The non-existence of a Hamel-basis and the general solution of Cauchy's functional equation for nonnegative numbers

By J. ACZÉL (Debrecen—Waterloo Ont) and P. ERDŐS (Budapest)

1. It is well known that the general continuous solution of Cauchy's functional equation

$$(1) \qquad f(x+y) = f(x)+f(y)$$

is (see e. g. [3] or [1])

$$(2) \qquad f(x)=cx,$$

with an arbitrary constant c. As DARBOUX has proved ([4]), this (with nonnegative c's) is also the most general solution of (1) which is nonnegative for positive x's (it is even enough to suppose the nonnegativity for small positive x's). But without any regularity-suppositions (2) isn't anymore the most general solution of (1), this can be shown and at the same time the most general solution of (1) can be constructed with the Hamel-basis of real numbers ([7]).

In all these results (1) was supposed valid for *all real* x, y and then also (2) is verified for all real x's moreover, the Hamel-basis also gives a representation of *all real* numbers. — But for many applications (see e. g. [1]), (1) can be supposed valid only for *nonnegative* x, y. It is easy to show that (2) (with nonnegative x) remains the most general continuous solution of (1) also with this restriction and with nonnegative x, c also the most general solution nonnegative for (small) positive variables. — But how to construct in this case *the most general solution of* (1) *for nonnegative* x, y? Are there *Hamel-bases of the nonnegative numbers?*

In this little note we answer the second question in the negative and construct nevertheless the general solution asked for in the first one.

2. We first recall the above-mentioned results of Cauchy, Darboux and Hamel with short proofs: (1) implies by induction

$$(3) \qquad f(x_1+x_2+\ldots+x_n) = f(x_1)+f(x_2)+\ldots+f(x_n)$$

and with $x_1=x_2=\ldots=x_n=x$

$$(4) \qquad f(nx)=nf(x).$$

If now $x=\frac{m}{n}t$ $(m>0,\ n>0)$ then by (4)

$$nf(x)=f(nx)=f(mt)=mf(t) \quad \text{i. e.} \quad f\left(\frac{m}{n}t\right)=\frac{m}{n}f(t)$$

or

$$f(rt)=rf(t) \tag{5}$$

for positive rational r's. For $x=0$ (1) gives

$$f(0)=0, \tag{6}$$

so that (5) remains valid for all *nonnegative rational* r's. If (1) was supposed for negative x's too, then from (1) with $y=-x$

$$f(-x)=-f(x) \tag{7}$$

follows and (5) becomes valid for all negative rational r's too. — In particular for $t=1$ we have from (5), by denoting $f(1)=c$,

$$f(r)=cr \tag{8}$$

for all nonnegative rational or for all rational r's, respectively.

If f is supposed to be continuous, then from (8)

$$f(x)=cx \tag{2}$$

follows also *for all nonnegative or all real* x's, respectively. But *this follows also from the nonnegativity of f for (small) positive* x's.

In fact, then for (small) $y>0$

$$f(x+y)=f(x)+f(y)\geqq f(x)$$

i. e. f is *monotonic* (increasing) and now for irrational x we take two sequences $\{r_n\}$ and $\{R_n\}$ tending to x monotonically from below or from above, respectively. Then $r_n<x<R_n$ and by (8) and by the monotony of f

$$cr_n=f(r_n)\leqq f(x)\leqq f(R_n)=cR_n.$$

But for $n\to\infty$ both $\{r_n\}$ and $\{R_n\}$ tend to x, so $cr_n\to cx$, $cR_n\to cx$ and $f(x)$, being between the two has to be cx too, so that (2) is proved ($f(x)=cx$ is nonnegative for positive x's if $c\geqq 0$).

Now, if *no regularity-suppositions are made at all, but* (1) *is supposed valid for all real* x, y, then we make use of the Hamel-bases of real numbers. Hamel ([7]) has proved that there exist subsets B (the „Hamel-bases") of the set of real numbers the elements of which are *rationally independent* (i. e. $r_1b_1+r_2b_2+\ldots+r_mb_m=0$ only if $r_1=r_2=\ldots=r_m=0$, here $b_j\in B$ and the r_j's are rational numbers; $j=1,2,\ldots$ $\ldots,m$, m being an arbitrary positive integer) and so that *all real numbers can be represented in a unique way as linear combinations of the basis-elements with rational coefficients:*

$$x=r_1b_1+r_2b_2+\ldots+r_nb_n \qquad (b_j\in B,\ r_j \text{ rational}). \tag{9}$$

So the general solution of (1) *can be constructed by choosing the values of f in an arbitrary way on B and then for any real x with the representation* (9) *by defining f by*

$$f(x)=r_1f(b_1)+r_2f(b_2)+\ldots+r_nf(b_n). \tag{10}$$

In fact, (10) satisfies (1), as one can see by substitution and on the other hand, from (3), (5) (which were consequences of (1), derived without any regularity-suppositions) and from (9)

$$f(x) = f(r_1b_1 + r_2b_2 + \ldots + r_nb_n) = r_1f(b_1) + r_2f(b_2) + \ldots + r_nf(b_n)$$

i. e. (10) follows, q. e. d.

3. This last proof shows that we could construct the general solution of the equation (1) valid for nonnegative x, y, if there would exist a "Hamel-basis B' of the nonnegative numbers" the elements of which were *nonnegative* real numbers and any non-negative real number could be represented in a *unique* way as linear combinations of the basis-elements with *non-negative rational* coefficients. We show that this is *not possible*.

In fact, if there would exist such a basis B', then its elements would have to be *rationally independent* as defined above (i. e. with r_j rational and not only positive rational numbers) for if e. g.

$$r_1b_1 + r_2b_2 + \ldots + r_kb_k - r_{k+1}b_{k+1} - \ldots - r_mb_m = 0$$

could hold with positive rational r_j's and with one r_j different from 0, then the representation of the nonnegative number

$$x = r_1b_1 + r_2b_2 + \ldots + r_kb_k = r_{k+1}b_{k+1} + \ldots + r_mb_m$$

would not be unique. — But for two basis elements b_1, b_2 there exist positive numbers n big enough so that

$$y = nb_1 - b_2 > 0.$$

If y had a representation

$$y = r_1b_1 + r_2b_2 + r_3b_3 + \ldots + r_nb_n \qquad (r_j > 0 \text{ rational } b_j \in B')$$

then the basis elements would *not* be rationally independent as we would have

$$nb_1 - b_2 = r_1b_1 + r_2b_2 + r_3b_3 + \ldots + r_nb_n,$$

i. e. $(r_1 - n)b_1 + (r_2 + 1)b_2 + r_3b_3 + \ldots + r_nb_n = 0$, $r_2 + 1 \neq 0$. Thus we have proved the following

Theorem 1. *There don't exist Hamel-bases of the set of nonnegative numbers, i. e. there does not exist any set B' whose elements are nonnegative and any non-negative number is representable as a linear combination of a finite number of elements of B' with nonnegative rational coefficients in a unique way.*

The question of existence of such a Hamel-basis of the nonnegative numbers was raised by J. Z. Yao (verbally).

4. Theorem 1 shows that the general solution of (1) for nonnegative x, y cannot be constructed *in analogy* to Hamel's construction ([7]) of the general solution of (1) for all real x, y. We will show, however, that this can be done *in consequence* of Hamel's theorem. In fact, we shall prove that *every solution of* (1) *for nonnegative*

x, y can be extended to a solution of (1) *for all real x, y.* This can be done by *defining*

$$f(-x) = -f(x) \tag{7}$$

for $-x<0$. We have to prove that *the so extended $f(x)$ satisfies* (1) *for all real x, y if the original $f(x)$ $(x \geqq 0)$ satisfies* (1) *for all nonnegative x, y.*

We distinguish 6 cases:

1. $x \geqq 0$, $y \geqq 0$: there is nothing to be proved.

2. $x \geqq 0$, $y<0$, $x+y \geqq 0$. Then by (7) and by the validity of (1) for nonnegative variables:

$$f(x) = f((x+y)+(-y)) = f(x+y)+f(-y) = f(x+y)-f(y),$$

thus

$$f(x+y) = f(x)+f(y).$$

3. $x<0$, $y \geqq 0$, $x+y \geqq 0$: proof analogous to 2.

4. $x \geqq 0$, $y<0$, $x+y<0$. Again by (7) and by (1) being valid for nonnegative variables:

$$f(y) = -f(-y) = -f(-(x+y)+x) = -(f(-(x+y))+f(x)) = f(x+y)-f(x)$$

thus

$$f(x+y) = f(x)+f(y),$$

in this case too.

5. $x<0$, $y \geqq 0$, $x+y<0$: proof analogous to 4. And finally

6. $x<0$, $y<0$:

$$f(x+y) = -f(-x-y) = -f((-x)+(-y)) = -(f(-x)+f(-y)) =$$
$$= f(x)+f(y) \text{ q. e. d.}$$

On the other hand every solution of (1) for all real x, y eo ipso satisfies (1) also for all nonnegative x, y.

So we have proved the following

Theorem 2. *Every solution of Cauchy's functional equation*

$$f(x+y) = f(x)+f(y) \tag{1}$$

for nonnegative x, y can be extended by the definition

$$f(-x) = -f(x) \tag{7}$$

to a solution of (1) *for all real x, y.*

So every solution of (1) *for nonnegative x, y can be given by taking an arbitrary solution $f(x)$ of the functional equation* (1) *considered for all real x, y, and by restricting it to nonnegative x's.*

Everywhere in this note the word "*nonnegative*" *can be replaced by* "*positive*" only then in **4** also

$$f(0) = 0 \tag{6}$$

has to be considered as a *definition*.

The content of this little note falls also under the heading "functional equations on restricted domains" (see e. g. [8], [5], [9], [2], [10], [6]), which is a newer field of research in the theory of functional equations.

References

[1] J. Aczél, Vorlesungen über Funktionalgleichungen und ihre Anwendungen, *Basel—Stuttgart*, 1961.

[2] J. Aczél, Über die Begründung der Additions- und Multiplikationsformeln von bedingten Wahrscheinlichkeiten, *Publ. Math. Inst. Hung. Acad. Sci.* **5** (1961), 111—121.

[3] A. L. Cauchy, Cours d'analyse de l'École Polytechnique, 1. Analyse algébrique, V. *Paris* 1821. (Oeuvres (2)) 3. *Paris*, 1897, pp. 98—113, 220.)

[4] G. Darboux, Sur le théorème fondamental de la géométrie projective, *Math. Ann.* **17** (1880), 33—42.

[5] P. Erdős, Problem 310, *Colloq. Math.* **7** (1960), 311.

[6] S. Gołab—L. Losonczi, Über die Funktionalgleichung der Arccos-Funktion, I, in this journal,

[7] G. Hamel, Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung $f(x+y)=f(x)+f(y)$, *Math. Ann.* **60** (1905), 459—462.

[8] S. Hartman, Problème 203. *Colloq. Math.* **5** (1958), 236—237.

[9] S. Hartman, A remark about Cauchy's equation, *Colloq. Math.* **8** (1961), 77—79.

[10] W. B. Jurkat, On Cauchy's functional equation, *Proc. Amer. Math. Soc.* **16** (1965), 683—686.

(Received 28. September 1964.)