PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 127, Number 10, Pages 2923–2929
S 0002-9939(99)04863-7
Article electronically published on April 23, 1999

SOLUTION OF A FUNCTIONAL EQUATION ARISING FROM UTILITY THAT IS BOTH SEPARABLE AND ADDITIVE

JÁNOS ACZÉL, ROMAN GER, AND ANTAL JÁRAI

(Communicated by Frederick W. Gehring)

Abstract. The problem of determining all utility measures over binary gambles that are both separable and additive leads to the functional equation

$$f(v) = f(vw) + f[vQ(w)], \qquad v, vQ(w) \in [0, k), \quad w \in [0, 1]\,.$$

The following conditions are more or less natural to the problem: f strictly increasing, Q strictly decreasing; both map their domains onto intervals (f onto a $[0, K)$, Q onto $[0, 1]$); thus both are continuous, $k > 1$, $f(0) = 0$, $f(1) = 1$, $Q(1) = 0$, $Q(0) = 1$. We determine, however, the general solution without any of these conditions (except $f : [0, k) \to \mathbb{R}_+ := [0, \infty)$, $Q : [0, 1] \to \mathbb{R}_+$, both into). If we exclude two trivial solutions, then we get as general solution $f(v) = \alpha v^\beta$ ($\beta > 0$, $\alpha > 0$; $\alpha = 1$ for $f(1) = 1$), which satisfies all the above conditions.

The paper concludes with a remark on the case where the equation is satisfied only almost everywhere.

1. Introduction

R. Duncan Luce (Univ. of California, Irvine) suggested the following problem.

Let $(x, E; y)$ denote the binary gamble in which x is the consequence when event E occurs and y otherwise. Let $\succsim$ denote a preference ordering over gambles and e the status quo, i.e. no consequence is received. *Event commutativity*, a behavioral property, means

$$((x, D; e), E; e) \sim ((x, E; e), D; e)$$

for all consequences x and for all independent events D and E ($\sim$ denotes that both $\succsim$ and $\precsim$ hold). Under conditions of adequate density of events and consequences,

Received by the editors June 17, 1997 and, in revised form, December 23, 1997.

1991 *Mathematics Subject Classification.* Primary 39B22, 90D05.

Key words and phrases. Utility measures, binary gambles, bounded, monotonic, continuous, differentiable functions, elevating boundedness to differentiability, functional equation with two unknown functions.

The research of the first author has been supported in part by the NSERC (Canada), grant no. OGP002972. He is also grateful for the hospitality of the Institute for Mathematical Behavioral Sciences, University of California, Irvine.

The research of the second author has been supported by the KBN (Poland) grant no. 2 P03A 049 09.

The research of the third author has been supported by the OTKA (Hungary) grant no. T-016846.

this event commutativity is sufficient for the existence of an order preserving *separable* representation of utility of the form

$$V^*(x, E; e) = V(x)W(E),$$

where W is onto $[0,1]$ and V is onto an interval that can be taken $[0,1]$ with $V(e) = 0$. Another axiomatization, due to Wakker [6] and which is closely related to additive conjoint measurement (Krantz, Luce, Suppes, and Tversky [3]), shows when there is an *additive* representation of the following type:

$$U(x, E; y) = U_1(x, E) + U_2(y, E),$$

where $U_i(e, E) = 0$, $i = 1, 2$. Furthermore, U may be normalized onto the interval $[0,1]$. The question here is (cf. Luce [4]) when do both representations hold simultaneously.

The following are additional noncontroversial assumptions. For all consequences x and y and events E, *idempotence:* $(x, E; x) \sim x$; *complementarity:* $(x, E; y) \sim (y, \overline{E}; x)$, where $\overline{E} = \Omega \setminus E$ and Ω is the certain (universal) event; and *certainty:* $(x, \Omega; y) \sim x$; and the usual form of *conjoint Archimedeanness* hold.

Because V^* and U are both order preserving, there is a strictly increasing function $f : [0, k) \to [0, \infty)$ such that $U_2(x, E) = f[V(x)W(\overline{E})]$. Setting $v = V(x)$ and $w = W(E)$, one shows that there is a function $Q : [0,1] \to [0,1]$ such that

$$f(v) = f(vw) + f[vQ(w)].$$

We solve this functional equation for both f and Q without any supposition (except the boundedness of Q and f from below by zero, implicit in the statement of ranges). We will point out, however, those places where positivity on $(0, k)$ and/or the supposition that the functions map their domains onto proper intervals can shorten the proof.

The first attempt at a proof used differentiability of the functions f and Q. Next we succeeded in reducing this assumption to continuity, while saving essential steps of the previous argument. Subsequently we derived continuity from strict monotonicity and finally we eliminated even that assumption. So (except for two trivial solutions) all these conditions are now consequences of boundedness of Q and f from below and of the functional equation.

On hearing of this drastic reduction of regularity suppositions, Duncan Luce asked in jest whether we planned also to eliminate the functional equation. This, of course, is not possible. But we can weaken the equation by supposing its validity only "almost everywhere". Then also the result holds almost everywhere.

2. Main result

In what follows the symbol $\mathbb{R}_+$ will stand for $[0, \infty)$.

Theorem. *Among the functions* $f : [0, k) \to \mathbb{R}_+$, $k \in (0, \infty]$, *and* $Q : [0,1] \to \mathbb{R}_+$, *the functional equation*

$$f(v) = f(vw) + f[vQ(w)] \qquad (v, vQ(w) \in [0, k), \quad w \in [0,1]), \tag{1}$$

has the trivial solutions

$$f = 0, \quad Q \textit{ arbitrary} \quad (\leq 1 \textit{ if } k < \infty), \tag{2}$$

and

$$(3)\qquad \begin{cases} f(v) = c > 0 \quad (v \in (0,k)), & f(0) = 0, \\ Q(w) = 0 \quad (w \in (0,1]), & Q(0) > 0 \text{ arbitrary } (\le 1 \text{ if } k < \infty). \end{cases}$$

For all other solutions there exist constants $\alpha > 0$, $\beta > 0$ *such that*

$$(4)\qquad f(v) = \alpha v^{\beta}, \qquad Q(w) = (1 - w^{\beta})^{\frac{1}{\beta}}.$$

If $k > 1$ *and* $f(1) = 1$, *then* $\alpha = 1$.

Conversely, all pairs of functions of the form (2), (3) *and* (4) *satisfy* (1).

Note 1. The functions f and/or Q may map *into* the given ranges; "onto" is not assumed. Neither is any regularity (monotonicity, continuity) supposed. These follow from the result. (Actually, the ranges state 0 as a lower bound). — Because of $v, vQ(w) \in [0,k)$ in (1), we have to have $Q(w) \le 1$ *for all* $w \in [0,1]$ *except if* $k = \infty$.

We present the proof in several lemmas.

Lemma 1. *We have* $f(0) = 0$ *and* f *is increasing.*

Notice that "increasing" is claimed in the weaker sense (not yet "strictly increasing").

Proof. Substituting $v = 0$ into (1) we get $f(0) = 0$. Since $f(vQ(w)) \ge 0$, we obtain from (1) that $f(v) \ge f(vw) \ge 0$. So f is increasing. □

Note 2. If the reader is willing to assume that there are z arbitrarily close to 0 with $f(z) > 0$, then, f being increasing, $f(v) > 0$ for all $v \in (0,k)$ and thus strictly increasing. If not, then we prove the following.

Lemma 2. *The function* Q *maps the interval* $(0,1]$ *into* $[0,1]$ *provided that* $f \ne 0$. *The function* f *is positive and either constant or strictly increasing on* $(0,k)$.

Proof. Suppose to the contrary that $Q(w_0) > 1$ for some $w_0 \in (0,1]$. Then (see Note 1) $k = \infty$. Thus, for every $v > 0$, in view of Lemma 1 and of equation (1), we would have

$$f(v) = f(vw_0) + f[vQ(w_0)] \ge f(vw_0) + f(v),$$

whence by the nonnegativity of f we deduce that $f(vw_0) = 0$ for all $v \in [0,\infty)$ (recall that $f(0) = 0$), i.e. $f = 0$.

To prove that f is positive on $(0,k)$ suppose to the contrary that there exists a $v_0 \in (0,k)$ with $f(v_0) = 0$. Then, f being increasing and nonnegative, $f(v) = 0$ for all $v \in [0,v_0]$. Let v_1 be the largest number (possibly ∞) for which $f(v) = 0$ on $[0,v_1)$. Then there exists no $v > v_1$ with $f(v) = 0$. We show that $v_1 = k$. Indeed, fix $w_0 \in (0,1)$ arbitrarily and take a $v \in (v_1, v_1/w_0)$. This implies $vw_0 \in (v_1 w_0, v_1)$ and thus $f(v) > 0$ and $f(vw_0) = 0$. So, by (1), $f(vQ(w_0)) = f(v) > 0$ and therefore $vQ(w_0) > v_1$, hence $Q(w_0) > v_1/v$. Letting $v \to v_1$ we get $Q(w_0) \ge 1$. Since also $Q(w_0) \le 1$, this means that $Q(w_0) = 1$ and $f(v) = f(vw_0) + f(v)$, that is $f(vw_0) = 0$ for *all* $v \in (0,k)$; thus for all $z \in (0,k)$ we have $f(z) = 0$. Since also $f(0) = 0$, we got the excluded $f = 0$ (which gives the trivial solution (2)). Thus there is no $v_0 \in (0,k)$ with $f(v_0) = 0$.

Suppose that $Q(w_0) = 0$ for some $w_0 \in (0,1)$. By (1) this means that $f(v) = f(vw_0)$ for all $v \in (0,k)$ whence, by an obvious induction, $f(v) = f(vw_0^n)$ for all

$v \in (0,k)$ and all $n \in \mathbb{N}$. Since vw_0^n tends to zero as n grows to infinity, we infer that f is constant on $(0,v)$ for every $v \in (0,k)$. This implies, since we already excluded $f=0$, that $0 < f(v) \equiv c$ (constant) on $(0,k)$. Then, by (1), $Q(w)=0$ for all $v \in (0,1]$. This exceptional case is the trivial solution (3).

Therefore, we may assume that $Q(w) > 0$ for all $w \in (0,1)$. Then, again by (1) and the fact that $f|_{(0,k)}$ is positive, we have

$$f(v) = f(vw) + f[vQ(w)] > f(vw)$$

for all $v \in (0,k)$ and $w \in (0,1)$ which states that f is strictly increasing. □

From here on we exclude the trivial solutions (2) *and* (3).

Lemma 3. *The function Q is strictly decreasing on* $[0,1]$ *and* $Q(1)=0$, $Q(0)=1$, *so* $0 < Q(w) < 1$ *for* $w \in (0,1)$.

Proof. If $w_1 > w_2$, then, by (1), $f(vQ(w_1)) = f(v) - f(vw_1) < f(v) - f(vw_2) = f[vQ(w_2)]$ since f is strictly increasing. So $vQ(w_1) < vQ(w_2)$, therefore Q is strictly decreasing. Putting $w=1$ or $w=0$ into (1) gives $f(vQ(1)) = 0$ and $f(vQ(0)) = f(v)$, respectively, so by Lemmas 1 and 2, $Q(1)=0$, $Q(0)=1$. □

Note 3. If the reader is willing to accept that f maps $[0,k)$ onto an interval and thus, being monotonic, is continuous, then the following Lemma 4 can be skipped.

Lemma 4. *The function f is continuous on* $[0,k)$ *and Q is continuous on* $[0,1]$.

Proof. As an increasing function, f has left-hand and right-hand limits at every point in $(0,k)$ and if the two are equal, then f is continuous at that point. Also, since it is bounded from below, f has a finite right-hand limit at 0, $L = \lim_{t\to 0+} f(v)$. Letting $v \searrow 0$ (that is, v tending to zero from the right) in (1), we get $L = L + L$ (except if $Q(w)=0$ for all $w \in (0,1]$, in which case $f(v)=c$ on $(0,k)$ and we have the excluded (2) or (3)), i.e. $L = \lim_{v\to 0+} f(v) = 0$. Since $f(0)=0$, *f is continuous at* 0. For an *arbitrary* $v_0 \in (0,k)$, take sequences $v_n \searrow v_0$, $w_n \searrow 0$ such that $v_n < v_0/Q(w_n)$ (which is possible by Lemma 3). If we let $n \to \infty$ in $f(v_n w_n) + f[v_n Q(w_n)] = f(v_n)$, then the right-hand side tends to the right-hand limit R of f at v_0. Since $v_n w_n \searrow 0$, the first term on the left tends to 0 by what we have just shown, so the second term has to converge to R. On the other hand, since $v_n Q(w_n) < v_0$, it can converge only to the left-hand limit of f at v_0. Thus the two limits are equal and *f is continuous at every* $v_0 \in (0,k)$ too.

From (1) with an arbitrary $v = c \in (0,k)$,

$$Q(w) = \frac{1}{c} f^{-1}(f(c) - f(cw)), \tag{5}$$

and so on $[0,1]$ the function Q is also continuous. □

Lemma 5. *The functions F and G, defined by*

$$F(z) = \int_0^z f(v)\,dv \quad (z \in [0,k)), \quad G(z) = \frac{F(z)}{z} \quad (z \in (0,k)), \quad G(0)=0, \tag{6}$$

are continuous and strictly increasing on $[0,k)$, *positive and continuously differentiable with* $F' > 0$, $G' > 0$ *on* $(0,k)$ *and G satisfies the functional equation* (1).

Proof. The statements about F are obvious, since $F' = f$ on $[0,k)$. Furthermore,

$$\lim_{z\to 0+} G(z) = \lim_{z\to 0+} \frac{F(z)}{z} = F'(0) = f(0) = 0,$$

so G is continuous also at 0 (and obviously so elsewhere). Clearly $G(z) > 0$ on $(0, k)$. Also

$$G'(z) = -z^{-2}F(z) + z^{-1}f(z) = z^{-2}\left(zf(z) - \int_0^z f(v)\,dv\right)$$
$$= z^{-2}\int_0^z (f(z) - f(v))\,dv$$

is continuous and positive on $(0, k)$, so G is strictly increasing there. Finally, integration of equation (1) with respect to v from 0 to z gives (with $s = vw$, $t = vQ(w)$)

$$F(z) = \int_0^z f(vw)\,dv + \int_0^z f(vQ(w))\,dv = \frac{1}{w}\int_0^{zw} f(s)\,ds + \frac{1}{Q(w)}\int_0^{zQ(w)} f(t)\,dt$$
$$= \frac{F(zw)}{w} + \frac{F[zQ(w)]}{Q(w)}.$$

If we divide by z and take (6) into consideration, we indeed get

$$G(z) = G(zw) + G[zQ(w)] \tag{7}$$

for $z \in (0, k)$, $w \in (0, 1)$ and, because of the continuity of G and Q, also for $z = 0$ and/or $w = 0$ or $w = 1$. □

Now we are in a position to conclude the

Proof of the theorem. (We disregard the trivial solutions (2) and (3), which have been completely established.) Equation (7) with $v = c \in (0, k)$ gives

$$Q(w) = \frac{1}{c}G^{-1}(G(c) - G(cw)) \quad (w \in (0, 1)).$$

So, since G is continuously differentiable and $G' > 0$ on $(0, 1)$, Q is also continuously differentiable on $(0, 1)$. Differentiating (7) with respect to z or w, we get

$$G'(z) = wG'(zw) + Q(w)G'(zQ(w)),$$

and

$$0 = G'(zw) + Q'(w)G'(zQ(w)), \tag{8}$$

respectively. Multiplying the first equation by $Q'(w)$, the second by $Q(w)$ and subtracting, we obtain

$$Q'(w)G'(z) = (wQ'(w) - Q(w))G'(zw) \quad (z \in (0, k),\ w \in (0, 1)). \tag{9}$$

If there existed a w_0 such that $w_0Q'(w_0) - Q(w_0) = 0$, then by (9) $Q'(w_0) = 0$ and by (8) $G'(zw_0) = 0$, contradicting $G' > 0$ (Lemma 5). So, with

$$H(w) = Q'(w)/[wQ'(w) - Q(w)],$$

we get the Pexider equation

$$G'(zw) = G'(z)H(w) \quad (z \in (0, k),\ w \in (0, 1)).$$

It is known (Radó and Baker [5], Aczél [1]) that for continuous positive G' the general solution of this equation is $G'(z) = Az^B$ with $A > 0$. If $B = -1$, then $G(z) = A\ln z + C$, which is impossible since it would imply $G(z) < 0$ for small z. So $B \neq -1$ and with $\beta = B + 1$, $a = A/(B + 1)$ we have $G(z) = az^\beta + C$ $(\beta \neq 0)$. But (see Lemma 5),

$$0 = \lim_{z \to 0+} G(z) = \lim_{z \to 0+} (az^\beta + C)$$

gives $C = 0$, $\beta > 0$. By (6) $F(z) = zG(z) = az^{\beta+1}$ ($a > 0$ since F is positive) and

$$f(v) = F'(v) = a(\beta+1)v^\beta = \alpha v^\beta \quad (v \in (0,k))$$

(where $\alpha > 0$, $\beta > 0$), and, by (5),

$$Q(w) = \frac{1}{c} f^{-1}(f(c) - f(cw)) = (1 - w^\beta)^{\frac{1}{\beta}} \quad (w \in (0,1)).$$

By continuity, these equations hold also for $v = 0$ and/or $v, w \in \{0, 1\}$. If also $k > 1$ and $f(1) = 1$ is supposed, then $\alpha = 1$. Since (2), (3) and (4) satisfy (1) and are nonnegative, the theorem is proved. $\square$

3. Weakening of the equation

In the previous section we eliminated all assumptions originally imposed upon the unknown functions (except nonnegativity). We now look briefly at weakening the equation itself. Below we sketch, omitting technical details, the idea how to solve equation (1) when it is assumed to hold only for almost all pairs $(v, w) \in [0, k) \times [0, 1]$ (with respect to the planar Lebesgue measure). Keeping the assumption that both f and Q are nonnegative, one can show that there exists exactly one pair (g, P) of functions $g : [0, k) \to \mathbb{R}_+$ and $P : [0, 1] \to \mathbb{R}_+$ satisfying equation (1) *everywhere*, i.e.

$$g(v) = g(vw) + g[vP(w)] \quad \text{for every } v \in [0, k),\ w \in [0, 1],$$

and such that $f(x) = g(x)$ for almost all $x \in [0, k)$ and $Q(x) = P(x)$ for almost all $x \in [0, 1]$ (in the sense of linear Lebesgue measure). Clearly, all such pairs (g, P) are entirely described by the theorem.

The following are key steps to the proof.

A set of measure 0 can be excluded from $(0, k)$ so that any v from the remaining subset satisfies (1) with almost every w in $(0, 1)$. A similar statement holds with w and $(0, k)$ in place of v and $(0, 1)$, respectively, and vice versa.

One then defines the *increasing* function g on $[0, k)$ by

$$g(v) = \inf\{\, r \geq 0 \mid f(z) \leq r \quad \text{for almost all } z \in (0, v) \,\} \qquad \text{for} \quad v \in (0, k) \quad \text{and} \quad g(0) = 0$$

and shows that it is almost everywhere equal to f but satisfies (1) everywhere.

Notice that the above result is the best that could be expected in this regard. Indeed the fact that (1) is satisfied almost everywhere does not imply that it is satisfied everywhere. If $f(v) = av^b$ almost everywhere and $Q(w) = (1 - w^b)^{1/b}$ almost everywhere but they are different on the remaining sets of measure 0, then (1) is satisfied almost everywhere but not everywhere on $[0, k) \times [0, 1]$.

Sets of Lebesgue measure 0 can be replaced in this proof for example by sets of first Baire category. We have also results for a wider class of "small" sets to be excluded from the domain, namely the members of an axiomatically given nonempty family of sets (a "linearly invariant proper ideal") $\Im \subset 2^{\mathbb{R}} \setminus \{\mathbb{R}\}$ which is closed under finite set-theoretical unions, hereditary with respect to descending inclusions and such that $\{x - S \mid x \in \mathbb{R}\} \subset \Im$ whenever $S \in \Im$ (cf. also [2]).

References

1. J. Aczél, *A short course on functional equations based upon recent applications to the social and behavioral sciences*, Reidel-Kluwer, Dordrecht-Boston, 1987. MR **88d:**39013
2. R. Ger, *Almost additive functions on semigroups and a functional equation*, Publicationes Mathematicae Debrecen **26** (1979), 219–228. MR **81e:**39010
3. D. H. Krantz, R. D. Luce, P. Suppes, and A. Tversky, *Foundations of measurement, Vol. 1*, Academic Press, New York, 1971. MR **56:**17265
4. R. D. Luce, *Coalescing, event commutativity, and theories of utility*, Journal of Risk and Uncertainty (to appear).
5. F. Radó and J. A. Baker, *Pexider's equation and aggregation of allocations*, Aequationes Mathematicae **32** (1987), 227–239. MR **89i:**39014
6. P. P. Wakker, *Additive representations of rank ordered sets, I. The algebraic approach*, Journal of Mathematical Psychology **35** (1991), 501–531. MR **93a:**92021

Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

E-mail address: jdaczel@math.uwaterloo.ca

Institute of Mathematics, Silesian University, Bankowa 14, PL-40-007 Katowice, Poland

E-mail address: romanger@us.edu.pl

Universität GH Paderborn, FB 17, D-33095 Paderborn, Germany

E-mail address: jarai@uni-paderborn.de

Current address: Department of Numerical Analysis, Eötvös Lóránd University, Pázmány Péter Sétány 1/D, H-1117 Budapest, Hungary

E-mail address: ajarai@moon.inf.elte.hu