Electronics Research Laboratory
University of California
Berkeley, California

ON THE CONVERGENCE OF A SERIES SOLUTION OF CERTAIN NONLINEAR DIFFERENCE EQUATIONS

by

A. G. Dewey

This research was supported by the Air Force Office of Scientific Research under Contract No. AF 49(628)-102.

July 11, 1963

ACKNOWLEDGEMENT

The author wishes to express his gratitude to Professor E. I. Jury for his useful comments and discussions during the course of this work.

SUMMARY AND REVIEW OF LITERATURE

This paper provides the necessary justification of the convolution z-transform method applied to nonlinear sampled-data systems proposed by Jury and Pai.[1, 2] The method is analogous to the method of applying the convolution Laplace transform to nonlinear differential equations, developed by Weber.[3] Rigorous mathematical justification of Weber's method was provided by Wasow[4] and further generalized in a paper by Golomb.[5]

The formal procedure for finding a series solution of certain nonlinear difference equations, and the necessary assumptions involved, were developed by Pai.[4] However, Pai did not prove the convergence of the solution and merely pointed out the final conclusion, i. e., that the initial state of the system should be sufficiently close to the equilibrium state.

The convergence of the solution is studied by the method of dominating series in a manner completely analogous to that used by Wasow in the case of nonlinear differential equations.

I. INTRODUCTION

We will consider a system of autonomous difference equations of the form

$$\underline{x}(n+1) = \underline{F}(\underline{x}(n)) \qquad n = 0, 1, \ldots \tag{1.1}$$

where $\underline{x}(n)$ is an r-dimensional vector.

Any r th-order scalar difference equation

$$\sum_{p=o}^{r} c_p \, x(n+p) + F\left[x(n),\ x(n+1), \ldots\ x(n+r) \right] = 0 \tag{1.2}$$

can be put into the vector form (1.1) by introduction of the variables:

$$\underline{x}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_r(n) \end{bmatrix} \qquad \text{where} \qquad \begin{array}{l} x_1(n) = x(n) \\ x_2(n) = x(n+1) \\ \vdots \\ x_r(n) = x(n+r) \end{array}$$

Under the assumptions to be stated in Section II, we can construct a solution of Eq. (1.1) of the form

$$\underline{x}(n) = \underline{\xi} + \sum_{p=1}^{\infty} \underline{u}_p \, a_p^n \tag{1.3}$$

where $\underline{\xi}$ is a solution of the equation

$$\underline{\xi} = \underline{F}[\underline{\xi}]$$

a_p $(p = 1, 2, \ldots\ r)$ can be determined from the original equation and a_p $(p = r+1,\ r+2, \ldots)$ are certain combinations of the form $a_1^{\alpha_1}\ a_2^{\alpha_2} \ldots a_r^{\alpha_r}$ where $\left\{\alpha_i\right\}_1^r$ are all nonnegative integers and $\alpha_1 + \alpha_2 + \ldots + \alpha_r \geq 2$. $\underline{u}_p$ $(p = r+1,\ r+2, \ldots)$ are vectors which can be determined in terms of $\underline{u}_p$ $(p = 1, 2, \ldots, r)$ which, in turn, are functions of r scalar parameters $\sigma_1, \sigma_2, \ldots \sigma_r$. These parameters are then finally determined when the initial state of the system $\underline{x}(0)$ is specified. It will be shown that a convergent solution of the form (1.3) can be found when the initial vector is sufficiently close to $\underline{\xi}$.

II. THE FORMAL PROCEDURE

We start by stating certain assumptions:

ASSUMPTION 1. The equation (1.1) possesses a solution $\underline{\xi}$ given by $\underline{\xi} = \underline{F}[\underline{\xi}]$; i.e. $\underline{\xi}$ is the equilibrium state of the system. If there is more than one such value of $\underline{\xi}$, then there will be as many solutions of the form (1.3), providing other assumptions hold.

Definition. If $\underline{v}$ is a vector with components $\left\{v_i\right\}_1^n$, the symbol $\|\underline{v}\|$ will denote the norm

$$\|\underline{v}\| = \max_i |v_i|$$

For a matrix $\underline{A}$ with components a_{ij} the symbol $\|\underline{A}\|$ will denote the norm

$$\|\underline{A}\| = \max_i \sum_{j=1}^{n} |a_{ij}|$$

Then clearly $\|\underline{A}\,\underline{v}\| \leqslant \|\underline{A}\|\,\|\underline{v}\|$

ASSUMPTION 2. The components $f_i[\underline{x}(n)]$ of $\underline{F}[\underline{x}(n)]$ are analytic functions in the components $x_1(n), x_2(n) \ldots x_r(n)$ of vector $\underline{x}(n)$ in the region

$$\|\underline{x}(n) - \underline{\xi}\| \leqslant \rho$$

where ρ is a positive constant.

Expand $\underline{F}[\underline{x}(n)]$ around the equilibrium point $\underline{\xi}$ and let $\underline{A}$ denote the Jacobian of the components f_i of $\underline{F}$ w.r.t. the components $\left\{x_j\right\}_1^r$ of $\underline{x}(n)$ evaluated at $\underline{x}(n) = \underline{\xi}$.

I.e., $\underline{A} = [a_{ij}]$ $r \times r$

where $a_{ij} = \left.\dfrac{\partial f_i[\underline{x}(n)]}{\partial x_j}\right|_{\underline{x}(n) = \underline{\xi}}$

Then denoting

$$\underline{x}(n) - \underline{\xi} = \underline{y}(n) \qquad (2.1)$$

we have $\underline{y}(n+1) + \underline{\xi} = \underline{F}(\underline{\xi}) + \underline{A}\,\underline{y}(n) + \underline{g}[\underline{y}(n)]$

where the components $g_i \,{}_1^r$ of $\underline{g}[\underline{y}(n)]$ possess series expansions in powers of $\{y_i\}_1^r$ without constant or linear terms.

These series converge in the domain

$$\|\underline{y}(n)\| \leqslant \rho \tag{2.2}$$

So the equation reduces to

$$\underline{y}(n+1) = \underline{A}\,\underline{y}(n) + \underline{g}[\underline{y}(n)]$$

where $\underline{A}$ is a constant r x r matrix.

ASSUMPTION 3. The eigenvalues of $\underline{A}$, $a_1, a_2, \ldots, a_r$ are all distinct and have modulus less than unity:

$$1 > |a_1| > |a_2| > \ldots |a_r|$$

We also assume that

$$|a_r| > |a_1|^2$$

This ensures proper ordering of the a_p's ($p = 1, 2, \ldots$) according to their magnitude. In fact, $a_1, a_2, \ldots, a_r$ are the poles of the generating function for the scalar nonlinear difference equation (1.2). This fact will be used in the application of this method.

The difference equation (2.3) can be formally satisfied by a series of the form

$$\underline{y}(n) = \sum_{p=1}^{\infty} \underline{u}_p \, a_p^n \tag{2.4}$$

where the a_p's are ordered in decreasing order of magnitude. By virtue of Assumption 3, the first r of these will be the eigenvalues of the matrix $\underline{A}$. The succeeding a_p's ($p = r+1, r+2, \ldots$) are of the form

$a_1^{\alpha_1} \; a_2^{\alpha_2} \; \ldots . \; a_r^{\alpha_r}$ where $\left\{\alpha_i\right\}_1^r$ are all nonnegative integers with $\alpha_1 + \alpha_2 + \ldots + \alpha_r \geq 2.$

Since all the eigenvalues of $\underline{A}$ have modulus less than unity, all the a_p's lie inside the unit circle.

Now substitute (2. 4) into Eq. (2. 3)

$$\sum_{p=1}^{\infty} \underline{u}_p \, a_p^{n+1} = \sum_{p=1}^{\infty} \underline{A} \, \underline{u}_p \, a_p^n + \underline{g}\left[\sum_{p=1}^{\infty} \underline{u}_p \, a_p^n\right]$$

and compare the coefficients of like powers of a_p^n.

For p = 1, 2, ..., r we get

$$\left(\underline{A} - \underline{I} \, a_p\right) \underline{u}_p = 0 \quad p = 1, 2, \ldots, r \tag{2.5}$$

Hence $\underline{u}_p$ is an eigenvector of $\underline{A}$ corresponding to the eigenvalue a_p. Let $\underline{u}_p = \sigma_p \underline{e}_p$ $(p = 1, 2, \ldots, r)$ where the σ_p's are scalar parameters to be determined and the $\underline{e}_p$'s are unit vectors, linearly independent, in the direction of the $\underline{u}_p$'s.

For p = r+1, r+ 2,.... we get

$$\left(\underline{A} - \underline{I} \, a_p\right) \underline{u}_p = \underline{h}_p\left(\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_{p-1}\right) \; p > r \tag{2.6}$$

where the components of the vector function $\underline{h}_p$ are polynomials in the components of $\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_{p-1}$ without constant or linear terms. Because of the manner in which the a_p's have been ordered, the right side of (2.6) can always be expressed in terms of the preceeding (p-1) vectors.

From Assumption 3, $a_{r+1}, a_{r+2}, \ldots$ are not eigenvalues of $\underline{A}$ and hence the matrices $\left(\underline{A} - \underline{I} \, a_p\right)$ $p > r$ are nonsingular. Clearly then we can compute $\underline{u}_{r+1}, \underline{u}_{r+2}, \ldots$ successively in terms of $\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_r$ — i. e., finally in terms of the r scalar parameters $\sigma_1, \sigma_2, \ldots, \sigma_r$.

We now have a solution for $\underline{x}(n)$ in (1.3) expressed in terms of $\sigma_1, \sigma_2, \ldots \sigma_r$. This must produce a convergent series for all values of n ($n \geq 0$).

For any given problem, the parameters are specified once the initial vector $\underline{x}(0)$ is specified. From Eq. (2.4) we have

$$\underline{x}(0) - \underline{\xi} = \underline{y}(0) = \sum_{p=1}^{\infty} \underline{u}_p \tag{2.7}$$

This gives r nonlinear relations involving $\sigma_1, \sigma_2, \ldots, \sigma_r$ which can then be determined.

III. THE CONVERGENCE OF THE SERIES

The convergence of the series (2. 4) will be studied by the method of dominating series.

From above, the matrices $(\underline{A} - \underline{I}\, a_p)$, $p > r$ are nonsingular, so let

$$c = \sup_{q>r} \left\| (\underline{A} - \underline{I}\, a_q)^{-1} \right\| < \infty$$

Then from Eq. (2. 6)

$$\left\| \underline{u}_p \right\| \leqslant c \left\| \underline{h}_p (\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_{p-1}) \right\| \quad p > r \tag{3.1}$$

As described in Wasow's paper, a function $\hat{h}_p$ will be constructed which dominates the r components of the vector function $\underline{h}_p$ for $p = r+1, r+2, \ldots$.

Referring to Eq. (2. 3), let M be some upper bound for $\left\| \underline{g} [\underline{y}(n)] \right\|$ in the domain defined by Eq. (2. 2).

Then we know, from the theory of analytic functions, that the coefficients of the terms of degree k in the power series for a component g_j of $\underline{g}$ are numerically not greater than $M\rho^{-k}$.

Hence the series

$$M \sum_{s_1 + s_2 + \ldots + s_r \geqslant 2} \rho^{-(s_1+s_2+\ldots+s_r)} y_1^{s_1} y_2^{s_2} \ldots y_r^{s_r}$$

represents a function that dominates all the g_j.

The series $M \sum_{s=2}^{\infty} \left(\dfrac{y_1+y_2+\ldots+y_r}{\rho}\right)^s$ (3. 2)

when expanded by the multinomial theorem has still larger positive coefficients. Hence the scalar function

$$\hat{g}(\underline{y}) = M \left[\left(1 - \rho^{-1} \sum_{j=1}^{r} y_j \right)^{-1} - 1 - \rho^{-1} \sum_{j=1}^{r} y_j \right] \tag{3.3}$$

whose power series expansion is (3. 2), dominates all the g_j. So if the vector $\underline{y}$ in (3.3) is replaced formally by the vectional series

$$\underline{v}(n) = \sum_{p=1}^{\infty} \underline{b}_n \, a_p^n \tag{3.4}$$

and the resulting products are expanded and rearranged as before, we obtain the scalar series

$$\hat{g}(\underline{v}) = \sum_{p=r+1}^{\infty} \hat{h}_p \, (\underline{b}_1, \underline{b}_2 \ldots \underline{b}_{p-1}) \, a_p^n \tag{3.5}$$

where the coefficients $\hat{h}_p$ are polynomials in the components of $\underline{b}_1, \underline{b}_2, \ldots \underline{b}_{p-1}$ without constant or linear terms. These polynomials dominate the r components of the vector function $\underline{h}_p \, (\underline{b}_1, \underline{b}_2 \ldots \underline{b}_{p-1})$ in the sense that the coefficients of $\hat{h}_p$ are positive and not less than the moduli of the corresponding coefficients of the components of $\underline{h}_p$.

Furthermore, if all components of $\underline{b}_p$ are equal, i.e., $\underline{b}_p = \beta_p \underline{e}_o$ where $\underline{e}_o$ denotes the vector

$$\underline{e}_o = (1, 1, \ldots, 1)$$

and $\beta_k \geq \|\underline{u}_k\|$ for $k = 1, 2, \ldots, (p-1)$, then we have from the dominating property of $\hat{h}_p$

$$\hat{h}_p \, (\underline{b}_1, \underline{b}_2 \ldots \underline{b}_{p-1}) \geq \|\underline{h}_p \, (\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_{p-1})\| \quad p>r \tag{3.6}$$

We will now show that the vector $\underline{v}(n) = v(n) \, \underline{e}_o$ where $v(n)$ is defined by the equation

$$v(n) - c \, \hat{g} \, (v \, \underline{e}_o) = \sum_{p=1}^{r} \|\underline{u}_p\| \, a_p^n \tag{3.7}$$

dominates $\underline{y}(n)$ in the sense that the norms of the coefficients of $\underline{v}(n)$ are not less than the norms of the corresponding coefficients of $\underline{y}(n)$; i.e.

$$\|\underline{b}_p\| \geq \|\underline{u}_p\| \qquad p = 1, 2, \ldots \tag{3.8}$$

To see this, let (3.4) with $\underline{b}_p = \beta_p \underline{e}_o$ be substituted into Eq. (3.7).

$$\sum_{p=1}^{\infty} \beta_p a_p^n - \epsilon \sum_{p=r+1}^{\infty} \hat{h}_p (\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_{p-1}) a_p^n = \sum_{p=1}^{r} \|\underline{u}_p\| a_p^n$$

As before, one obtains recursive relations for the β_p:

$$\beta_p = \|\underline{u}_p\| \qquad p = 1, 2, \ldots, r \tag{3.9}$$

and

$$\beta_p = c \hat{h}_p (\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_{p-1}) \qquad p > r \tag{3.10}$$

Thus we have $\|\underline{u}_p\| = \|\underline{b}_p\|$ $p = 1, 2, \ldots, r$ and from Eqs. (3.1), (3.6), and (3.10),

$$\begin{aligned} \|\underline{u}_p\| &\leq c \|\underline{h}_p (\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_{p-1})\| \\ &\leq c \hat{h}_p (\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_{p-1}) \\ &= \beta_p = \|\underline{b}_p\| \qquad p > r \end{aligned}$$

providing that $\|\underline{b}_k\| \geq \|\underline{u}_k\|$ for $k = 1, 2, \ldots, p-1$.

Thus Eq. (3.8) follows by induction. To prove the convergence of the dominating series (3.4), consider (3.7) as an equation for v in terms of

$$\xi = \sum_{p=1}^{r} \|\underline{u}_p\| a_p^n$$

As in Wasow's paper, it is immediately seen that there is a positive number γ_o such that (3.7) defines v as an analytic function of ξ in a certain circle $|\xi| \leq \gamma_o$. I. e., for $\left| \sum_{p=1}^{r} \|\underline{u}_p\| a_p^n \right| \leq \gamma_o$ now $|a_p| < 1$ for $p = 1, 2, \ldots,$ and therefore this inequality will be satisfied if

$$\sum_{p=1}^{r} \left\| \underline{u}_p \right\| \leqslant \gamma_o \tag{3.11}$$

since $\left| \sum_{p=1}^{r} \left\| \underline{u}_p \right\| \; a_p^{\,n} \right| < \sum_{p=1}^{r} \left\| \underline{u}_p \right\| \leqslant \gamma_o$, for $n \geqslant 0$.

Thus v is representable as an absolutely convergent series for $n \geqslant o$:

$$v(n) = \sum c_{\alpha_1 \alpha_2 \cdots \alpha_r} \left(a_1^{\alpha_1} a_2^{\alpha_2} \cdots a_r^{\alpha_r}\right)^n \tag{3.12}$$

where $\sum_{i=1}^{r} \alpha_i \geqslant 1$

The terms of (3.12) may be assumed to be arranged according to the magnitudes of the a_p's as before.

The coefficients $c_{\alpha_1 \alpha_2 \cdots \alpha_r}$ could be calculated recursively by insertion of (3.12) into (3.7) and identifying the corresponding terms right and left. However, these recursion formulas become identical with Eqs. (3.9), (3.10) if $c_{\alpha_1 \alpha_2 \cdots \alpha_r}$ is replaced by β_p. Hence the expansion (3.4) is obtained.

This proves the absolute convergence of series (3.4) for $n \geqslant 0$ and therefore of the dominated series (2.4) whenever condition (3.11) is satisfied.

It follows from Eq. (3.7) that there exists a positive number γ_1 such that $\left\| \underline{v} \right\| \leqslant \rho$ for $n \geqslant 0$ if

$$\sum_{p=1}^{r} \left\| \underline{u}_p \right\| \leqslant \gamma_1 \tag{3.13}$$

Since $\left\| \underline{y}(n) \right\| \leqslant \left\| \underline{v}(n) \right\|$ it follows that the constructed series (2.4) satisfies Eq. (2.2) whenever conditions (3.11) and (3.13) are satisfied.

Thus the following theorem has been proved:

THEOREM: A difference equation

$$\underline{x}(n+1) = \underline{F}\,[\underline{x}(n)] \qquad n = 0, 1, \ldots$$

satisfying Assumptions 1, 2, and 3 possesses a solution admitting a series expansion of the form

$$\underline{x}(n) = \underline{\xi} + \sum_{p=1}^{\infty} \underline{u}_p \, a_p^n$$

which converges absolutely for $n \geqslant 0$ provided that

$$\sum_{p=1}^{r} \| \underline{u}_p \| \leqslant \gamma \tag{3.14}$$

where γ is a positive constant that depends only on $\underline{F}$.

Since $\underline{u}_p = \sigma_p \underline{e}_p$ $(p = 1, 2, \ldots, r)$ and the σ_p's are determined by Eq. (2.7), we see that Eq. (3.14) immediately places a bound on $\underline{y}(0)$.

Hence $\underline{x}(0)$ must be sufficiently close to $\underline{\xi}$.

CONCLUSION

Although we have considered an autonomous difference equation (1. 2), this method is immediately applicable to equations with constant forcing functions, since the constant can be absorbed into the nonlinear function $\underline{F}$.

For more general forcing functions, we will obtain in Eq. (1. 3), in place of the equilibrium state $\underline{\xi}$, a particular solution $\underline{\xi}(n)$ of the equation which will be of the same form as the forcing function. This particular solution will not affect the convergence of the transient solution, except that we will now require the initial vector $\underline{x}(0)$ to be sufficiently close to $\underline{\xi}(0)$.

Clearly, we may consider the difference equation as a recurrence relation for $\underline{x}(n)$, and given $\underline{x}(0)$, we can calculate $\underline{x}(n)$ recursively for $n > 0$. However, the advantage of the method described by this paper lies in the fact that it gives the solution of the equation in closed form, and by inspection of the eigenvalues a_p of the matrix $\underline{A}$, we obtain information on the various modes of the solution.

In the case of poor convergence of the series solution (1. 3), we can always calculate $\underline{x}(1)$, $\underline{x}(2)$, etc., and using one of these as the initial vector, we may improve the convergence of the solution. This method has been used and illustrated by Pai.[1]

APPLICATION

Since nonlinear sampled-data feedback control systems are described by nonlinear difference equations, the above procedure leads to a method of analysis of such systems.

For a first-order system with constant input and which satisfies all the assumptions of Section II, we will obtain for the output a convergent series of the form

$$c(nT) = A_o + \sum_{p=1}^{\infty} A_p a_p^n$$

Taking the z-transform of this equation we obtain

$$C^*(z) = \sum_{\alpha=0}^{\infty} \frac{A_\alpha}{1-a_\alpha z^{-1}} \qquad \text{where } a_o = 1$$

This is exactly the form of $C^*(z)$ assumed by Jury and Pai.[1, 2] The poles a_α of $C^*(z)$ are all known, and the coefficients A_α can be calculated recursively using the original difference equation and the convolution z-transform.

REFERENCES

1. M. A. Pai, "The Analysis of Nonlinear Feedback Sampled-Data Systems," Ph. D. Dissertation, Dept. of Elec. Engrg., University of California, Berkeley, Calif.; Sept. 1961.

2. E. I. Jury and M. A. Pai, "Convolution Z-Transform Method Applied to Certain Nonlinear Discrete Systems," IRE Trans. on Automatic Control, Vol. AC-7, No. 1.; Jan. 1962.

3. E. Weber, "Complex Convolution Applied to Nonlinear Problems," Proc. Symp. on Nonlinear Circuit Analysis, Polytech Inst. of Brooklyn, N. Y., Vol. 6 (1959).

4. W. Wasow, "Solution of Certain Nonlinear Differential Equations by Series of Exponential Functions," Illinois J. Math., Vol. 2 (1958), pp. 254-260.

5. M. Golomb, "Solution of Certain Nonautonomous Differential Systems by Series of Exponential Functions," Illinois J. Math., Vol. 3 (1959), pp. 45-65.