). So gilt aber $H(1 \circ x_2) = 0, H(1) = 0$, $H(x_2) \neq 0$, so daß $0 = H(1 \circ x_2) = 1 \square x_2$ nicht im Inneren von $(0, H(x_2))$ liegt, im Gegensatz zu (13). ($2^{-x_2} + 2^{-1} < 2^{-E} + 2^{-1} < 1$ wegen $x_2 > E > 1$.) Also gilt mit (15)

$$H(x) = 0 \quad \text{für} \quad x \in [1, \infty].$$

c) Endlich beweisen wir (14) auch für alle $x \in [0, 1)$. Der Beweis geht ähnlich wie in b), nur muß man wegen der Voraussetzung $2^{-x} + 2^{-y} \leqq 1$ etwas vorsichtiger sein. Es sei C die kleinste nichtnegative Zahl derart, daß $H(x) = 0$ für alle $x \geqq C$. Wir wollen $C = 0$ beweisen. Wäre nämlich $C > 0$, so gäbe es für jedes positive $\varepsilon < C$ ein $x \in (C - \varepsilon, C)$ derart, daß $H(x) \neq 0$ wäre. Wir wählen ein festes

$$D > -\log(1-2^{\varepsilon - C}) \tag{16}$$

und in $(C - \varepsilon, C)$ ein — wegen der Stetigkeit von $x \circ y$ wieder existierendes — genügend nahe zu C liegendes x_0 für das

$$x_0 \circ D \geqq C, \tag{17}$$

und $H(x_0) \neq 0$, $H(D) = 0$,

$$0 = H(x_0 \circ D) = x_0 \square D \in (H(x_0), 0)$$

was wieder ein Widerspruch ist. ($2^{-x_0} + 2^{-D} < 2^{-(C-\varepsilon)} + 2^{-D} < 1$ wegen $x_0 > C - \varepsilon$ und wegen (16)). Also ist $C = 0$, d. h.

$$H(x) = 0 \quad \text{für alle} \quad x \geqq 0. \tag{14}$$

Damit haben wir das Lemma und zugleich die Formel (7) bewiesen.

Eine andere Beweisanordnung für (7) ersetzt — auch eine Idee des Herrn Daróczy verwendend — die Teile a) und b) des obigen Lemmas durch ein bekanntes Hilfsmittel aus [2] und [4].

Aus (6) folgt nämlich mit

$$h(g^{-1}(z)) = G(z) \tag{18}$$

das Bestehen von

$$G\left(\frac{2^{-x}g(x) + 2^{-y}g(y)}{2^{-x} + 2^{-y}}\right) = \frac{2^{-x}G(g(x)) + 2^{-y}G(g(y))}{2^{-x} + 2^{-y}} \qquad (2^{-x} + 2^{-y} \leqq 1)\,.$$

Dies bedeutet aber, daß an der Kurve $w = G(z)$ zwischen $z = g(1)$ und $z = g(N)$ mit beliebig großen N (da $2^{-x} + 2^{-y} \leqq 1$ für $x \geqq 1$, $y \geqq 1$ immer erfüllt ist), jede Sehne wenigstens einen Punkt außer den Endpunkten mit der Kurve gemeinsam hat. (Die Sehne mit den Endpunkten $\{u, G(u)\}$, $\{v, G(v)\}$ hat nämlich auch den Punkt

$$\left\{\frac{2^{-x}u + 2^{-y}v}{2^{-x} + 2^{-y}},\ \frac{2^{-x}G(u) + 2^{-y}G(v)}{2^{-x} + 2^{-y}}\right\}$$

mit der Kurve gemeinsam. Die „Gewichte" 2^{-x}, 2^{-y} hängen natürlich von den Endpunkten $u = g(x)$, $v = g(y)$ ab, das macht aber nichts: ein behaupteter innerer gemeinsamer Punkt existiert immer.) Dann ist aber (s. [4], S. 73; [2] S. 245) G linear:

$$G(z) = Az + B\,,$$

und wegen (18)

$$h(x) = A\,g(x) + B \tag{7}$$

für alle $x \geqq 1$. Um (7) auch für alle $x \in [0, 1)$ zu beweisen, verfahren wir wie in c):

Es sei C die kleinste Zahl derart, daß (7) für alle $x \geqq C$ bestehen. Wäre $C > 0$, so gäbe es wieder für alle $0 < \varepsilon < C$ ein $x \in (C - \varepsilon, C)$, so daß (7) *nicht* bestehen würde:

$$h(x) \neq A g(x) + B \quad \text{für} \quad x \in (C - \varepsilon, C)\,. \tag{19}$$

Andererseits folgt aus (6)

$$h\left(g^{-1}\left(\frac{2^{-x}g(x) + 2^{-y}g(y)}{2^{-x} + 2^{-y}}\right)\right) = \frac{2^{-x}h(x) + 2^{-y}h(y)}{2^{-x} + 2^{-y}} \qquad (2^{-x} + 2^{-y} \leqq 1) \tag{20}$$

und für die durch (11_1) definierte *stetige* Operation $x \circ y$ gilt wieder (13_1). Wenn wir also D wieder laut (16) wählen, so gibt es ein

$$x_0 \in (C - \varepsilon, C)\,, \tag{21}$$

so daß (17) und (19) bestehe und $2^{-x_0} + 2^{-D} \leqq 1$. Da aber (7) für $x \geqq C$ schon gültig ist, wird aus (20) für $x = x_0$, $y = D$

$$A\,\frac{2^{-x}g(x_0) + 2^{-D}g(D)}{2^{-x} + 2^{-D}} + B = \frac{2^{-x_0}h(x_0) + 2^{-D}(A g(D) + B)}{2^{-x_0} + 2^{-D}}$$

d. h.

$$h(x_0) = A\,g(x_0) + B$$

im Gegensatz zu (19) und (21) womit $C = 0$ und (7) für alle $x \geqq 0$ wieder bewiesen ist.

Ende des Beweises vom Satz 1. Schreiben wir in (7) den Index t (vgl. (5)) aus, so wird

$$h_t(x) = A_t g(x) + B_t$$

oder (mit (5))

$$g(x+t) = A(t) g(x) + B(t) \qquad (x \geqq 0, t \geqq 0).$$

Die einzigen streng monotonen Lösungen dieser Funktionalgleichung (s. [*1*] S. 120—121, vgl. [*3*], [*4*], S. 68—69, [*5*] S. 557—558) sind aber

$$(22) \quad g(x) = ax + b \quad \text{und} \quad g(x) = ae^{cx} + b = a2^{(1-\alpha)x} + b \, (a \neq 0, \; c \neq 0, \; \alpha \neq 1)$$

womit der Satz 1 bewiesen ist.

3.

Um $\alpha \leqq 0$ auszuschließen, fordern wir noch die als natürlich erscheinende Bedingung

$$(23) \qquad \lim_{q \to 0} I(p, q) = I(p) = -\log p,$$

die tatsächlich (vgl. (2))

$$(24) \qquad \lim_{q \to 0} q g(-\log q) = 0$$

zu Folge hat. Unter den Funktionen (22) erfüllt $g(x) = ax + b$ die Bedingung (24) immer, aber $g(x) = a2^{(1-\alpha)x} + b$ nur wenn $\alpha > 0$ ist.

Damit haben wir das folgende

Korollar. *Unter den stetigen und streng monotonen Lösungen der Funktionalgleichung* (4) *erfüllen nur*

$$g(x) = ax + b \quad \textit{und} \quad g(x) = a2^{(1-\alpha)x} + b \quad \textit{mit} \quad a \neq 0, \quad \alpha \neq 1 \quad \textit{und} \quad \alpha > 0$$

die Bedingung (24), *d. h. die Shannonsche Entropie und die Entropien positiver Ordnung sind die einzigen translativen Entropiefunktionen, die der Bedingung* (23) *genügen.* —

Wir wollen endlich noch unser Lemma teilweise verallgemeinern. Ähnlich wie dieses Lemma (ja sogar wegen des Wegfallens der Bedingung $2^{-x} + 2^{-y} \leqq 1$ noch leichter) kann man folgendes beweisen:

Satz 2. *Sind in einem (endlichen oder unendlichen, offenen, halboffenen oder abgeschlossenen) Intervall $\langle a, b \rangle$ die Funktionen $H(x)$ und $x \circ y$ als Funktion von x und y stetig, gelten ferner*

$$(12) \qquad H(x \circ y) = x \square y,$$

$$(13) \qquad x \circ y \in (x, y), \quad x \square y \in (H(x), H(y))$$

in $\langle a, b \rangle$ und gibt es zwei beliebige Werte $a', b' \in \langle a, b \rangle$ mit $a' \neq b'$ derart, daß

$$H(a') = H(b'),$$

dann ist die Funktion H im ganzen Intervall $\langle a, b \rangle$ konstant.

Insbesondere gilt noch folgendes: Es sei für reelle x, y eine Mittelwertoperation $x * y$ erklärt mit der Eigenschaft $x * y \in (x, y)$, so daß die Funktionalgleichung

$H(x \circ y) = H(x) * H(y)$ erfüllt ist. Wenn man dann die Operation $x \square y$ durch $H(x) * H(y)$ ersetzt, bleibt die Behauptung des Satzes 2 richtig.

Einige Spezialfälle dieses als wichtig erscheinenden Satzes wurden schon in [1] (S. 50, 85—86, 203—206) angewendet.

Literatur

[1] ACZÉL, J.: Vorlesungen über Funktionalgleichungen und ihre Anwendungen. Basel-Stuttgart 1961.

[2] BAJRAKTAREVIĆ, M.: Sur une équation fonctionnelle aux valeurs moyennes. Glasnik Mat.-Fiz. i Astr. **13**, 243—248 (1958).

[3] DARÓCZY, Z.: Über die gemeinsame Charakterisierung der zu den nicht vollständigen Verteilungen gehörigen Entropien von Shannon und von Rényi. Z. Wahrscheinlichkeitstheorie **1**, 381—388 (1963).

[4] HARDY, G. H., J. E. LITTLEWOOD and G. PÓLYA: Inequalities. 2nd edition, Cambridge 1952.

[5] RÉNYI, A.: On measures of entropy and information. Proc. Fourth Berkeley Sympos. math. Statist. Probability 1961, 547—561.

Mathematisches Institut der Universität
L. Kossuth, Debrecen 10/Ungarn

(Eingegangen am 8. Juli 1963)