SOLUTIONS OF THE FUNCTIONAL EQUATION

$$(f(x))^2 - f(x^2) = h(x)$$

I.N. Baker

(received April 16, 1959)

Introduction. In a recent paper [2] Lambek and Moser have introduced the functional equations

$$(f(x))^2 - f(x^2) = h(x) \tag{1}$$

and

$$(f(x))^2 - f(2x) = h(x) \tag{2}$$

in connection with some problems of number theory, in particular in dealing with the sums by pairs of sets of integers. The second may be put into the same form as (1) by the substitutions $x = \ln z$, $f(\ln z) = F(z)$, $h(\ln z) = H(z)$.

In [2] solutions of (1) of a special form were discussed, e.g., $f(x)$ a polynomial with coefficients 0 or 1. If one is interested in more general solutions natural starting points are formal expansions for $f(x)$ about $x = 0$ or 1 which are the fixpoints of the function x^2. We shall assume $h(x)$ to be a polynomial except for theorem 1 where regularity at the origin is enough. Convergence of the formal series under certain assumptions is proved (theorem 1) but in general these series are not polynomials (§ 2), nor are the expansions about the different fixpoints in general elements of the same analytic configuration.

1. The existence of solutions of (1) regular at the origin. The assumption of a solution $f(x)$ of (1) in the form

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \tag{3}$$

https://doi.org/10.4153/CMB-1960-012-8 Published online by Cambridge University Press

yields two values for a_o. If $h(0) \neq 0$ the values of a_o are not 0; if $h(0) = 0$ then one has $a_o = 0$ or $a_o = 1$. If $h(0) = -1/4$ the two roots a_o reduce to the single value $\frac{1}{2}$. We prove

THEOREM 1. If $h(x) = \beta_o - \beta_1 x - \beta_2 x^2 \ldots\ldots$ is a power series with a positive radius of convergence and if a_o is a non-zero root of

$$a_o^2 - a_o = \beta_o ,$$

then there is a unique series (3) which solves (1) and has a positive radius of convergence.

Thus if $\beta_o \neq 0, -1/4$ there are just two solutions of the form (3). If $\beta_o = -1/4$ there is just one. If $\beta_o = 0$ there is one solution with constant term $a_o = 1$ and if $\beta_o = \beta_1 \ldots = \beta_{r-1} = 0$, $\beta_r \neq 0$, there are further solutions of the form

$$f(x) = x^k \sum_{n=o}^{\infty} a_n x^n ,$$

where k is an integer satisfying $1 \leqslant k \leqslant [\frac{1}{2}r]$.

Before proving the theorem we derive the following

LEMMA 1. For fixed $q \geqslant 3$ and all sufficiently large n

$$\sum_{j=q}^{n-q} j^{-2}(n-j)^{-2} < 3(q-1)^{-1}n^{-2}$$

Proof.

$$\sum_{j=q}^{n-q} j^{-2}(n-j)^{-2} < \int_{q-1}^{n-q+1} x^{-2}(n-x)^{-2}dx = 2(q-1)^{-1}n^{-2} + o(n^{-2}).$$

Proof of theorem 1. Since the series for h(x) has a positive radius of convergence there is a $K > 0$ such that $|\beta_n| < K^n$ for $n = 1, 2, \ldots\ldots$. Put (3) in (1). The following recursions are found.

(4) $$a_o^2 - a_o = \beta_o$$

(5) $$2a_o a_1 = - \beta_1$$

https://doi.org/10.4153/CMB-1960-012-8 Published online by Cambridge University Press

and in general

(6) $$2a_0 a_n = -\beta_n - \sum_{j=1}^{n-1} a_j a_{n-j} + (a_{\frac{1}{2}n})$$

where the term $(a_{\frac{1}{2}n})$ occurs only if n is even. We show that for a suitable finite $R > 0$

(7) $$|a_n| < R^n/n^2.$$

We choose an integer $q \geq 3$ with $(q-1)|a_0| \geq 3$, m so large that lemma 1 is satisfied for all $n > m$ and R so that

(a) (7) holds for $n = 1, 2, 3, \ldots, m$,

(b) $|a_0| (m-j)^2 R^j > 2(q+1) |a_j| m^2$ for $j = 1, 2, \ldots, q-1$,

(c) $(q+1) n^2 K^n < R^n |a_0|$ for all $n > m$,

and

(d) $4 R^{-n/2} < |a_0| (q+1)^{-1}$ for all $n > m$.

Then by induction (7) holds for all n. For suppose (7) holds for $1, 2, \ldots, n-1$. From (6) it follows that

$$2|a_0 a_n| \leq |\beta_n| + 2 \left|\sum_{j=1}^{q-1} a_j a_{n-j}\right| + \left|\sum_{j=q}^{n-q} a_j a_{n-j}\right| + |a_{\frac{1}{2}n}|$$

$$\leq K^n + 2 \sum_{j=1}^{q-1} |a_j| R^{n-j} (n-j)^{-2}$$

$$+ R^n \sum_{j=q}^{n-q} j^{-2} (n-j)^{-2} + 4 R^{\frac{1}{2}n}/n^2$$

$$< |a_0| R^n n^{-2} (q+1)^{-1} \{1 + (q-1) + 3(q+1) |a_0|^{-1} (q-1)^{-1} + 1\}$$

$$\leq 2|a_0| R^n/n^2$$

and the induction is complete. In the above estimation we have used (b) as follows:

$$2|a_j| R^{n-j} (n-j)^{-2} < 2|a_j| R^{n-j} n^{-2} m^2 (m-j)^{-2} < |a_0| R^n n^{-2} (q+1)^{-1}$$

in dealing with the second term on the right.

https://doi.org/10.4153/CMB-1960-012-8 Published online by Cambridge University Press

The above analysis shows that the series (3) is determined completely by a_0 and converges in the circle $|x| < R^{-1}$.

2. Entire solutions when h(x) is a polynomial. From now on h(x) will be assumed to be a polynomial. The next question we consider is the nature of the solutions whose existence is proven by theorem 1. A first result is that they are certainly not entire transcendental functions.

THEOREM 2. If h(x) is a polynomial any entire solution of (1) is a polynomial.

Proof. Suppose a solution f(x) of (1) is entire and transcendental. Let F(r) be the maximum modulus of f(x) for x = r. From (1) it follows that

$$F(r^2) = \text{Max}_{|x|=r} |(f(x))^2 - h(x)|$$

$$< (F(r))^2 + H(r)$$

where $H(r) = \text{Max}_{|x|=r} |h(x)|$. Let the degree of h(x) be n. There is a positive constant A such that

$$H(r) < A r^n$$

and

$$(F(r))^2 > A r^n$$

hold simultaneously for all sufficiently large r so that for such r values one has

$$F(r^2) < 2(F(r))^2 . \tag{8}$$

But, as in [1, p. 140, Hilfssatz 4], Hadamard's theorem on the convexity of log F(r) as a function of log r shows that such an inequality as (8) can hold for no entire transcendent function. This completes the proof of theorem 2. One may restate it in the form

THEOREM 2'. If h(x) is such that (1) has no polynomial solution f(x) then there is no entire solution at all.

The occurrence of polynomial solutions is often excluded. From

$$(f(x))^2 - f(x^2) = h(x)$$

follows

https://doi.org/10.4153/CMB-1960-012-8 Published online by Cambridge University Press

$$(f(-x))^2 - f(x^2) = h(-x)$$

and by subtraction of these two equations

$$\{f(x) + f(-x)\} \{ f(x) - f(-x)\} = h(x) - h(-x) .$$

Put

$$f_0(x) = \text{even part of } f(x) = \tfrac{1}{2}[f(x) + f(-x)]$$

$$f_1(x) = \text{odd part of } f(x) = \tfrac{1}{2}[f(x) - f(-x)]$$

$$h_1(x) = \text{odd part of } h(x) = \tfrac{1}{2}[h(x) - h(-x)] .$$

Then

(9) $$f_0(x)\, f_1(x) = \tfrac{1}{2}\, h_1(x) .$$

To obtain a case where (1) has no polynomial solutions take $h(x)$ of degree at least $2n+1$, n an integer > 1, with $h_1(x) \not\equiv 0$ but of degree at most n. Suppose that (1) has a polynomial solution $f(x)$. Then from (9) neither $f_0(x)$ nor $f_1(x)$ can vanish identically and the degree of each and thus of $f(x)$ is at most n. Then in (1) the left hand side has degree at most $2n$ and the right has degree at least $2n+1$, which affords a contradiction.

Thus cases exist where $h(x)$ is a polynomial but (1) has no polynomial solutions $f(x)$. If $h(0) \neq 0$ the recursive calculation of theorem 1 shows that there are at most two polynomial solutions since there are only two solutions regular at 0.

3. Analytic continuation of the solutions when these are not entire. The radius of convergence of the expansions in theorem 1 cannot be greater than 1 unless they are entire. Suppose the exact radius of convergence is ρ with $1 < \rho < \infty$. Then one may use the functional equation (1) in the form

(10) $$f(x) = -h(x^{\frac{1}{2}}) + (f(x^{\frac{1}{2}}))^2$$

to continue $f(x)$ - say along radial paths emanating from the origin - throughout the circle $|x| < \rho^2$. Since $\rho^2 > \rho$ this is a contradiction and the assumption $\rho > 1$ is false.

Moreover if $f(x)$ is not a polynomial it is a (possibly many valued) function which can be continued along some radial paths from the origin to points arbitrarily close to the unit circle but it is impossible to continue $f(x)$ radially over

https://doi.org/10.4153/CMB-1960-012-8 Published online by Cambridge University Press

any point on $|x| = 1$. For if ρ is the exact radius of convergence of the expansion of theorem 1 we have $\rho \leq 1$.

(a) $\rho = 1$. In this case the unit circle forms a natural boundary for if $f(x)$ is regular at $x = e^{i\theta}$ then it is regular in a whole sector of the form

$$\{ |x| < \rho', \rho' > 1, |\arg x - \theta| < 1/n \}$$

and by repeated application of

$$f(x^2) = (f(x))^2 - h(x)$$

we may obtain sectors of regularity of $f(x)$ of opening $2/n$, $4/n$, $8/n$, ... and so eventually continue $f(x)$ into the full circle $|x| < \rho'$, $\rho' > 1$ which is impossible.

(b) Suppose $\rho < 1$. The argument of case (a) shows that $f(x)$ cannot be continued along a radial path over any point on the unit circle. Further if $f(x)$ can even be continued to the boundary of the unit circle in a whole sector

$$\{ |x| < 1, |\arg x - \theta| < \varepsilon \}$$

the same argument would give a continuation throughout the whole circle $|x| < 1$ and so $\rho = 1$. Thus if $\rho < 1$ the radial distribution of the singularities met on radial continuation from $x = 0$ is dense in the neighbourhood of $|x| = 1$.

At least one singularity x_o lies on $|x| = \rho < 1$. (1) shows that $(f(x))^2$ is regular at x_o; x_o is a branch point of order 1 for $f(x)$ which interchanges two branches $\pm f(x)$ there. One has $f(x_o) = 0$ or

$$f(x_o^2) + h(x_o) = 0$$

so that x_o^2 is a root of

$$f(t) + h(t^{\frac{1}{2}}) = 0$$

lying in $|t| \leq \rho^2$ so there is at most a finite set of possible x_o with $|x_o| = \rho$. Excluding the rays $\{ \arg x = \arg x_o, |x| \geq |x_o| \}$ we can continue $f(x)$ further by radial continuation using (1) in the form

$$f(x) = [f(x^2) + h(x)]^{\frac{1}{2}}$$

https://doi.org/10.4153/CMB-1960-012-8 Published online by Cambridge University Press

and excluding (i) rays of x values for which x^2 falls on rays already excluded, (ii) rays to the boundary from any fresh branch points which appear. The number of rays (i) and (ii) intersecting any circle $|x| \leq \rho_1 < 1$ is finite.

The question arises whether it might not be possible to continue f(x) outside the unit circle by a non-radial path. If in particular one considers continuation over the point 1 it is interesting to note that two formal expansions

$$\sum_{n=o}^{\infty} a'_n (x-1)^n$$

satisfying (1) may be obtained by direct calculation. For a'_o one has

$$(a'_o)^2 - a'_o = h(1)$$

and from a'_o all subsequent a'_n are determined by the recursion, unless a'_o has the form 2^m, m a positive integer, when the recursion breaks down. We prove

THEOREM 3. If $g(x) = \sum_{n=o}^{\infty} a'_n (x-1)^n$ is a power series satisfying (1) formally then its radius of convergence σ is either 0, 1 or ∞ and in the last case it is a polynomial. If $\sigma = 1$ the continuation of g(x) is restricted only by the point $x = 0$ and g(x) is not part of the same analytic configuration as the solutions in theorem 1.

Proof. From theorem 2, g(x) is a polynomial if $\sigma = \infty$. Suppose $\sigma \neq 0, \infty$. g(x) certainly converges in a neighbourhood of $x = 1$ and satisfies (1) there. By repeated use of (1) in the form (10) we see that g(x) may be continued into a "sector" of the form first of

$$\{1 - \varepsilon < |x| < 1 + \varepsilon \text{ , } |\arg x| < \varepsilon, \varepsilon > 0\}$$

and then

$$\{(1-\varepsilon)^{2^n} < |x| < (1+\varepsilon)^{2^n} \text{ , } |\arg x| < \varepsilon\}$$

and thus

$$\{0 < |x| < \infty, \ |\arg x| < \varepsilon\}.$$

By further use of (10) g(x) may be continued along any circular path around $x = 0$ paying due care to the appropriate branches of x, $x^{\frac{1}{2}}$, $\sqrt{g(x)}$, $g(x^{\frac{1}{2}})$. The radius of convergence σ is therefore at least 1. If σ were greater than 1 direct transformation of the original expansion of g(x) about $x = 1$ to the new centre $x = 0$ would give a solution of (1) of the form $\sum_{n=o}^{\infty} b_n x^n$ and

https://doi.org/10.4153/CMB-1960-012-8 Published online by Cambridge University Press

thus one of the expansions discussed in sections 1, 2 and the beginning of section 3. This expansion would be radially continuable (along the positive real axis) beyond the unit circle, which from the first part of this section is possible only if g(x) is a polynomial, i.e. $\sigma = \infty$. Thus $\sigma = 0$, 1 or ∞. It is also clear that in case $\sigma = 1$, i.e. when $x = 0$ is the only finite singularity of g(x), it is impossible to have any regular element of the function g(x) over $x = 0$.

4. Generalization. As a generalization of (1) one could introduce the equation

$$g(f(x)) - f(g(x)) = h(x) \tag{11}$$

where g(x) and h(x) are given polynomials and f(x) is to be found. A variety of formal expansions about the fixpoints of g(x) could be found. A fixpoint ξ of g(x) is a point with $g(\xi) = \xi$. The number of expansions about a given fixpoint ξ would in general be the same as the degree of g(x) and this would also equal the number of different ξ if all fixpoints were distinct. Theorem 2 continues to hold; there are no entire transcendent solutions f(x). If f(x) is a convergent expansion but not a polynomial its continuation will be a complicated process depending on the iteration properties of g(x).

REFERENCES

1. I.N. Baker, Zusammensetzungen ganzer Funktionen, Math. Z. 69 (1958), 121-163.

2. J. Lambek and L. Moser, On some two-way classifications of integers, Canad. Math. Bull. 2 (1959), 85-89.

University of Alberta
and
Imperial College of Science, London.