Research Papers

Remarks on One-Parameter Subsemigroups of the Affine Group and Their Homo- and Isomorphisms

Dedicated to the Memory of O. Varga on His 60th Birthday

J. Aczél (Waterloo, Ont., Canada) and St. Gołąb (Krakow, Poland)

The classical method of determining subgroups of a given continuous group of transformations, due to S. Lie, has an analytical character and therefore presupposes the regularity of the unknown functions. Another method in investigating the above mentioned problem is to solve the functional equations to which the group property (really the associativity alone) of the respective transformations lead. This can give irregular solutions or show that all solutions are regular, depending on the type of the equations obtained.

In this paper we explain this method on the example of determining certain one-parameter subsemigroups of the affine group of transformations

$$p \to \bar{p} = \alpha p + \beta \quad (\alpha \neq 0)$$

in a one-dimensional space, and then we determine all homomorphisms and, in particular, all isomorphisms between some pairs of subsemigroups thus found (cf. also [2], where all endomorphisms of the – two-parameter – affine group where determined). In the semigroups we permit also $\alpha = 0$

1. The general form of the transformations of a one-parameter subsemigroup would be

$$p \to \bar{p} = \alpha(u)\, p + \beta(u). \tag{1}$$

The supposition that this set of transformations is closed under superposition means that

$$\bar{\bar{p}} = \alpha(v)\,\bar{p} + \beta(v) = \alpha(v)\,\alpha(u)\, p + \alpha(v)\,\beta(u) + \beta(v) \tag{2}$$

and at the same time

$$\bar{\bar{p}} = \alpha(u \circ v)\, p + \beta(u \circ v) \tag{3}$$

where $u \circ v$ is the new parameter value resulting from the composition of the parameter values u and v.

Comparison of (2) and (3) leads to the system of functional equations

$$\alpha(u \circ v) = \alpha(v)\,\alpha(u) \tag{4}$$

$$\beta(u \circ v) = \alpha(v)\,\beta(u) + \beta(v). \tag{5}$$

Received November 20, 1967 and in revised form, December 31, 1968.

We will consider this rather difficult system of functional equations only when either α or β is injective. If the function α is injective, then $s=\alpha(u)$ can be introduced as new parameter and, with the notation $\phi(s)=\beta[\alpha^{-1}(s)]$, (1) goes over into

$$p \to \bar{p} = sp + \phi(s) \tag{6}$$

and (4), (5) into

$$\phi(st) = t\phi(s) + \phi(t). \tag{7}$$

If, on the other hand, β is injective, then we introduce $x=\beta(v)$ as new parameter and denote $\psi(x)=\alpha[\beta^{-1}(x)]$, so that (1) and (4), (5) become

$$p \to \bar{p} = \psi(x)\, p + x \tag{8}$$

and

$$\psi[x + y\psi(x)] = \psi(x)\,\psi(y), \tag{9}$$

respectively. We will restrict ourselves to transformations of the forms (6) and (8) and so to the functional equations (7) and (9).

Much of our considerations remains valid in more general commutative fields, but here we will deal with reals.

Now, it is very easy to solve equation (7). By the commutativity of multiplication we have

$$t\phi(s) + \phi(t) = \phi(st) = \phi(ts) = s\phi(t) + \phi(s),$$

or, by choosing a constant $t=t_0\neq 1$ and denoting $\gamma=\phi(t_0)/(t_0-1)$, we have

$$\phi(s) = \gamma(s-1) \tag{10}$$

which satisfies (7) for arbitrary constant γ.

Notice that we have not supposed that (7) is satisfied for all real s, t, only that the set S of (s, t) for which it is satisfied is *symmetric*: $(s, t)\in S \Rightarrow (t, s)\in S$ (and that there is a $t\neq 1$; but if $S=\{(1, 1)\}$ then (7) gives $\phi(1)=0$, so (10) is true also in this case). We summarize:

THEOREM 1. *All solutions of the functional equation* (7) *over a symmetric subset of the real plane, whose domains contain with each s and t also st, are given by* (10) *with arbitrary constant* γ.

Thus all one-parameter subsemigroups of form (6) *of the one-dimensional affine group are given by*

$$p \to sp + \gamma(s-1) \tag{11}$$

where γ *is some constant. The composition* (4) *of parameters is of course* $(s, t)\to st$.

So in this case *all* subsemigroups are analytic. If we are interested in *subgroups* then $s=0$ has to be excluded. The largest subgroup is corresponding to $(-\infty, \infty)\setminus\{0\}$. The *largest proper subgroup* of this is $(0, \infty)$. The *unit element* is 1.

The result (10) means that for $\beta = \phi\alpha$

$$\beta(u) = \gamma[\alpha(u) - 1],$$

thus *whenever* α *is injective and* $\gamma \neq 0$, *then also* β *is injective.* If $\gamma = 0$ we get from (11) $p \to sp$, that is, subsemigroups of *the centroaffine group.*

With equation (9) things are much more complicated, even if (9) is supposed valid for all pairs (x, y) of real numbers. In this case it was recently completely solved by S. Wołodźko [5], while all differentiable resp. all continuous solutions were found earlier ([1] and [3], respectively). All *differentiable* solutions are

$$\psi(x) = 0 \tag{12}$$

and

$$\psi(x) = 1 + \delta x \tag{13}$$

while all *continuous but not differentiable* solutions are of the form

$$\psi(x) = \begin{cases} 1 + \delta x & \text{for} \quad x \leqslant -1/\delta \\ 0 & \text{for} \quad x \geqslant -1/\delta \end{cases} \quad (\delta < 0) \tag{14}$$

or

$$\psi(x) = \begin{cases} 0 & \text{for} \quad x \leqslant -1/\delta \\ 1 + \delta x & \text{for} \quad x \geqslant -1/\delta \end{cases} \quad (\delta > 0), \tag{15}$$

where δ in (13), (14), and (15) is an arbitrary constant in $(-\infty, \infty)$, $(-\infty, 0)$, or $(0, \infty)$, respectively.

As it was shown in [3], the equation (9) has also non-continuous but measurable and many non-measurable solutions. The Dirichlet function

$$\psi(x) = \begin{cases} 1 & \text{for rational } x \\ 0 & \text{for irrational } x \end{cases}$$

is an example of a non-continuous, but measurable and bounded solution. This solution is trivial in the following sense: Equation (9) shows that with y_1 and y_2 in the range of ψ, also their product $y_1 y_2$ will be in the range of ψ. So the range is an infinite set, except if ψ does not take any values different from 0, 1, -1. These latter solutions which do not take any values different from 0, 1, -1 are called *trivial solutions.*

An example of a non-bounded, non measurable solution is

$$\psi(x) = 1 + a(x)$$

where a is an arbitrary additive function whose values are rational and not all zero on a Hamel-basis. (As it is well known, [2], any function given arbitrarily on a Hamel-basis can be extended to an additive function for the reals). If the value of a on an element $h_0 \neq 0$ of the Hamel-basis is chosen to be 0, then this solution has arbitrarily small periods, because for any (arbitrarily small) rational r

$$\psi(x + rh_0) = 1 + a(x + rh_0) = 1 + a(x) + ra(h_0) = 1 + a(x) = \psi(x).$$

Solutions which are periodic with arbitrarily small periods are called *microperiodic*.

In [3] also *all non trivial, non microperiodic solutions* of (9) have been determined. These are

$$\psi(x) = \begin{cases} 1 + \delta x & \text{if} \quad 1 + \delta x \in G \\ 0 & \text{if} \quad 1 + \delta x \notin G, \end{cases} \tag{16}$$

where δ is an arbitrary constant, while G is an arbitrary multiplicative subgroup of the multiplication group of real numbers which contains elements different from 1 and -1.

As the solutions [(12) and] (14), (15), (16) show, if β is injective, even bijective, $\alpha = \psi\beta$ might not be injective. But in (8), ψ should be different from 0, so we will consider only solutions which are 0 at most in one point. This excludes among others the solutions (12), (14), (15), and (16) if $G \neq (-\infty, \infty)$.

We prove here the following Theorem (cf. [1]):

THEOREM 2. *The solutions of the functional equation* (9) *over the real plane, which are* 0 *at most in one point, are given by* (13).

Thus all one-parameter subsemigroups of the form (8) *of the one-dimensional affine group with x running through the reals and* $\psi(x) = 0$ *for at most one x are given by*

$$p \to (1 + \delta x)\, p + x \tag{17}$$

where δ *is an arbitrary constant. The composition* (5) *of parameters is*

$$(x, y) \to x \circ y = x + y + \delta xy.$$

In the case of Theorem 2, we have

$$\alpha(u) = 1 + \delta\beta(u).$$

If $\delta = 0$, then we get from (1) $p \to p + x$, that is, *the group of translations. If* $\delta \neq 0$, *then,* β *being bijective, also* α *is bijective.*

Proof of Theorem 2. First of all, by putting $y = 0$ into (9) we have

$$\psi(x) = \psi(x)\,\psi(0),$$

so either we have (12), which is now excluded, or

$$\psi(0) = 1. \tag{18}$$

If $\psi(x)=1$ for all real x, then we have (13) with $\delta=0$. If however there exist x with

$$\psi(x) \neq 1, \tag{19}$$

then put into (9) such an x and

$$y = \frac{x}{1-\psi(x)}$$

in order to get

$$\psi\left(\frac{x}{1-\psi(x)}\right) = \psi(x)\,\psi\left(\frac{x}{1-\psi(x)}\right).$$

Because of (19)

$$\psi\left(\frac{x}{1-\psi(x)}\right) = 0.$$

But, by supposition, there exists at most one x_0 such that $\psi(x_0)=0$, so

$$\frac{x}{1-\psi(x)} = x_0$$

for all x satisfying (19). By (18), $x_0 \neq 0$, and writing $-1/x_0=\delta$, we get

$$\psi(x) = 1 + \delta x \quad \text{for all } x \text{ for which } \psi(x) \neq 1. \tag{20}$$

If there were $x \neq 0$ for which $\psi(x)=1$ and also $y \neq 0$ for which $\psi(y)=1+\delta y$, then for such x, y equation (9) states

$$\psi(x+y) = 1 + \delta y. \tag{21}$$

But, by (20) either $\psi(x+y)=1+\delta(x+y)$ or $\psi(x+y)=1$. In both cases, when compared with (21), we get $\delta=0$. Thus either $\psi(x)=1$ or $\psi(x)=1+\delta x$ for all x, that is (13) holds (by (18) also for $x=0$), what was to be proved. The rest of the statements is obvious.

It should be emphasized that Theorem 1 determines *all subsemigroups of the form* (6) – just take (11) with s in an arbitrary real set closed under multiplication – while Theorem 2 determines only those subsemigroups of the form (8), where x takes all real values and ψ is 0 at most in one point. These are given by (17). It is easy to see, that (13) *with* $\delta \neq 0$ *is the most general injective solution of* (9), *even if* (9) *is supposed only on an arbitrary symmetric set.*

If we are interested in *groups*, then the value $x=-1/\delta$ has to be excluded from the real line, so the *largest subgroup* is that corresponding to $(-\infty, \infty)\setminus\{-1/\delta\}$. The *largest proper subgroup* of this is that corresponding to whichever of the intervals $(-\infty, -1/\delta)$, $(-1/\delta, \infty)$ contains 0. The *unit element* of such a subgroup corresponds evidently to $x=0$.

We can say moreover, *if a subgroup contains a whole neighborhood of the unit element, then it has to contain the whole subgroup corresponding to whichever of the intervals* $(-\infty, -1/\delta)$, $(-1/\delta, \infty)$ *contains* 0 *and thus it is either this subgroup itself or that corresponding to the entire* $(-\infty, \infty)\setminus\{-1/\delta\}$. In fact, for (17), the composition is, as mentioned in Theorem 2, $x \circ y = x + y + \delta xy$, while the inverse element has to be defined by $x \circ x^{(-1)} = 0$, that is

$$x^{(-1)} = -\frac{x}{1 + \delta x}.$$

A subgroup has to contain $x^{(n)}$ for all integer n ($x^{(0)} = 0$, $x^{(n+1)} = x \circ x^{(n)}$, $n = 0, 1, 2, \ldots$, $x^{(-n)} = [x^{(-1)}]^{(n)}$). The limits z of $x^{(n)}$ as $n \to \pm\infty$, satisfy $z \circ z = z$. The finite fixed points of $z \to z^{(2)} = z \circ z = 2z + \delta z^2$ are $z = 0$ and $z = -1/\delta$, moreover $z = 0$ is a repulsive fixed point, while $z = -1/\delta$ and also ∞ and $-\infty$ are attractive fixed points (cf. [4]).

2. Now we turn to the determination of the homomorphisms between semigroups of the forms (11) and (17) (from (17) into (11) and from (11) into (17)).

If the subgroup (17) is homomorphic to a subgroup (11), then $x = f(s)$, $y = f(t)$ and

$$f(st) = f(s) + f(t) + \delta f(s) f(t). \tag{22}$$

In the case $\delta = 0$ we have

$$f(st) = f(s) + f(t). \tag{23}$$

If (23) is supposed for a set of *real, nonzero* s, t (including -1 and with s, t also st), then its general solution is

$$f(s) = a(\log|s|) \quad (s \neq 0) \tag{24}$$

where a is additive:

$$a(u + v) = a(u) + a(v). \tag{25}$$

If on the other hand (23) is supposed only for *positive* s, t, then the general solution is ([2])

$$f(s) = a(\log s) \quad (s > 0), \tag{26}$$

where a again satisfies (25). (The formula (24) follows from (23) and (26) by observing $f(1) = 0$, and $f(-1) = 0$ by putting $s = t = -1$ into (23) and then, for $s \neq 0$, $t = \operatorname{sign} s$ $f(s) = f(|s|) + f(\operatorname{sign} s) = f(|s|) = a(\log|s|)$: this is, why we have supposed that $t = -1$ is an admissible substitution in (23).)

Now take the cases $\delta \neq 0$. Multiply both sides of (22) by δ and add 1:

$$1 + \delta f(st) = 1 + \delta f(s) + \delta f(t) + \delta^2 f(s) f(t)$$

or with $m(s) = 1 + \delta f(s)$:

$$m(st) = m(s)\, m(t) \quad (st \neq 0). \tag{27}$$

The most general real solutions of (27) are (see [2])

$$m(s) = 0 \quad \text{and} \quad m(s) = e^{a(\log s)},$$

if only *positive* x, y are considered, but

$$m(x) = 0, \quad m(s) = e^{a(\log |s|)} \quad \text{and} \quad m(s) = e^{a(\log |s|)} \operatorname{sign} s,$$

if both *positive and negative* s, t are considered (including -1). So

$$f(s) = -1/\delta \tag{28}$$

or

$$f(s) = \frac{1}{\delta} e^{a(\log s)} - \frac{1}{\delta} \quad (s > 0) \tag{29}$$

in the first case and

$$f(s) = -1/\delta, \quad f(s) = \frac{1}{\delta} e^{a(\log |s|)} - \frac{1}{\delta} \tag{30}$$

or

$$f(s) = \frac{1}{\delta} e^{a(\log |s|)} \operatorname{sign} s - \frac{1}{\delta} \quad (s \neq 0) \tag{31}$$

in the second, where a again satisfies (25).

Thus we have proved the following Theorem.

THEOREM 3. *If subsemigroups of the form* (11) *contain only positive parameters* s, *then all their homomorphisms into the semigroup* (17) *have the forms* (26), (28) *or* (29), *while if both positive and negative* s *are considered (including* -1*), then they are all of the forms* (24), (30) *or* (31), *where* a *is an arbitrary function additive on the set of logarithms of the absolute values of the parameters* s.

Of these only (26), (29) *and* (31) *can be injective and these if and only if the values of* a *on the Hamel-basis elements (the exponentials of which are* s*-values belonging to the semigroup of transformations* (11)*) are (rationally) linearly independent. If in particular* s *in* (11) *runs through all positive or all nonzero numbers, then* (26), (29) *and* (31) *respectively with* a *as above give all isomorphisms with the group* (17) *on* $(-\infty, \infty)\setminus\{-1/\delta\}$.

If moreover f *is bounded in a neighborhood (or on a set of positive measure), then in all these formulas* $a(z)$ *should be replaced by* cz.

(The last statements follow from well-known results on the functional equation (25), see e.g. [2].)

If we look for homomorphisms g of the groups (17) into (11) (with s arbitrary), then we get in the same way

$$g(x + y + \delta xy) = g(x)\, g(y)$$

with *the general solutions*

$$g(x) = 0 \tag{32}$$

and

$$g(x) = e^{a(x)} \tag{33}$$

in the case $\delta = 0$, *while for* $\delta \neq 0$ *the general solutions are*

$$g(x) = 0, \tag{34}$$

$$g(x) = e^{a(\log|1+\delta x|)} \tag{35}$$

and

$$g(x) = e^{a(\log|1+\delta x|)} \operatorname{sign}(1 + \delta x), \tag{36}$$

if (17) *is taken for all* x *in* $(-\infty, \infty) \setminus \{-1/\delta\}$, *while all solutions are given by* (34) *and* (35) *if* (17) *is taken only for that one of the intervals* $(-\infty, -1/\delta)$, $(-1/\delta, \infty)$ *which contains* 0.

Of these, if the values of a on the Hamel basis are linearly independent, then (33) *in both cases,* (36) *in the case of two intervals and* (35) *in the case of one interval are injective, but only* (36) *is bijective* [(33), (35) *are not surjective*] and so only (36) gives an isomorphism – the same as (31).

In case where $\gamma\delta = 1$, *the subgroups described by* (11) *and* (17) *are* not only isomorphic, but, up to notations, *identical.*

Theorem 3 took care also of the case where only subsemigroups of the groups (11) were considered, that is, essentially subsemigroups of the multiplicative group of real numbers. Now there can also be homo- and isomorphisms between different subsemigroups of this multiplicative group. We show how functional equations work in determining them on an example.

The nonzero rational numbers evidently form a group under multiplication. Let us look for homomorphisms of this group into other subsemigroups of the multiplicative semigroups of reals. We evidently have to do with the functional equation [cf. (27)]

$$m(xy) = m(x)\, m(y) \quad x, y \text{ nonzero rationals}. \tag{37}$$

In order to solve this, first notice that, except for the trivial solution

$$m(x) = 0, \tag{38}$$

(37) implies with $y = 1$

$$m(1) = 1 \tag{39}$$

and, with $x = y = -1$, either

$$m(-1) = 1 \tag{40}$$

or

$$m(-1) = -1. \tag{41}$$

Notice also that if $m(x) = 0$ for one $x \neq 0$ then it is always 0, i.e. we have (38) again, because $m(z) = m(x \cdot z/x) = m(x)\, m(z/x) = 0$ for all z.

Now we express the values of m on the integers greater than 1 with aid of its values on the primes: As (37) evidently implies

$$m(x_1 x_2 \dots x_k) = m(x_1)\, m(x_2) \dots m(x_k)$$

and

$$m(x^k) = m(x)^k,$$

so,

$$\left.\begin{array}{l} \text{for} \quad n = p_1^{k_1} p_2^{k_2} \dots p_j^{k_j} \quad (n > 1, \text{integer}, p_1, \dots, p_j \text{ primes}) \\ m(n) = m(p_1)^{k_1} m(p_2)^{k_2} \dots m(p_n)^{k_n}. \end{array}\right\} \tag{42}$$

For positive rational r we get $m(r)$ by observing that (37) and (39) imply

$$m(1/y)\, m(y) = 1 \quad \text{and} \quad m(x/y) = m(x)/m(y).$$

So

$$m(r) = m\left(\frac{n_1}{n_2}\right) = \frac{m(n_1)}{m(n_2)} \quad (n_1, n_2 \text{ positive integers}). \tag{43}$$

Finally, for negative rationals we get in the case (40)

$$m\left(-\frac{n_1}{n_2}\right) = m\left(\frac{n_1}{n_2}\right) = \frac{m(n_1)}{m(n_2)} \tag{44}$$

and in the case (41)

$$m\left(-\frac{n_1}{n_2}\right) = -m\left(\frac{n_1}{n_2}\right) = -\frac{m(n_1)}{m(n_2)}. \tag{45}$$

Reciprocally, it is easy to see, that choosing the values of m arbitrarily, but different from 0, on the positive primes, the functions defined by (39), (42), (43) and either (44) or (45) all satisfy the functional equation (37). We have yet to state that the definition (43) is unambiguous, because

$$m\left(\frac{kn_1}{kn_2}\right) = \frac{m(kn_1)}{m(kn_2)} = \frac{m(k)\, m(n_1)}{m(k)\, m(n_2)} = \frac{m(n_1)}{m(n_2)} = m\left(\frac{n_1}{n_2}\right).$$

So we have proved the following Theorem:

THEOREM 4. *The general real solution m of* (37), *which is not identically zero, is given by choosing the values of m for positive prime arguments arbitrarily different from zero and extending its definition by the formulas* (39), (42), (43) *and either* (44) *or* (45). *– These give thus all homomorphisms of the multiplicative group of rationals into other multiplicative subsemigroups (subgroups) of the multiplicative group of reals.*

If and only if the logarithms of the absolute values of the values of m chosen on the positive primes are linearly independent (integer coefficients) and the further definition of m happens by (39), (42), (43) *and* (45), *then m is injective and thus an isomorphism onto its range.*

Such isomorphisms evidently can not exist if the considered second multiplicative group is not countable. But by Theorem 4 this *second, image group* also *can not consist* of the powers of one number or *of products of powers of a finite number of integers (and possibly their negatives) alone*, because then there would be only one resp. a finite number of linearly independent numbers (in the sense of integer coefficients) in the range of $\log|m|$ and not infinitely many (for all $\log|m(p_j)|$, p_j running through all primes).

We see that in Theorem 4 the logarithms of the prime numbers take the role previously played by the Hamel bases (they are bases for the logarithms of the positive integers in the sense of integer coefficients).

We wish to finish these remarks by stating that we did not aim at completeness nor absolute novelty, we only wanted to show how a general method works in a rather broad field.

REFERENCES

[1] ACZÉL, J., *Beiträge zur Theorie der geometrischen Objekte*, III–IV, Acta Math. Acad. Sci. Hung. *8*, 19–52 (1957).

[2] ACZÉL, J., *Lectures on Functional Equations and Their Applications* (Academic Press, New York–London 1966 [Mathematics in Science and Engineering, Vol. 19]), Sections 2.1.1, 2.1.2 and 5.1.1.

[3] GOŁĄB, St. and SCHINZEL, A., *Sur l'équation fonctionnelle* $f[x+yf(x)]=f(x)f(y)$, Publ. Math. Debrecen *6*, 113–125 (1959).

[4] KUCZMA, M., *Functional Equations in a Single Variable* (PWN – Polish Scientific Publishers, Warszawa 1968 [Monografie Mat., Vol. 46]).

[5] WOŁODŹKO, S., *Solution générale de l'équation fonctionnelle* $f[x+yf(x)]=f(x)f(y)$, Aequationes Mathematicae *2*, 12–29 (1968–69).

University of Waterloo
Jagello University