[19] B. D. O. Anderson, "Structure of a class of unimodular matrices," in *Linear Algebra and Its Applications*. to be published.
[20] ——, "The inverse problem of stationary covariance generation," *J. of Statistical Physics*, vol. 1, no. 1, pp. 133–147, 1969.
[21] J. C. Willems, "Time reversibility in deterministic and stochastic dynamical systems," *Proc. 1977 U.S.-Italy seminar on Variable Structure Systems*, in *Springer Lecture Notes in Economics and Mathematical Systems*, to be published.

Brian D. O. Anderson (S'62–M'66–SM'74–F'75), for a photograph and biography please see page 873 of the October 1979 issue of this TRANSACTIONS.

Thomas Kailath (S'57–M'62–F'70), for a photograph and biography please see page 873 of the October 1979 issue of this TRANSACTIONS.

Letters to the Editor

Practical System Inversion Without Impulse Response Approximation

BRIAN D. O. ANDERSON

Abstract—**A procedure is described for finding the input sequence of a linear system given the output sequence. The system need not be minimum phase though it cannot have unit circle zeros. The method does not involve impulse response approximations, but does rely in the general case on the availability of a storage medium such as tape. Errors are introduced by initial or terminal condition effects but these are guaranteed to have decaying influence as one moves away from the initial or terminal time.**

I. INTRODUCTION

Let S be a linear time-invariant finite-dimensional discrete time system with transfer function matrix $W(z)$. Suppose further that S is bounded-input bounded-output stable. We consider the problem of determining from output sequence measurements $\{y(n)\}$ on S, assumed to be bounded, what the input sequence $\{u(n)\}$ was which gave rise to this output sequence. To this end, it is necessary to suppose that the normal rank of $W(z)$ equals the number of columns of $W(z)$.

This problem has been well studied, using frequency domain and state-variable concepts, e.g., [1]–[3]. Particular attention has been paid to the case when $W(z)$ has more rows than columns, and thus the inverse is not uniquely defined. In this note, our aim is to focus on the problem of obtaining an inverse implemented using numerically stable recursions. Since the poles of any inverse must include the (multivariable) zeros of the original transfer function matrix [4], the core of the problem is what to do about zeros outside the region $|z|<1$.

There is one approximate approach that has found extensive technological application, in telephone channel equalizers constructed with tapped delay lines [5]. Here, the inverse system is being approximated with a finite impulse response filter, and achieves approximate causal inversion, normally with delay.

Manuscript received September 19, 1978; revised December 28, 1978. This paper was supported by the Australian Research Grants Committee.

The author is with the Department of Electrical Engineering, University of Newcastle, New South Wales, 2308, Australia.

Consider the scalar case. If $W(z)=p(z)/q(z)$ form $W^{-1}(z)=q(z)/p(z)$, and then regarding $W^{-1}(z)$ as a 2-sided z-transform form the inverse z-transform $\{w(n)\}$ with the property that $w(\pm n)\to 0$ as $n\to\infty$, assuming this possible. Then one truncates this for some negative n, say $-n=M$, and defines $\bar{w}(n)=w(n+M)$, for $n\leqslant 0$, $\bar{w}(n)=0$, $n<0$. Thus $\bar{w}(n)$ is a causal impulse response. One then has the task of obtaining a finite-dimensional realization of $\bar{w}(n)$. This can be done by truncating $\bar{w}(n)$ for large positive n, yielding a finite-impulse-response filter. When this filter is driven causally by $\{y(n)\}$, the output at time n away from the initial time is approximately $u(n-M)$.

The truncation of $w(n)$ for large negative n and perhaps also for large positive n means that errors will be present throughout the interval over which inversion takes place, to an amount which depends on the truncation point. There will be further error in the neighborhood of the initial time due to initial condition effects.

In contrast, the method following introduces errors, in general, in the neighborhood of initial and terminal times which decay to zero in moving away from these times, as explained below. The penalty is that bulk storage is normally required, though of a convenient form.

II. SIMPLE CASES

Consider the system

$$y(n)=u(n)+\alpha u(n-1). \tag{1}$$

If $|\alpha|<1$, we can recover $\{u(n)\}$ from $\{y(n)\}$ by solving in the half-line $n\geqslant s_i$ the equation

$$\bar{u}(n)+\alpha\bar{u}(n-1)=y(n) \tag{2}$$

with initial condition $\bar{u}(s_i-1)=0$. This recursively yields $\bar{u}(s_i)$, $\bar{u}(s_i+1),\cdots$, with $|u(n)-\bar{u}(n)|\leqslant K|\alpha|^{(n-s_i)}$, for some K. On the other hand, if $|\alpha|>1$, we could solve the difference equation (2) with a *terminal* condition $\bar{u}(s_t+1)=0$ in the half-line $n\leqslant s_t$, recursively obtaining $\bar{u}(s_t), \bar{u}(s_t-1),\cdots$. Since (1) and (2) imply that

$$\bar{u}(n-1)-u(n-1)=-\alpha^{-1}[\bar{u}(n)-u(n)]$$

it is clear that $|u(n)-\bar{u}(n)|\leqslant K|\alpha|^{n-s_t}$.

0098-4094/79/1100-0965$00.75 ©1979 IEEE

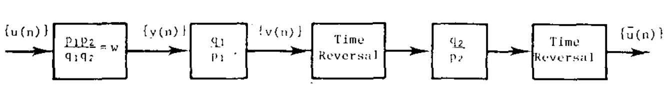

Fig. 1. Inversion procedure.

On the other hand, neither of these approaches nor indeed any other approach will be successful if $|\alpha|=1$. Any initial error would propagate in both $n \geqslant s_i$ and $n \leqslant s_t$ without attenuation.

These three simple cases exemplify the fact that the general problem will only be solvable if

$$W(z) \text{ has full rank on } |z|=1 \tag{3}$$

where by solvability, we have in mind the possibility of endpoint errors which propagate with reducing effect as we move from the end point. This requirement is consistent with the observation noted above [4] that the poles of any inverse of $W(z)$ include all multivariable zeros of $W(z)$; of course, with all poles of $W(z)$ lying in $|z|<1$, any rank drop of $W(z)$ on $|z|=1$ corresponds to a zero. For the approximate approach of the introduction to work (3) is also a requirement, since without it, no inverse z-transform will decay to zero as $|n|\to\infty$.

III. Scalar Systems

Suppose $W(z)=p(z)/q(z)$ where p, q are prime polynomials, q has all roots inside $|z|<1$, p has no roots on $|z|=1$, and deg $q \geqslant$ deg$p$. Write

$$\frac{p(z)}{q(z)}=\frac{p_1(z)p_2(z)}{q_1(z)q_2(z)} \tag{4}$$

where all zeros of p_1 lie in $|z|<1$, all zeros of p_2 lie in $|z|>1$ and $\deg q_1 \leqslant \deg p_1$. (This factorization may cause numerical problems, and this difficulty is not addressed in this paper.) Then as illustrated in Fig. 1, the following steps will recover $\{u(n)\}$ from $\{y(n)\}$, modulo endpoint errors.

1) Running forwards in time, obtain a signal $\{v(n)\}$ from $\{y(n)\}$ by passing $\{y(n)\}$ into the stable causal system with transfer function q_1/p_1. (This involves selection of an initial time s_i). Modulo decaying initial condition effects $\{v(n)\}$ is also the signal which would result from passing $\{u(n)\}$ into a filter of transfer function p_2/q_2. If this signal is $\bar{v}(n)$, then $|v(n)-\bar{v}(n)| \leqslant K_1\alpha_1^{n-s_i}$, for some constants $K_1>0$ and $\alpha_1 \in [0,1)$, with α_1 approximately the largest magnitude of a zero of p_1.

2) Record $\{v(n)\}$ on tape. By reversing the tape, play $\{v(n)\}$ backwards into a system with transfer function $q_2(z)/p_2(z)$. More precisely, with $p_2(z)=a_0z^{\alpha}+a_1z^{\alpha-1}+\cdots+a_{\alpha}$ and $q_2(z)=b_0z^{\beta}+b_1z^{\beta-1}+\cdots+b_{\beta}$, one solves

$$a_0\bar{u}(n+\alpha)+a_1\bar{u}(n+\alpha-1)+\cdots+a_{\alpha}\bar{u}(n)$$
$$=b_0v(n+\beta)+b_1v(n+\beta-1)+\cdots+b_{\beta}v(n) \tag{5}$$

backwards in time, using zero terminal conditions, i.e., one sets $\bar{u}(s_t+\alpha)=\bar{u}(s_t+\alpha-1)=\cdots=\bar{u}(s_t+1)=0$ and uses (5) to find $\bar{u}(s_t)$, then $\bar{u}(s_t-1),\cdots$. (Notice that $a_{\alpha}\neq 0$ because $p_2(z)$ has no zeros in $|z|<1$.) This calculation is a stable one and forgets the terminal conditions in view of the zero positions of $p_2(z)$.

It follows that for some constants $K>0$ and $\alpha_1>1$ with α_2 approximately the smallest magnitude of a zero of p_2

$$|u(n)-\bar{u}(n)| \leqslant K(\alpha_1^{n-s_i}+\alpha_2^{n-s_t}). \tag{6}$$

3) Record the values of $\bar{u}(n)$ on tape and reverse the tape. Of course, in certain cases not all the above steps will be necessary.

The trick of reversing time when all modes are unstable is known in numerical analysis as backward recurrence, and is discussed in, e.g., [6].

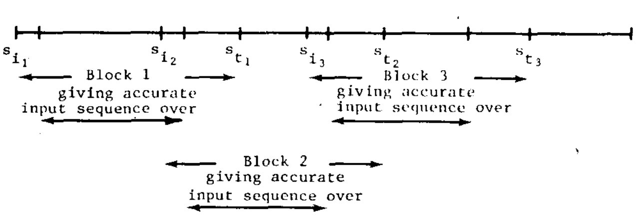

Fig. 2. Division of long interval into blocks.

IV. Matrix Systems

Suppose $W(z)$ is a $p\times q$ matrix with rank q everywhere on $|z|=1$ and every entry analytic in $|z|<1$. Construct the Smith–McMillan form of $W(z)$, see [7], as

$$W(z)=U_1(z)\begin{bmatrix}\Delta(z)\\ \hdashline 0\end{bmatrix}U_2(z) \tag{7}$$

where $U_1(z)$ and $U_2(z)$ are polynomial in z with nonzero constant determinant, $\Delta(z)$ is $q\times q$ diagonal matrix of nonzero rational functions, not necessarily proper, and the zero block is $(p-q)\times q$.

Next write

$$\Delta(z)=\Delta_1(z)\Delta_2(z)$$

where $\Delta_i(z)$ are diagonal matrices of rational functions, and after all cancellations are done, numerators of $\Delta_1(z)$ have all zeros in $|z|<1$ and numerators of $\Delta_2(z)$ have all zeros in $|z|>1$.

A left inverse for $W(z)$ is provided by

$$W_l^{-1}(z)=\{U_2^{-1}(z)\Delta_2^{-1}(z)\}\{[\Delta_1^{-1}(z)\quad 0]U_1^{-1}(z)\}$$
$$=V_2(z)V_1(z)$$

with obvious definitions of $V_i(z)$. Select s such that $\bar{V}_1(z)=V_1(z)z^{-s}$ is proper, and set $\bar{V}_2(z)=V_2(z)z^{s}$. Then

$$W_l^{-1}(z)=\bar{V}_2(z)\bar{V}_1(z) \tag{8}$$

and $\bar{V}_2(0)$ is finite. The computation of $\bar{V}_1$ and $\bar{V}_2$ may be difficult numerically.

The steps for constructing an approximation $\{\bar{u}(n)\}$ to $\{u(n)\}$ are now as follows.

1) Running forwards in time from some initial time s_i, obtain $\{v(n)\}$ by passing $\{y(n)\}$ into the stable causal system with transfer function matrix $\bar{V}_1(z)$.

2) Record $\{v(n)\}$ on tape. By reversing the tape play $\{v(n)\}$ backwards into a system with transfer function $\bar{V}_2(z)$, starting at terminal time s_t. The output is $\bar{u}(n)$.

3) Record $\bar{u}(n)$ on tape and reverse the tape. Away from the endpoints, $u(n)=\bar{u}(n)$, while (6) holds on an entry-by-entry basis for all n.

In step 2) the fact that $\bar{V}_2(0)$ is finite is helpful. Let $\bar{V}_2(z)$ have a coprime left-matrix fraction description $A_2^{-1}(z)B_2(z)$, see [1]. Then $A_2(0)$ is nonsingular. The matrix version of (5) then has $A_{\alpha}=A_2(0)$ nonsingular, making the recursion straightforward.

V. Conclusions

In this note, we have suggested a procedure for system inversion which relies on the availability of a significant amount of storage in what amounts to a standard stack format. Inversion is asymptotically exact, in the sense that at times remote from the endpoints, the error goes to zero, in contrast to another popular method for inversion.

It may be that the storage is not available, or the delay in processing is not acceptable. In this case, the data should be divided into overlapping blocks, the size of the overlap corresponding to the interval over which initial condition effects are important, see Fig. 2. Blocks would then be processed one at a time. As depicted, each block calculation could be done independently of others. However, one could use the block 1 calculations to eliminate the initial interval-of-inaccuracy of block 2, making block 2 (and similarly all later blocks) shorter.

References

[1] W. A. Wolovich, *Linear Multivariable Systems*. New York: Springer Verlag, 1974.
[2] P. J. Moylan, "Stable Inversion of Linear Systems," *IEEE Trans. Auto. Contr.*, vol. AC-22, pp. 74–78, Feb. 1977.
[3] L. M. Silverman, "Inversion of multivariable linear systems," *IEEE Trans. Auto. Contr.*, vol. AC-14, pp. 272–274, June 1969.
[4] G. Bengtsson, "Minimal system inverses for linear multivariable systems," *J. Math. Anal. Appl.*, vol. 46, pp. 261–274, 1974.
[5] R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication*. New York: McGraw-Hill, 1968.
[6] L. Fox and D. F. Mayers, *Computing Methods for Scientists and Engineers*. Oxford, England: Clarendon, 1968.
[7] R. W. Brockett, *Finite Dimensional Linear Systems*. New York: Wiley, 1970.

A Simultaneous Analog-to-Quaternary Converter

K. WAYNE CURRENT

***Abstract*—The function of an analog-to-quaternary (A/Q) converter is to convert an analog input voltage into a quaternary output word. In this correspondence, a simple simultaneous analog-to-quaternary converter which makes use of new current-mode quaternary threshold logic circuits is described.**

I. Introduction

The potential packing density advantages of integrated multiple valued logic circuitry are gradually being acknowledged by increasing numbers of designers (e.g., [1]–[4]). Rapid increases in packing densities are now being obtained through an intensive industrial effort to minimize integrated devices and metal lines. But even greater packing densities could be obtained if these silicon technology advances were coupled with the packing density advantages of multiple valued logic. In some applications, nearly 50 percent fewer integrated devices are required to implement certain digital signal processing functions [5]. In most applications of multiple valued logic, binary inputs are converted to multiple valued logical signals for internal signal processing and then decoded into binary equivalents for delivery off-chip. In this correspondence, we consider the conversion of analog signals to four-valued or quaternary logical signals. A simple simultaneous analog-to-quaternary (A/Q) converter design which uses new current-mode quaternary threshold logic circuits is presented.

Manuscript received November 6, 1978; revised January 30, 1979. This work was supported in part by the Intercampus Faculty Grant 3-5225061900-5.

The author is with the Department of Electrical Engineering, University of California, Davis, CA 95616.

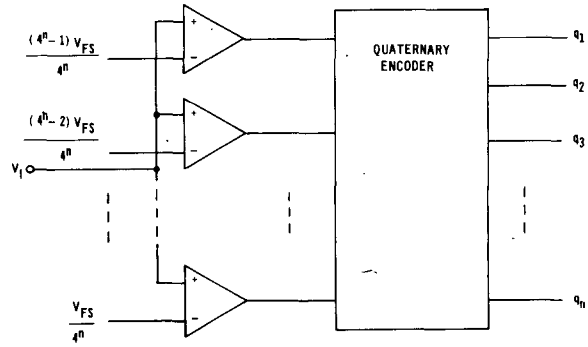

Fig. 1. Simultaneous A/Q conversion.

II. Simultaneous Analog-To-Quaternary Conversion

Simultaneous A/Q conversion is the fastest and most simple technique. The analog input voltage is approximated as a quaternary fraction of a reference full-scale voltage. The output of the A/Q converter can be defined as

$$Q=\frac{V_I}{V_{FS}}=q_1 4^{-1}+q_2 4^{-2}+\cdots+q_n 4^{-n}$$

where V_I and V_{FS} are the input and full-scale voltages, respectively, q_i are the quaternary digits having values zero, one, two, or three, and n is the number of digits in the output word. The n digits of the A/Q converter output word are presented simultaneously in the form of n parallel outputs. This type of converter uses (4^n-1) separate analog comparators with separate fixed reference voltages from zero to full scale for all the quantization levels in the n quaternary digit output word. A block diagram of this standard technique is shown in Fig. 1. The outputs of these comparators are then processed with an encoding logic circuit to develop the quaternary-coded output word. Simultaneous A/Q conversion is the fastest technique since all encoder circuit inputs are processed simultaneously and no feedback or iterative procedure is required; thus the encoding operation is completed in a single clock cycle.

One drawback to this technique is the rapid growth of circuit complexity for converters with n larger than about 2. For $n=2$, 15 comparators and their 15 reference voltages plus the encoding logic are required. With $n=3$, 63 comparators and their 63 separate reference voltages, plus encoding logic, are required, thus significantly increasing the circuitry required for implementation. Simultaneous converters with only two-digit output words can be advantageously employed as sectional building blocks in forming other types of A/Q converters with reduced aperture times.

III. Current-Mode Threshold Logic A/Q Converter

A two-digit simultaneous A/Q converter block diagram organized for implementation with current-mode threshold logic is shown in Fig. 2. The usual "thermometer" type arrangement of voltage references and current switching comparators is used to generate the encoding circuit inputs. The encoding circuit then generates the two-digit quaternary output word by counting the number of comparator outputs activated.

The encoding circuit block diagram is shown in more detail in Fig. 3. The overall circuit operation is described as follows. Each