NASA TECHNICAL NOTE

NASA TN D-3124

c. 1

LOAN COPY: RETURN
AFWL (WLIL-2)
KIRTLAND AFB, N MEX

01301 52

TECH LIBRARY KAFB, NM

NASA TN D-3124

# ERROR ANALYSIS OF BINARY RATE MULTIPLIER

*by George J. Moshos*

*Lewis Research Center*

*Cleveland, Ohio*

Technical Library
Air Weapons
Laboratory
20 DEC 1965
RECEIVED

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • DECEMBER 1965

TECH LIBRARY KAFB, NM
0130152

NASA TN D-3124

# ERROR ANALYSIS OF BINARY RATE MULTIPLIER

By George J. Moshos

Lewis Research Center
Cleveland, Ohio

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – Price $2.00

# ERROR ANALYSIS OF BINARY RATE MULTIPLIER

by George J. Moshos

Lewis Research Center

## SUMMARY

The binary rate multiplier is studied as a means of achieving approximate multiplication. The difference between the actual and desired output is defined as the error of the binary rate multiplier, and closed formulas are obtained for expressing this error in explicit form depending on the starting conditions of the binary rate multiplier counter. As a result of analyzing these error formulas, error bounds are obtained.

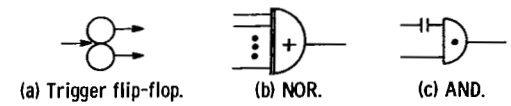
## INTRODUCTION

An integral part of many special purpose digital computers used for real time control is the binary rate multiplier (BRM) (e.g., refs. 1 to 4). In these applications this unit is used as a means of scaling down a pulse stream to some specified fraction. A logic diagram of a BRM, which is built out of the standard logic elements shown in figure 1, is shown in figure 2(a). The NOR element shown in figure 1(a) may have various number of inputs. The parallel lines shown on one of the inputs of figure 1(c) are included to indicate that the AND circuit is intended to act as a pulse gate dependent on the level setting of the other line. The following brief description explains the operations on the BRM.

The input pulse stream is applied directly to the binary counter whose value is denoted by $x_n x_{n-1} \cdots x_2 x_1$. Each flip-flop of the counter is operated as a trigger. For every two input pulses to a trigger, two output pulses are produced; one pulse when the flip-flop makes a 0 to 1 transition, called an $\alpha$ pulse, and one when the flip-flop makes a 1 to 0 transition, called a $\beta$ pulse. The $\beta$ pulse is used to trigger the next stage of the counter. The $\alpha$ pulses are gated through AND gates and mixed through a NOR element to produce the desired fraction of the input pulses. This simple mixing technique may be used because the $\alpha$ pulses from the various stages are separated in time from each other. This timing factor is shown in figure 2(b).

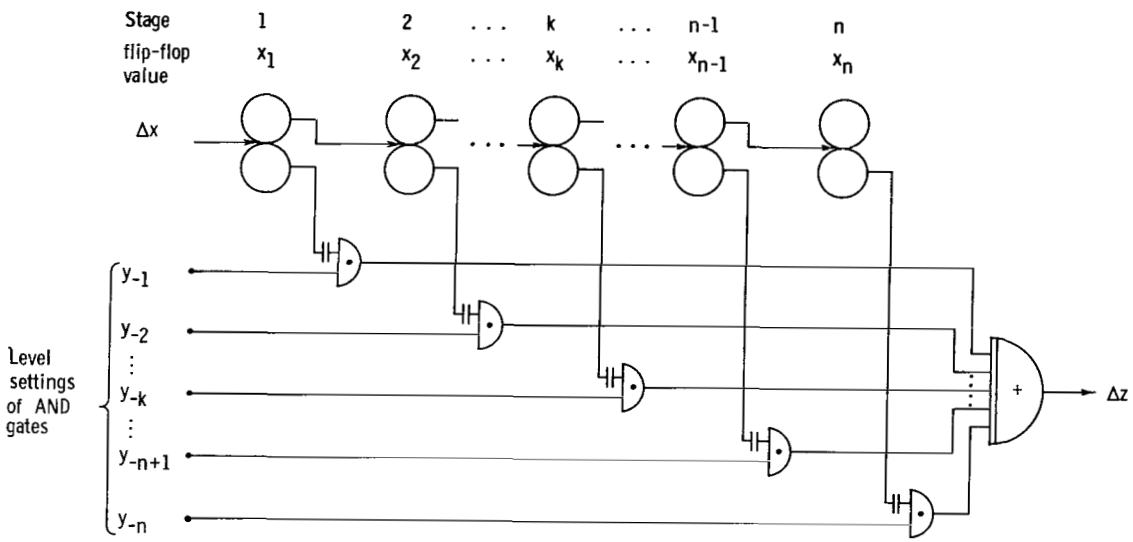
This unit may be used to achieve approximate multiplication. In particular, if $\Delta x$ is the number of input pulses and $y$ is a binary number less than 1, $\Delta z$, the number of output pulses, may be stated quantitatively as

$$\Delta z = y \, \Delta x \tag{1}$$

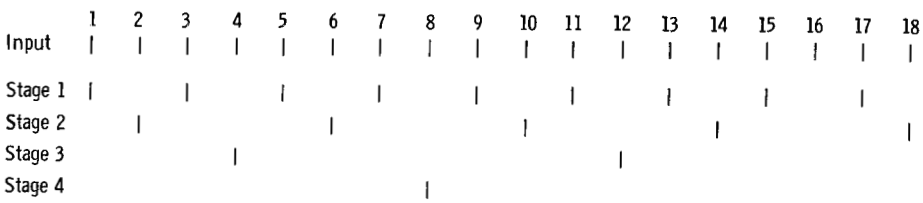


(a) Trigger flip-flop. (b) NOR. (c) AND.

Figure 1. - Logic elements.




(a) Logic diagram.




(b) Timing diagram.

Figure 2. - Binary rate multiplier.

The difference between the actual output of the BRM and $\Delta z$, as given by equation (1), is defined as the error of the unit. In this report, the formulation of error is explicitly given and studied. The analysis of the BRM is begun by deriving the approximate relation shown in equation (1).

## BINARY RATE MULTIPLIER

The quantitative relation of a BRM may be expressed as follows: If $\Delta x$ is the number of input pulses, the number of output pulses produced by the $k^{th}$ stage of the counter is $\Delta x \cdot 2^{-k}$. This multiplicative relation will remain valid over any interval for which $\Delta x$ is a multiple of $2^k$ pulses. If $y_{-k}$ is the level setting of the $k^{th}$ stage AND gate, the number of output pulses that

may be gated through this stage will be $y_{-k}\,\Delta x \cdot 2^{-k}$. Since the output pulses from the various stages are simply mixed, the number of output pulses $\Delta z$ of an n-stage BRM over any interval $\Delta x$, which is a multiple of $2^n$ pulses, will be the sum of all the pulses gated through all the stages. This output is

$$\Delta z = \Delta x \sum_{i=1}^{n} y_{-i} 2^{-i} \tag{2}$$

The quantity $y = \sum_{i=1}^{n} y_{-i} 2^{-i}$ is a binary number. Therefore, equation (2) may be written as

$$\Delta z = y\,\Delta x$$

where the range of $y$ is

$$0 \leq y \leq 1 - 2^{-n} \tag{3}$$

in steps of $2^{-n}$.

If $y$ is constant over a $\Delta x$ interval of $2^n$ pulses, the output given by equation (1) is exact. If $y$ is constant over a $\Delta x$ interval of less than $2^n$ pulses, however, this multiplicative relation may not be valid. In this latter case, the actual output depends not only on the values of $y$ and $\Delta x$ but also on the starting value of the BRM counter. If the output from a machine, whose BRM counter starting value is $x_s$, is denoted by $\Delta z_{x_s}$, it can be shown that the average output over all of the $2^n$ possible machines is also given by equation (1). This can be demonstrated as follows: If $\overline{\Delta z}$ is the average output over all these machines, $\overline{\Delta z}$ is, by definition,

$$\overline{\Delta z} = \frac{\sum_{x_s=0}^{2^{n-1}} \Delta z_{x_s}}{2^n} \tag{4}$$

The sum $\sum_{x_s=0}^{2^{n-1}} \Delta z_{x_s}$ given by equation (4) is the total pulse output over the $2^n$ different possible machines when each machine receives $\Delta x$ input pulses and its AND gates are set to the value $y$. It will be observed that the $\alpha$ transitions produced by these $2^n$ machines are the same as the ones produced by a single machine with $2^n\,\Delta x$ successive input pulses. For example, the $\alpha$ transition ending with counter value $x$ is attained $\Delta x$ times (once by each of the $\Delta x$ machines whose starting value is prior to $x$ in the counting sequence) of the $2^n$ machines used in the average and also $\Delta x$ times when $2^n\,\Delta x$ successive input pulses are applied to a single machine (since each counter value is traversed $\Delta x$ times in this case). Therefore, the total pulse output over all the $2^n$ possible machines with each receiving $\Delta x$ input pulses is equal to the pulse output of a single machine receiving $2^n\,\Delta x$ input pulses. In this latter case the pulse output is also given by equation (1) since the input pulse interval is a multiple of $2^n$ pulses. Therefore,

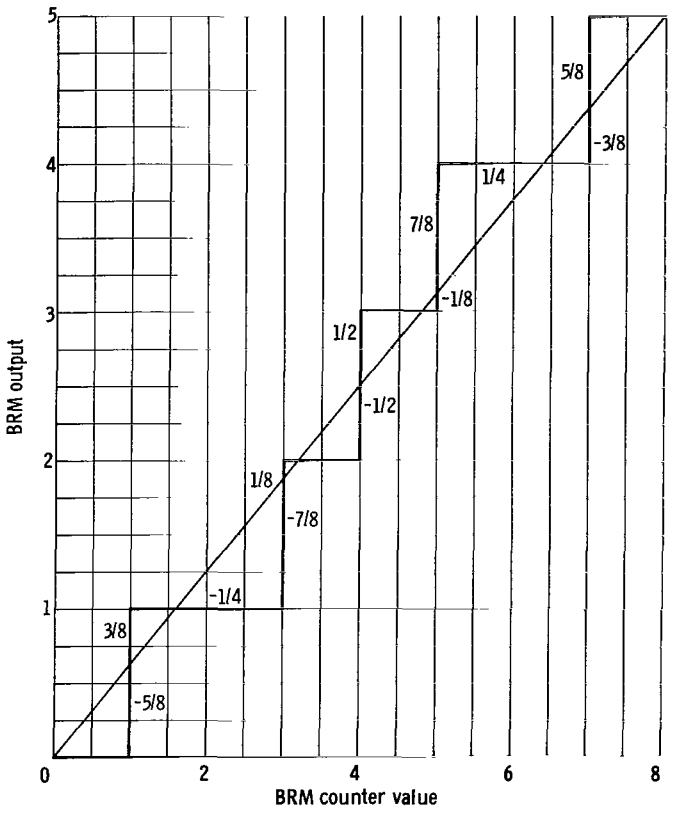



Figure 3. - Multiplication error resulting from y = .101.

$$\sum_{x_s=0}^{2^{n-1}} \Delta z_{x_s} = y \cdot 2^n \cdot \Delta x \qquad (5)$$

Combining equations (4) and (5) yields

$$\overline{\Delta z} = y\,\Delta x \qquad (6)$$

## MULTIPLICATION ERROR FORMULAS

The error for the BRM may be defined as the actual output minus $y\,\Delta x$ (i.e., the value predicted by eq. (1)). Since the actual output changes only when the input pulses arrive, it is only necessary to consider the error at these discrete times. For example, figure 3 gives the output for a three-stage BRM together with the values predicted by equation (1) when the BRM counter starting value is zero and $y = .101$. The difference between these two curves is the defined error. Yet, in discussing error, the error value will be given only when the abscissa is discrete values such as 0, 1, 2, etc. (which correspond to the BRM counter values). Moreover, it will be observed that when an output pulse is produced, the error may change by one quanta. Therefore, it is necessary to distinguish between the error immediately before the output pulse and the error immediately after the output pulse. When the BRM counter starting value is zero, these two values of the error will be denoted by F and E, respectively. When the BRM counter starting value is arbitrary, the starting value must also enter into the error formula as a parameter. The errors will be denoted prior to and after the arrival of the output pulse as H and G, respectively. It will be noted that the error defined this way makes E a special case of G, and F a special case of H. Nevertheless, this distinction is maintained, since it is convenient for our subsequent discussion.

Starting the BRM counter with zero gives the definition of the error E as the actual output after the output pulses are generated, minus $y\,\Delta x$. This difference, when only one stage of an n-stage BRM is gated by y, can be expressed systematically in tabular form as shown in table I.

TABLE I. - MULTIPLICATION ERROR E OF BRM WHEN ONE STAGE IS GATED

| $y_{-1} = 1$ | | $y_{-2} = 1$ | | $y_{-3} = 1$ | | $y_{-k} = 1$ | |
|---|---|---|---|---|---|---|---|
| $x_1$ | E | $x_2x_1$ | E | $x_3x_2x_1$ | E | $x_kx_{k-1} \ldots x_3x_2x_1$ | E |
| 0 | 0 | 0 0 | 0 | 0 0 0 | 0 | 0 0 . . . 0 0 0 | 0 |
| 1 | 1/2 | 0 1 | -1/4 | 0 0 1 | -1/8 | 0 0 . . . 0 0 1 | $-1/2^k$ |
| | | 1 0 | 2/4 | 0 1 0 | -2/8 | 0 0 . . . 0 1 0 | $-2/2^k$ |
| | | 1 1 | 1/4 | 0 1 1 | -3/8 | . | . |
| | | | | 1 0 0 | 4/8 | . | . |
| | | | | 1 0 1 | 3/8 | . | . |
| | | | | 1 1 0 | 2/8 | 0 1 . . . 1 1 1 | $-(2^{k-1} - 1)/2^k$ |
| | | | | 1 1 1 | 1/8 | 1 0 . . . 0 0 0 | 1/2 |
| | | | | | | 1 0 . . . 0 0 1 | $(2^{k-1} - 1)/2^k$ |
| | | | | | | . | . |
| | | | | | | . | . |
| | | | | | | . | . |
| | | | | | | 1 1 . . . 1 1 1 | $1/2^k$ |

An inspection of this table shows that the error associated with the various stages of a BRM may be expressed more concisely in algebraic form as shown in table II.

TABLE II. - MULTIPLICATION ERROR E

IN ALGEBRAIC FORM

| Stage | E |
|---|---|
| 1 | $y_{-1}(x_1/2)$ |
| 2 | $y_{-2}(x_2/2 - x_1/4)$ |
| 3 | $y_{-3}(x_3/2 - x_2/4 - x_1/8)$ |
| . | . |
| . | . |
| . | . |
| k | $y_{-k}(x_k/2 - x_{k-1}/4 - \ldots - x_1/2^k)$ |

For an arbitrary value of y, the value of E is the linear combination of the values shown in table II. This bilinear form is shown in equation (7) for an n-stage BRM. The element subscripts of the Boolean vectors x and y (i.e., vectors whose elements are 0 or 1), which are shown in this equation, correspond to the stage numbers of the BRM shown in figure 1:

$$E = (x_1, x_2 \ldots x_n)M\begin{pmatrix} y_{-1} \\ y_{-2} \\ \cdot \\ \cdot \\ \cdot \\ y_{-n} \end{pmatrix} \tag{7}$$

$$M = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & -\frac{1}{8} & -\frac{1}{16} & \cdots & -\frac{1}{2^{n-1}} & -\frac{1}{2^n} \\ 0 & \frac{1}{2} & -\frac{1}{4} & -\frac{1}{8} & \cdots & -\frac{1}{2^{n-2}} & -\frac{1}{2^{n-1}} \\ \cdot & \cdot & \frac{1}{2} & -\frac{1}{4} & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \frac{1}{2} & \cdots & \cdot & \cdot \\ \cdot & \cdot & & & & \cdot & \cdot \\ & & & & \cdot & & \\ & & & & \cdot & & \\ & & & & \cdot & & \\ & & & & \frac{1}{2} & -\frac{1}{4} & -\frac{1}{8} \\ & & & & & \frac{1}{2} & -\frac{1}{4} \\ 0 & \cdot & \cdot & \cdot & & 0 & \frac{1}{2} \end{pmatrix} \tag{8}$$

In the formulation of E, the maximum values of the output of the BRM were reflected at the points of discontinuities. It will be observed that just prior to these points the error is one quanta less than that shown by E. A formulation of F, in which the minimum values are reflected at the points of discontinuities, can be obtained in a manner similar to that for obtaining E. The quantity F, when only one stage of an n-stage BRM is gated by y, is shown in tabular form in table III (only a few cases are exhibited).

TABLE III. - MULTIPLICATION ERROR F WHEN

ONLY ONE STAGE IS GATED

| $y_{-1} = 1$ | | | $y_{-2} = 1$ | | | $y_{-3} = 1$ | | |
|---|---|---|---|---|---|---|---|---|
| $C_1$ | $x_1$ | F | $C_2C_1$ | $x_2x_1$ | F | $C_3C_2C_1$ | $x_3x_2x_1$ | F |
| 0 | 0 | 0 | 0 0 | 0 0 | 0 | 0 0 0 | 0 0 0 | 0 |
| 1 | 1 | -1/2 | 1 1 | 0 1 | -1/4 | 1 1 1 | 0 0 1 | -1/8 |
| | | | 1 0 | 1 0 | -2/4 | 1 1 0 | 0 1 0 | -2/8 |
| | | | 0 1 | 1 1 | 1/4 | 1 0 1 | 0 1 1 | -3/8 |
| | | | | | | 1 0 0 | 1 0 0 | -4/8 |
| | | | | | | 0 1 1 | 1 0 1 | 3/8 |
| | | | | | | 0 1 0 | 1 1 0 | 2/8 |
| | | | | | | 0 0 1 | 1 1 1 | 1/8 |

Comparing table III with table I shows that the values of F equal the values of E except at the points where the discontinuity occurs. At these points, F equals -1/2, while the corresponding value of E equals +1/2. The values of C shown in table III correspond to the 2's complement of the x values. It will be observed that the values of F are identical in terms of C to the negative values of E. Consequently, it can be asserted that F in terms of C is just the negative of E:

$$F = -(C_1, C_2 \ldots C_n)M\begin{pmatrix} y_{-1} \\ y_{-2} \\ \cdot \\ \cdot \\ \cdot \\ y_{-n} \end{pmatrix} \tag{9}$$

An example will help clarify these formulas. In this example the value of y is .101, and the values of E and F are calculated for successive BRM counter values that are notated by the subscripts on E and F:

$$\begin{pmatrix} E_0 \\ E_1 \\ E_2 \\ E_3 \\ E_4 \\ E_5 \\ E_6 \\ E_7 \end{pmatrix} = \begin{pmatrix} 000 \\ 100 \\ 010 \\ 110 \\ 001 \\ 101 \\ 011 \\ 111 \end{pmatrix} \begin{pmatrix} 1/2 & -1/4 & -1/8 \\ 0 & 1/2 & -1/4 \\ 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3/8 \\ -1/4 \\ 1/8 \\ 1/2 \\ 7/8 \\ 1/4 \\ 5/8 \end{pmatrix}$$

$$\begin{pmatrix} F_0 \\ F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \\ F_7 \end{pmatrix} = - \begin{pmatrix} 000 \\ 111 \\ 011 \\ 101 \\ 001 \\ 110 \\ 010 \\ 100 \end{pmatrix} \begin{pmatrix} 1/2 & -1/4 & -1/8 \\ 0 & 1/2 & -1/4 \\ 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -5/8 \\ -1/4 \\ -7/8 \\ -1/2 \\ -1/8 \\ 1/4 \\ -3/8 \end{pmatrix}$$

This example is shown in graphical form in figure 3 (p. 4).

When a BRM counter starts out with an arbitrary value, the starting value must enter into the error formula as a parameter. These error formulas are given explicitly by equations (10) and (11). In particular, these equations reflect the maximum (denoted by G) and minimum (denoted by H) values of the actual output at the points of discontinuities. In these formulations, x and $x_S$ represent the value and the initial value of the counter, and C and $C_S$ represents the 2's complements of these values. The subscripts on those literals represent, as before, the stage of the BRM. In equation (11), $x_{SR}$ identifies the rightmost (i.e., lowest order) counter bit, whose value is 1; for example, for the counter value 100, $x_{SR}y_{-R} = y_{-3}$, for counter value 011, $x_{SR}y_{-R} = y_{-1}$, etc:

$$G = (x_1 - x_{S1}, x_2 - x_{S2} \cdots x_n - x_{Sn})M\begin{pmatrix} y_{-1} \\ y_{-2} \\ \cdot \\ \cdot \\ \cdot \\ y_{-n} \end{pmatrix} \tag{10}$$

$$H = -x_{SR}y_{-R} - (C_1 - C_{S1}, C_2 - C_{S2} \cdots C_n - C_{Sn})M\begin{pmatrix} y_{-1} \\ y_{-2} \\ \cdot \\ \cdot \\ \cdot \\ y_{-n} \end{pmatrix} \tag{11}$$

The following two examples illustrate equations (10) and (11) for $y = .01$ and $(x_{S1}, x_{S2}) = (0,1)$ and are calculated for successive BRM counter values that are notated by the subscripts on G and H:

$$\begin{pmatrix} G_0 \\ G_1 \\ G_2 \\ G_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & -1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1/2 & -1/4 \\ 0 & 1/2 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1/4 \\ -1/2 \\ -3/4 \end{pmatrix}$$

$$\begin{pmatrix} H_0 \\ H_1 \\ H_2 \\ H_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 1 & -1 \\ 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1/2 & -1/4 \\ 0 & 1/2 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1/4 \\ -1/2 \\ -3/4 \end{pmatrix}$$

In these examples it will be noted that $(x_i - x_{Si})$ may be 0, 1, or -1.

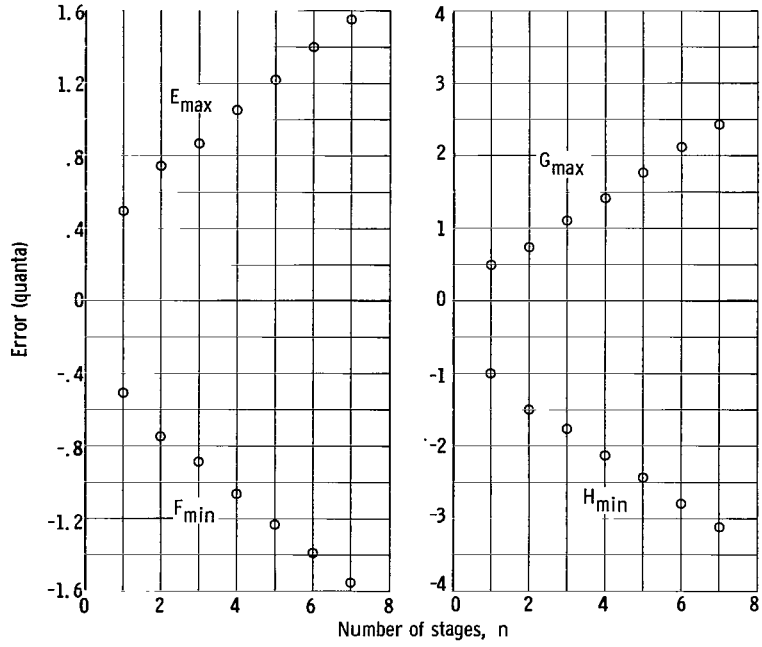
## MULTIPLICATION ERROR BOUNDS

The maximum positive error and the minimum negative error for a n stage BRM whose counter starts out with zero may be obtained by an analysis of equations (7) and (9), respectively. These values will then form a bound of the deviation of the BRM from that of exact multiplication. This analysis is presented in appendix A. It is shown in that analysis that for an n-stage BRM these values are:

$$E_{max}(n) = \frac{7}{18} + \frac{n}{6} + \frac{(-1)^n}{9 \cdot 2^n} \tag{12}$$

$$F_{min}(n) = -\frac{7}{18} - \frac{n}{6} - \frac{(-1)^n}{9 \cdot 2^n} \tag{13}$$

Equation (12) is plotted together with equation (13) in figure 4(a). As a by-product of developing equations (12) and (13), the points where these values




(a) Zero starting. (b) Arbitrary starting.

Figure 4. - Multiplication error bounds.

TABLE IV. - x AND y VALUES FOR $E_{max}$

| n | $x_7 \cdots x_1$ | $y_{-1} \cdots y_{-7}$ | $x_7 \cdots x_1$ | $y_{-1} \cdots y_{-7}$ | $E_{max}$ |
|---|---|---|---|---|---|
| 2 | 11 | 11 | 11 | 11 | 3/4 |
| 3 | 101 | 101 | 111 | 111 | 7/8 |
| 4 | 1011 | 1101 | 1101 | 1011 | 17/16 |
| 5 | 10101 | 10101 | 11011 | 11011 | 39/32 |
| 6 | 101011 | 110101 | 110101 | 101011 | 89/64 |
| 7 | 1010101 | 1010101 | 1101011 | 1101011 | 199/128 |

TABLE V. - x AND y VALUES FOR $F_{min}$

| n | $x_7 \cdots x_1$ | $y_{-1} \cdots y_{-7}$ | $x_7 \cdots x_1$ | $y_{-1} \cdots y_{-7}$ | $F_{min}$ |
|---|---|---|---|---|---|
| 2 | 01 | 11 | 01 | 11 | -3/4 |
| 3 | 011 | 101 | 001 | 111 | -7/8 |
| 4 | 0101 | 1101 | 0011 | 1011 | -17/16 |
| 5 | 01011 | 10101 | 00101 | 11011 | -39/32 |
| 6 | 010101 | 110101 | 001011 | 101011 | -89/64 |
| 7 | 0101011 | 1010101 | 0010101 | 1101011 | -199/128 |

TABLE VI. - x, $x_S$, AND y VALUES FOR $G_{max}$

| n | $x_{S7} \cdots x_{S1}$ | $x_7 \cdots x_1$ | $y_{-1} \cdots y_{-7}$ | $x_{S7} \cdots x_{S1}$ | $x_7 \cdots x_1$ | $y_{-1} \cdots y_{-7}$ | $G_{max}$ |
|---|---|---|---|---|---|---|---|
| 2 | 01 | 10 | 01 | 00 | 11 | 11 | 3/4 |
| 3 | 001 | 110 | 011 | 010 | 101 | 101 | 9/8 |
| 4 | 0101 | 1010 | 0101 | 0010 | 1101 | 1011 | 23/16 |
| 5 | 00101 | 11010 | 01011 | 01010 | 10101 | 10101 | 57/32 |
| 6 | 010101 | 101010 | 010101 | 001010 | 110101 | 101011 | 135/64 |
| 7 | 0010101 | 1101010 | 0101011 | 0101010 | 1010101 | 1010101 | 313/128 |

TABLE VII. - x, $x_S$, AND y VALUES FOR $H_{min}$

| n | $x_{S7} \cdots x_{S1}$ | $x_7 \cdots x_1$ | $y_{-1} \cdots y_{-7}$ | $x_{S7} \cdots x_{S1}$ | $x_7 \cdots x_1$ | $y_{-1} \cdots y_{-7}$ | $H_{min}$ |
|---|---|---|---|---|---|---|---|
| 2 | 11 | 01 | 11 | 11 | 01 | 11 | -3/2 |
| 3 | 101 | 011 | 101 | 111 | 001 | 111 | -7/4 |
| 4 | 1101 | 0011 | 1011 | 1011 | 0101 | 1101 | -17/8 |
| 5 | 10101 | 01011 | 10101 | 11011 | 00101 | 11011 | -39/16 |
| 6 | 110101 | 001011 | 101011 | 101011 | 010101 | 110101 | -89/32 |
| 7 | 1010101 | 0101011 | 1010101 | 1101011 | 0010101 | 1101011 | -199/64 |

occurred were found. These points are tabulated for various BRM in tables IV and V.

Appendix B presents an analysis of a BRM whose counter starts out with an arbitrary value. The basis of this analysis is to use equation (10) to obtain the maximum positive error and to use equation (11) to obtain the minimum negative error. For an n-stage BRM these values are

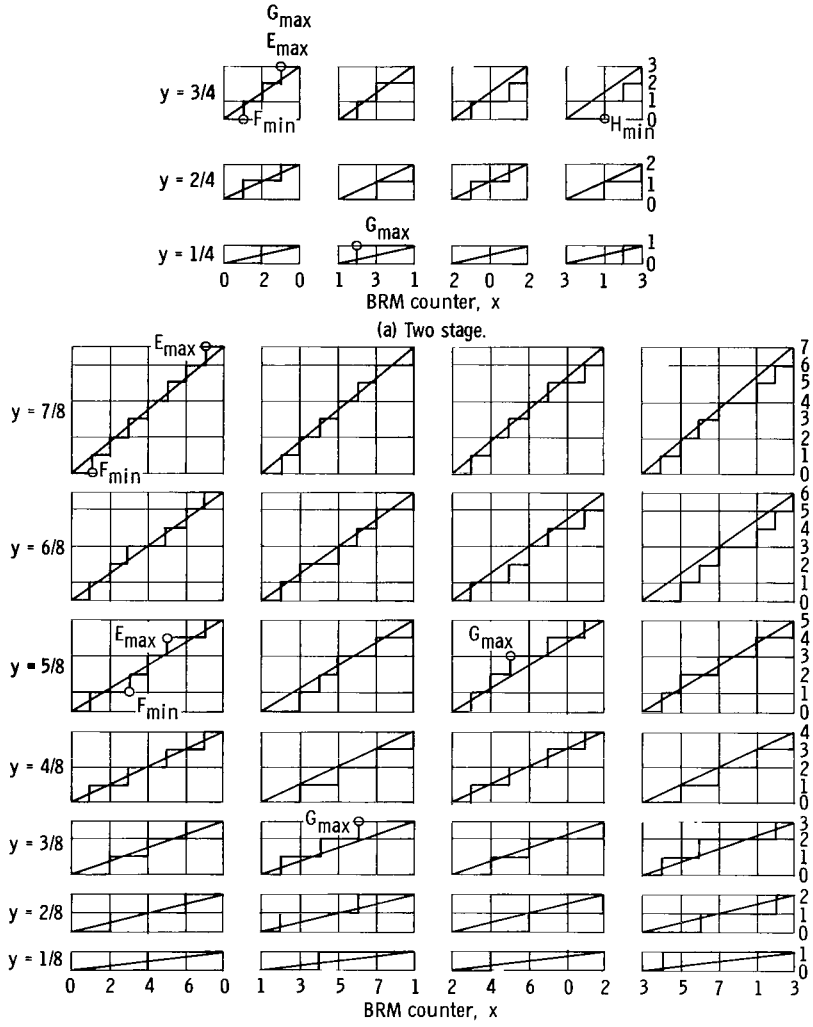
$$G_{max}(n) = \frac{1}{9} + \frac{n}{3} - \frac{(-1)^n}{9 \cdot 2^n} \tag{14}$$

$$H_{min}(n) = -\frac{7}{9} - \frac{n}{3} - \frac{(-1)^n}{9 \cdot 2^{n-1}} \tag{15}$$

These values form a bound for the generated round-off error. Fortunately, only two maximum and two minimum values may occur for a BRM (for $n > 2$), and therefore one can expect better results than would be predicted by these values. These values are plotted in figure 4(b) and are also presented together with the points at which they occur in tables VI and VII.

The problem that has been considered in appendixes A and B and in this section is illustrated in figure 5. The actual and desired outputs for a two- and a three-stage BRM are plotted in these figures for all starting values. As is illustrated, finding the multiplication error bounds by graphical means is not trivial. The points labeled $E_{max}$, $F_{min}$, $G_{max}$, and $H_{min}$ in these simple cases agree with those predicted in appendixes A and B.

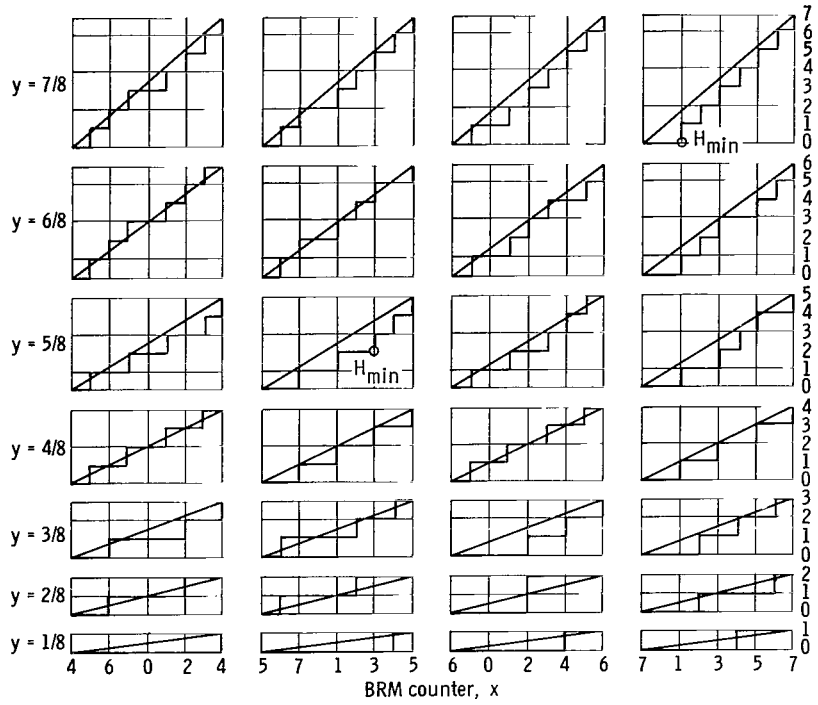
The error formulas arrived at in appendixes A and B can be arrived at directly by use of the exclusive OR operator. It was not used, however, because




(a) Two stage.

(b) Three stage.

Figure 5. - Output of BRM for all starting values.




(b) Concluded.

Figure 5. - Concluded.

some of the intermediate results obtained in the appendixes are interesting in their own right and would be bypassed by this alternate method of proof.

## CONCLUSIONS

The BRM has been shown to achieve approximate multiplication. The error has been defined as the difference between the actual output and $y\,\Delta x$ and is formulated as a bilinear expression. This error is shown to be dependent on the starting conditions of the BRM counter. The error formulas presented are analyzed in detail and explicit error bounds are given. These bounds are shown to increase by approximately 1/6 of a quanta per stage when the BRM counter starting value is zero, and by approximately 1/3 of a quanta per stage when the BRM counter starting value is arbitrary.

Lewis Research Center,
National Aeronautics and Space Administration,
Cleveland, Ohio, August 31, 1965.

## APPENDIX A

## MULTIPLICATION ERROR BOUNDS (ZERO STARTING)

In this section the error equation of an n-stage BRM whose counter starts with zero will be analyzed with the objective of obtaining tight error bounds. Nevertheless, some of the intermediate results that will be obtained in this section are interesting in their own right. Because this analysis is complex, formal methods will be used in this analysis. The basic outline is to use equation (7) to find the points where the maximum positive value is attained and then evaluate the equation at these points. In a similar manner, equation (9) will be used to find the minimum negative value.

The analysis is begun by stating and proving Lemma A.1.

Lemma A.1: A sufficient condition for E given by equation (7) to attain its maximum value is that $x_i = y_{-i}$.

Equation (7) may be rewritten as a bilinear expression so that the terms which are dependent on either $x_i$ or $y_{-i}$ are grouped together. The quantity A in the resultant expression is independent of either $x_i$ or $y_{-i}$:

$$E(x_i,y_{-i}) = A + \frac{1}{2} x_i y_{-i} - \frac{1}{2} x_i\left(\frac{1}{2} y_{-i-1} + \frac{1}{4} y_{-i-2} + \cdots + \frac{1}{2^{n-i}} y_{-n}\right) - \frac{1}{2} y_{-i}\left(\frac{1}{2} x_{i-1} + \frac{1}{4} x_{i-2} + \cdots + \frac{1}{2^{i-1}} x_1\right) \qquad (A1)$$

By direct evaluation, the value of this expression is as follows:

$$E(0,0) = A$$

$$E(1,0) = A - \frac{1}{2}\left(\frac{1}{2} y_{-i-1} + \cdots + \frac{1}{2^{n-i}} y_{-n}\right)$$

$$E(0,1) = A - \frac{1}{2}\left(\frac{1}{2} x_{i-1} + \cdots + \frac{1}{2^{i-1}} x_1\right) \qquad (A2)$$

Moreover, for the specific case when i is n, the value of equation (A1) is

$$E(0,0) = A$$

$$E(1,0) = A$$

$$E(0,1) = A - \frac{1}{2}\left(\frac{1}{2} x_{n-1} + \cdots + \frac{1}{2^{n-1}} x_1\right)$$

$$E(1,1) = A + \frac{1}{2} - \frac{1}{2}\left(\frac{1}{2} x_{n-1} + \cdots + \frac{1}{2^{n-1}} x_1\right) \qquad (A3)$$

This Lemma is proved by observing that the value of $E$ when $x_i = y_{-i} = 0$ is always greater than or equal to the value of $E$ in both cases when $x_i \neq y_{-i}$ and that for the $n^{th}$ component the value of $E$ is always greater when $x_n = y_{-n} = 1$. Although it was not needed in the proof of this Lemma, it can be similarly shown that this is also true for the first component.

Based on this Lemma, the maximum value of equation (7) will be obtained by finding the maximum of the quadratic form expression:

$$Q(x_1,x_2 \ldots x_n) = (x_1,x_2 \ldots x_n)M\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{A4}$$

Theorem A.1: For all values of the components $x_i$,

$$Q(1,x_2,x_3 \ldots x_n) > Q(0,x_2,x_3 \ldots x_n)$$

This theorem follows directly from the proof of Lemma A.1.

Theorem A.2: For all values of the components $x_i$,

$$Q(x_1,x_2 \ldots x_{n-1},1) > Q(x_1,x_2 \ldots x_{n-1},0)$$

This theorem follows directly from the proof of Lemma A.1.

Theorem A.3: For all values of the components $x_i$,

$$Q(1,x_2,x_3 \ldots x_{n-1},1) = Q(1,\bar{x}_2,\bar{x}_3 \ldots \bar{x}_{n-1},1)$$

where $\bar{x}_i$ is the complement of $x_i$.

The difference

$$Q(1,x_2,x_3 \ldots x_{n-1},1) - Q(1,\bar{x}_2,\bar{x}_3 \ldots \bar{x}_{n-1},1) =$$

$$(1,x_2 \ldots x_{n-1},1)M\begin{pmatrix} 1 \\ x_2 \\ \vdots \\ x_{n-1} \\ 1 \end{pmatrix} - (1,\bar{x}_2 \ldots \bar{x}_{n-1},1)M\begin{pmatrix} 1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_{n-1} \\ 1 \end{pmatrix} \tag{A5}$$

may be written equivalently as

$$(x_2 \cdots x_{n-1})K\begin{pmatrix} x_2 \\ \vdots \\ x_{n-1} \end{pmatrix} - (\bar{x}_2 \cdots \bar{x}_{n-1})K\begin{pmatrix} \bar{x}_2 \\ \vdots \\ \bar{x}_{n-1} \end{pmatrix} \tag{A6}$$

where

$$K = \begin{pmatrix} \frac{1}{2} - \frac{1}{2^2} - \frac{1}{2^{n-1}} & -\frac{1}{4} & -\frac{1}{8} & \cdots & \cdot & \cdots & \cdot \\ 0 & \frac{1}{2} - \frac{1}{2^3} - \frac{1}{2^{n-2}} & -\frac{1}{4} & \cdots & \cdot & \cdots & \cdot \\ \cdot & \cdot & \frac{1}{2} - \frac{1}{2^4} - \frac{1}{2^{n-3}} & -\frac{1}{4} \cdots & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \ddots & & & \\ & & & & \frac{1}{2} - \frac{1}{2^i} - \frac{1}{2^{n-i+1}} & \cdots & \vdots \\ & & & & & \ddots & \cdot \\ 0 & \cdot & \cdot & & & 0 & \frac{1}{2} - \frac{1}{2^{n-1}} - \frac{1}{2^2} \end{pmatrix}$$

Expanding equation (A6) and using the identity

$$x_i x_j - \bar{x}_i \bar{x}_j = x_i - \bar{x}_j \tag{A7}$$

to simplify the cross product terms give a typical term $x_i$ as

$$\left(x_i + \bar{x}_i\right)\left(\frac{1}{2^{n-i+1}} - \frac{1}{2^i}\right) \tag{A8}$$

Since $(x_i + \bar{x}_i) = 1$, equation (A6) is independent of the $x_i$ variable, and the contribution from this term is

$$\frac{1}{2^{n-i+1}} - \frac{1}{2^i} \tag{A9}$$

Similarly the contribution to the difference expressed in equation (A6) from the term involving $x_{n-i+1}$ is

$$\frac{1}{2^i} - \frac{1}{2^{n-i+1}} \tag{A10}$$

Consequently, the contribution to the difference expressed by equation (A6) by each element may be paired by the contribution from another element to cancel each other out of the expression. If $n$ is odd, the middle term cannot be paired. But, since this is the $(n + 1)/2$ term, by equation (A9), its contribution is

$$\frac{1}{2^{(n+1)/2}} - \frac{1}{2^{(n+1)/2}} = 0$$

Therefore, the value of the difference shown by equation (A5) is equal to zero. This implies that

$$Q(1,x_2 \ldots x_{n-1},1) = Q(1,\bar{x}_2 \ldots \bar{x}_{n-1},1)$$

Lemma A.2: For $v = x_2,x_3 \ldots x_i$ and $a = 1,0,1,0 \ldots 1,0$

$$Q(1,v,0,a,1) \geq Q(1,v,1,a,1)$$

where $\bar{v}$ and $\bar{a}$ are the component by component complement of $v$ and $a$, respectively.

By Theorem A.3 it is noted that

$$Q(1,v,1,a,1) = Q(1,\bar{v},0,\bar{a},1)$$

Therefore, the difference between the two quadratic forms of the Lemma can be expressed as

$$\delta = Q(1,\bar{v},0,a,1) - Q(1,\bar{v},0,\bar{a},1) \qquad (A11)$$

Partitioning the $M$ matrix of equation (A11) so that $M_1$ and $M_2$ are compatible with the vectors gives $\delta$ as

$$\delta = (1,\bar{v},0,a,1)M_1\begin{pmatrix}1\\ \bar{v}\\ 0\end{pmatrix} + (1,\bar{v},0,a,1)M_2\begin{pmatrix}a\\ 1\end{pmatrix}$$

$$- (1,\bar{v},0,\bar{a},1)M_1\begin{pmatrix}1\\ \bar{v}\\ 0\end{pmatrix} - (1,\bar{v},0,\bar{a},1)M_2\begin{pmatrix}\bar{a}\\ 1\end{pmatrix} \qquad (A12)$$

But it will be noted that

$$(1,\bar{v},0,a,1)M_1\begin{pmatrix}1\\ \bar{v}\\ 0\end{pmatrix} = (1,\bar{v},0,\bar{a},1)M_1\begin{pmatrix}1\\ \bar{v}\\ 0\end{pmatrix}$$

Therefore

$$\delta = (1,\bar{v},0,a,1)M_2\begin{pmatrix}a\\ 1\end{pmatrix} - (1,\bar{v},0,\bar{a},1)M_2\begin{pmatrix}\bar{a}\\ 1\end{pmatrix} \qquad (A13)$$

It will now be proved by induction on the length of $a$ that $\delta \geq 0$. It may be immediately verified that $\delta = 0$ for $a$ of length zero. Assume that $\delta_k \geq 0$, where $\delta_k$ is the value of $\delta$ when $a$ is of length $2k$. It will now be verified that $\delta_{k+1} \geq 0$:

$$\delta_{k+1} = \delta_k + \begin{pmatrix} 1 \\ \overline{x}_2 \\ \overline{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \overline{x}_i \\ 0_{i+1} \\ 1_{i+2} \\ 0_{i+3} \\ \cdot \\ \cdot \\ \cdot \\ 0_{i+2k-1} \\ 1_{i+2k} \\ 0_{i+2k+1} \\ 1_{i+2k+2} \\ 0_{i+2k+3} \\ 1_{i+2k+4} \end{pmatrix} \begin{pmatrix} -1/2^{i+2k+4} \\ -1/2^{i+2k+3} \\ -1/2^{i+2k+2} \\ \cdot \\ \cdot \\ \cdot \\ -1/2^{2k+5} \\ -1/2^{2k+4} \\ -1/2^{2k+3} \\ -1/2^{2k+2} \\ \cdot \\ \cdot \\ \cdot \\ -1/64 \\ -1/32 \\ -1/16 \\ -1/8 \\ -1/4 \\ 1/2 \end{pmatrix} - \begin{pmatrix} 1 \\ \overline{x}_2 \\ \overline{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \overline{x}_i \\ 0_{i+1} \\ 0_{i+2} \\ 1_{i+3} \\ \cdot \\ \cdot \\ \cdot \\ 1_{i+2k-1} \\ 0_{i+2k} \\ 1_{i+2k+1} \\ 0_{i+2k+2} \\ 1_{i+2k+3} \\ 1_{i+2k+4} \end{pmatrix} \begin{pmatrix} -1/2^{i+2k+3} \\ \cdot \\ \cdot \\ \cdot \\ \\ \\ \\ \cdot \\ -1/2^{2k+1} \\ \\ \\ \\ \cdot \\ \cdot \\ \cdot \\ -1/4 \\ 1/2 \\ 0 \end{pmatrix}$$

$$- \begin{pmatrix} 1 \\ \overline{x}_2 \\ \overline{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \overline{x}_i \\ 0_{i+1} \\ 0_{i+2} \\ 1_{i+3} \\ \cdot \\ \cdot \\ \cdot \\ 1_{i+2k-1} \\ 0_{i+2k} \\ 1_{i+2k+1} \\ 0_{i+2k+2} \\ 1_{i+2k+3} \\ 1_{i+2k+4} \end{pmatrix} \begin{pmatrix} -1/2^{i+2k+4} \\ \cdot \\ \cdot \\ \cdot \\ \\ \\ -1/2^{2k+5} \\ -1/2^{2k+4} \\ \cdot \\ -1/2^{2k+2} \\ \\ \\ \\ \cdot \\ \cdot \\ \cdot \\ -1/4 \\ 1/2 \end{pmatrix} + \begin{pmatrix} 1 \\ \overline{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \overline{x}_i \\ 0_{i+1} \\ 0_{i+2} \\ 1_{i+3} \\ \cdot \\ \cdot \\ \cdot \\ 0_{i+2k} \\ 1_{i+2k+1} \\ 1_{i+2k+2} \end{pmatrix} \begin{pmatrix} -1/2^{i+2k+2} \\ -1/2^{i+2k+1} \\ \cdot \\ \cdot \\ \cdot \\ -1/2^{2k+3} \\ -1/2^{2k+2} \\ -1/2^{2k+1} \\ -1/2^{2k} \\ \cdot \\ \cdot \\ \cdot \\ -1/8 \\ -1/4 \\ 1/2 \end{pmatrix} \qquad \text{(A14)}$$

The 0's and 1's in equation (A14) are subscripted to show their position in the vector, and the vector itself is displayed as a column rather than a row in order to show the correspondence between the terms that must be multiplied to form the value.

Multiplying out the terms of equation (A14) yields

$$\delta_{k+1} = \delta_k + \frac{1}{2^{2k+3}} - \frac{1}{2^{i+2k+3}} - \frac{1}{2^{2k+4}} \sum_{j=0}^{i-2} \frac{1}{2^j} \overline{x}_{i-j} \qquad \text{(A15)}$$

But, by direct evaluation

$$\frac{1}{2^{2k+4}} \sum_{j=0}^{i-2} \frac{1}{2^j} \bar{x}_{i-j} + \frac{1}{2^{2k+4}} \frac{1}{2^{i-1}} < \frac{2}{2^{2k+4}} = \frac{1}{2^{2k+3}}$$

Therefore, $\delta_{k+1} > 0$.

Lemma A.3: For $v = x_2, x_3 \ldots x_i$ and $a = 1,0,1,0 \ldots 1,0$

$$Q(1,\bar{v},1,0,a,1) \geq Q(1,v,0,0,a,1)$$

where $\bar{v}$ and $\bar{a}$ are defined as in Lemma A.2.

Proceeding in a manner similar to Lemma A.2 shows by Theorem A.3 that

$$Q(1,v,0,0,a,1) = Q(1,\bar{v},1,1,\bar{a},1)$$

Therefore, the difference between the two quadratic forms of the Lemma can be expressed as

$$\delta = Q(1,\bar{v},1,0,a,1) - Q(1,\bar{v},1,1,\bar{a},1) \tag{A16}$$

Partitioning the M matrix of equation (A16) so that $M_1$ and $M_2$ are compatible with the vectors gives $\delta$ in the form

$$\delta = (1,\bar{v},1,0,a,1)M_1\begin{pmatrix}1\\ \bar{v}\\ 1\end{pmatrix} + (1,\bar{v},1,0,a,1)M_2\begin{pmatrix}0\\ a\\ 1\end{pmatrix}$$

$$- (1,\bar{v},1,1,\bar{a},1)M_1\begin{pmatrix}1\\ \bar{v}\\ 1\end{pmatrix} - (1,\bar{v},1,1,\bar{a},1)M_2\begin{pmatrix}1\\ \bar{a}\\ 1\end{pmatrix} \tag{A17}$$

But expanding the first and third terms of equation (A17) shows that

$$(1,\bar{v},1,0,a,1)M_1\begin{pmatrix}1\\ \bar{v}\\ 1\end{pmatrix} = (1,\bar{v},1,1,\bar{a},1)M_1\begin{pmatrix}1\\ \bar{v}\\ 1\end{pmatrix}$$

Therefore,

$$\delta = (1,\bar{v},1,0,a,1)M_2\begin{pmatrix}0\\ a\\ 1\end{pmatrix} - (1,\bar{v},1,1,\bar{a},1)M_2\begin{pmatrix}1\\ \bar{a}\\ 1\end{pmatrix} \tag{A18}$$

If $\delta_k$ is used to denote the value of $\delta$ when a is of length 2k it will now be proved by induction on the length of a that

$$\delta_k \geq \frac{1}{2^{2k+3}} \bar{x}_i + \frac{1}{2^{2k+4}} \bar{x}_{i-1} + \ldots + \frac{1}{2^{i+k+1}} \bar{x}_2 + \frac{1}{2^{i+2k+2}} > 0 \tag{A19}$$

First, note that for a of length zero

$$\delta_0 = \begin{pmatrix} 1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \bar{x}_i \\ 1_{i+1} \\ 0_{i+2} \\ 1_{i+3} \end{pmatrix} \begin{pmatrix} -1/2^{i+3} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ -1/16 \\ -1/8 \\ -1/4 \\ 1/2 \end{pmatrix} - \begin{pmatrix} 1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \bar{x}_i \\ 1_{i+1} \\ 1_{i+2} \\ 1_{i+3} \end{pmatrix} \begin{pmatrix} -1/2^{i+2} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ -1/8 \\ -1/4 \\ 1/2 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \bar{x}_i \\ 1_{i+1} \\ 1_{i+2} \\ 1_{i+3} \end{pmatrix} \begin{pmatrix} -1/2^{i+3} \\ -1/2^{i+2} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ -1/16 \\ -1/8 \\ -1/4 \\ 1/2 \end{pmatrix} \qquad \text{(A20)}$$

$$= \frac{1}{2^3}\bar{x}_i + \frac{1}{2^4}\bar{x}_{i-1} + \cdots + \frac{1}{2^{i+1}}\bar{x}_2 + \frac{1}{2^{i+2}}$$

The notation used in equation (A20) is similar to that used in proving Lemma A.2. Assume that $\delta_k > 0$. It will now be shown that $\delta_{k+1} > 0$:

$$\delta_{k+1} = \delta_{k+} + \begin{pmatrix} 1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \bar{x}_i \\ 1_{i+1} \\ 0_{i+2} \\ 1_{i+3} \\ 0_{i+4} \\ \cdot \\ \cdot \\ \cdot \\ 1_{i+2k+1} \\ 0_{i+2k+2} \\ 1_{i+2k+3} \\ 0_{i+2k+4} \\ 1_{i+2k+5} \end{pmatrix} \begin{pmatrix} -1/2^{i+2k+5} \\ -1/2^{i+2k+4} \\ \cdot \\ \cdot \\ \cdot \\ -1/2^{2k+6} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ -1/8 \\ -1/4 \\ 1/2 \end{pmatrix} - \begin{pmatrix} 1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \bar{x}_i \\ 1_{i+1} \\ 1 \\ 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \\ 0 \\ 1 \\ 1_{i+2k+5} \end{pmatrix} \begin{pmatrix} -1/2^{i+2k+4} \\ \cdot \\ \cdot \\ \cdot \\ -1/2^{2k+5} \\ -1/2^{2k+4} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ -1/4 \\ 1/2 \\ 0 \end{pmatrix}$$

$$+ \begin{pmatrix} 1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \bar{x}_i \\ 1_{i+1} \\ 0_{i+2} \\ 1_{i+3} \\ 0_{i+4} \\ \cdot \\ \cdot \\ \cdot \\ 1_{i+2k+1} \\ 0_{i+2k+2} \\ 1_{i+2k+3} \\ 0_{i+2k+4} \\ 1_{i+2k+5} \end{pmatrix} \begin{pmatrix} -1/2^{i+2k+5} \\ \\ \\ -1/2^{2k+6} \\ \\ -1/2^{2k+4} \\ \cdot \\ \cdot \\ \cdot \\ \\ \cdot \\ \cdot \\ \cdot \\ -1/4 \\ 1/2 \end{pmatrix} + \begin{pmatrix} 1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \cdot \\ \cdot \\ \cdot \\ \bar{x}_i \\ 1_{i+1} \\ 1_{i+2} \\ 0_{i+3} \\ 1_{i+4} \\ \cdot \\ \cdot \\ \cdot \\ 0_{i+2k+1} \\ 1_{i+2k+2} \\ 1_{i+2k+3} \end{pmatrix} \begin{pmatrix} -1/2^{i+2k+3} \\ \cdot \\ \cdot \\ \cdot \\ -1/2^{2k+4} \\ \\ \\ \\ \\ \cdot \\ \cdot \\ \cdot \\ -1/4 \\ 1/2 \end{pmatrix} \qquad \text{(A21)}$$

Multiplying equation (A21) gives

$$\delta_{k+1} = \delta_k - \frac{1}{2^{2k+5}} \bar{x}_i - \cdots - \frac{1}{2^{i+2k+3}} \bar{x}_2 - \frac{1}{2^{i+2k+4}} \qquad \text{(A22)}$$

Using equation (A19) gives the right side of equation (A22) as

$$> \frac{1}{2^{2k+3}} \bar{x}_i + \cdots + \frac{1}{2^{i+2k+1}} \bar{x}_2 + \frac{1}{2^{i+2k+2}} - \frac{1}{2^{2k+5}} \bar{x}_i - \cdots$$

$$- \frac{1}{2^{i+2k+3}} \bar{x}_2 - \frac{1}{2^{i+2k+4}} > \frac{1}{2^{2k+5}} \bar{x}_i + \cdots \frac{1}{2^{i+2k+3}} \bar{x}_2 + \frac{1}{2^{i+2k+4}} > 0$$

Theorem A.4: There exists a $v^*$ such that for all $v$

$$Q(1,v^*,0,a,1) \geq Q(1,v,x_{i+1},a,1)$$

where $v$ and $a$ are defined as before.

First, note that $Q(1,v^{**},0,a,1) \geq Q(1,v,1,a,1)$ because, suppose it were false, then there would exist a $\overset{*}{v}{}^{**}$ such that for all $v$

$$Q(1,\overset{*}{v}{}^{**},1,a,1) > Q(1,v,0,a,1)$$

But from Lemma A.2

$$Q(1,\overline{\overset{*}{v}}{}^{**},0,a,1) \geq Q(1,\overset{*}{v}{}^{**},1,a,1)$$

Therefore, a contradiction exists. Moreover, note that

$$Q(1,\overset{**}{v}{}^{**},0,a,1) \geq Q(1,v,0,a,1)$$

that is, there is a largest. Therefore, Theorem A.4 is proved by choosing either $v^{**}$ or $\overset{**}{v}{}^{**}$ for $v^*$; that is, whichever makes $Q(1,v^*,0,a,1)$ the largest.

Theorem A.5: There exists a $v^*$ such that for all $v$

$$Q(1,v^*,1,0,a,1) \geq Q(1,v,x_{i+1},0,a,1)$$

First, note that $Q(1,v^{**},1,0,a,1) \geq Q(1,v,0,0,a,1)$ because, suppose it were false, then there would exist a $\overset{*}{v}{}^{**}$ such that for all $v$

$$Q(1,\overset{*}{v}{}^{**},0,0,a,1) > Q(1,v,1,0,a,1)$$

But from Lemma A.3

$$Q(1,\overline{\overset{*}{v}}{}^{**},1,0,a,1) \geq Q(1,\overset{*}{v}{}^{**},0,0,a,1)$$

Therefore, a contradiction exists. Moreover, a $v^{**}$ can be chosen so that

$$Q(1,v^{**},0,0,a,1) \geq Q(1,v,0,0,a,1)$$

Therefore

$$Q(1,v^{*},1,0,a,1) \geq Q(1,v,x_{i+1},0,a,1)$$

Theorem A.6: There exists a $v^{*}$ such that for all $v$

$$Q(1,v^{*},0,1) \geq Q(1,v,x_{i+1},1)$$

First, note that there exists a $v^{**}$ such that $Q(1,v^{**},0,1) \geq Q(1,v,1,1)$, because; suppose it were false, then there would exist a $v^{**}$ such that for all $v$

$$Q(1,v^{**},1,1) > Q(1,v,0,1)$$

But by Theorem A.3

$$Q(1,\bar{v}^{**},0,1) = Q(1,v^{**},1,1)$$

Therefore, a contradiction exists. Moreover, a $v^{**}$ can be chosen such that

$$Q(1,v^{**},0,1) \geq Q(1,v,0,1)$$

Therefore,

$$Q(1,v^{*},0,1) \geq Q(1,v,x_{i+1},1)$$

It will now be demonstrated by an example how these theorems can be used to obtain the value of $x$ so that the error is the maximum positive value. Consider a seven-stage BRM. By Theorem A.1 and Theorem A.2

$$Q(1,x_2,x_3,x_4,x_5,x_6,1) \geq Q(x_1,x_2,x_3,x_4,x_5,x_6,x_7)$$

by Theorem A.6

$$Q(1,x_2^{*},x_3^{*},x_4^{*},x_5^{*},0,1) \geq Q(1,x_2,x_3,x_4,x_5,x_6,1)$$

by Theorem A.5

$$Q(1,x_2^{**},x_3^{**},x_4^{**},1,0,1 \geq Q(1,x_2^{*},x_3^{*},x_4^{*},x_5^{*},0,1)$$

by Theorem A.4

$$Q(1,x_2^{*},x_3^{*},0,1,0,1) \geq Q(1,x_2^{**},x_3^{**},x_4^{**},1,0,1)$$

by Theorem A.5

$$Q(1,x_2^{**},1,0,1,0,1) \geq Q(1,x_2^{*},x_3^{*},0,1,0,1)$$

by Theorem A.6

$$Q(1,0,1,0,1,0,1) \geq Q(1,x_2^{**},1,0,1,0,1)$$

Putting these inequalities together results in

$$Q(1,0,1,0,1,0,1) \geq Q(x_1,x_2,x_3,x_4,x_5,x_6,x_7)$$

Moreover, by Theorem A.3

$$Q(1,1,0,1,0,1,1) \geq Q(x_1,x_2,x_3,x_4,x_5,x_6,x_7)$$

Using the theorems in the pattern illustrated by the example makes it easy to verify that the maximum positive value will occur at the points shown in table IV (p. 9).

The maximum positive value of the error may be expressed concisely as follows: Let $E_{max}(k)$ denote this value for a k-stage BRM. If k is odd,

$$E_{max}(k+2) = E_{max}(k) + \begin{pmatrix} 1_1 \\ 0_2 \\ 1_3 \\ \cdot \\ \cdot \\ \cdot \\ 1_k \\ 0_{k+1} \\ 1_{k+2} \end{pmatrix} \begin{pmatrix} -1/2^{k+2} \\ \\ \\ \cdot \\ \cdot \\ \cdot \\ -1/8 \\ -1/4 \\ 1/2 \end{pmatrix} \qquad \text{(A23)}$$

Evaluating equation (A23) yields the difference equation

$$E_{max}(k+2) = E_{max}(k) + \frac{1}{3}\left(1 + \frac{1}{2^{k+2}}\right) \qquad \text{(A24)}$$

Solving equation (A24) for an n-stage BRM in terms of $E_{max}(1)$ yields

$$E_{max}(n) = E_{max}(1) - \frac{1}{9} + \frac{n}{6} - \frac{1}{9 \cdot 2^n}$$

But $E_{max}(1) = 1/2$.

Therefore, the maximum positive error of an n-stage BRM, where n is odd, is:

$$E_{max}(n) = \frac{7}{18} + \frac{n}{6} - \frac{1}{9 \cdot 2^n} \qquad \text{(A25)}$$

If k is even,

$$E_{max}(k+2) = E_{max}(k) + \begin{pmatrix} 1_1 \\ 1_2 \\ 0_3 \\ 1_4 \\ \cdot \\ \cdot \\ \cdot \\ 0_{k-1} \\ 1_k \\ 0_{k+1} \\ 1_{k+2} \end{pmatrix} \begin{pmatrix} -1/2^{k+2} \\ -1/2^{k+1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ -1/8 \\ -1/4 \\ 1/2 \end{pmatrix} = E_{max}(k) + \frac{1}{3} - \frac{1}{3}\frac{1}{2^{k+2}} \tag{A26}$$

Solving this difference equation for an n-stage BRM in terms of $E_{max}(2)$ and then evaluating the resultant expression for $E_{max}(2) = 3/4$ yield

$$E_{max}(n) = \frac{7}{18} + \frac{n}{6} + \frac{1}{9 \cdot 2^n} \tag{A27}$$

Combining equations (A25) and (A27) gives a closed-form equation for the maximum positive error of an n-stage BRM:

$$E_{max}(n) = \frac{7}{18} + \frac{n}{6} + \frac{(-1)^n}{9 \cdot 2^n} \tag{A28}$$

The minimum negative value for an n-stage BRM can be obtained by applying equation (9). Comparing the form of equation (7) with equation (9) shows that the previous results can be utilized with a slight modification. In particular, the value of the minimum is equal to the negative of the maximum and occurs at points that are the 2's complement of the maximum value. Consequently, the minimum negative values will occur at the points shown in table V (p. 9).

## APPENDIX B

## MULTIPLICATION ERROR BOUNDS (ARBITRARY STARTING)

The error formulas given by equations (10) and (11) express the multiplication error of a BRM whose counter starts with an arbitrary value. Equation (10) is the error formula resulting when the maximum value of the actual output is considered at the points of discontinuities. Equation (11) is the companion equation resulting when the minimum value of the actual output is considered at these points. In this section, these error formulas will be analyzed with the objective of obtaining error bounds for a BRM with this added degree of freedom. First, equation (10) is analyzed to obtain the maximum positive error of an n-stage BRM.

It is convenient for this discussion to define a vector $b$ so that

$$\begin{pmatrix} b_{-1} \\ b_{-2} \\ \cdot \\ \cdot \\ \cdot \\ b_{-n} \end{pmatrix} = M \begin{pmatrix} y_{-1} \\ y_{-2} \\ \cdot \\ \cdot \\ \cdot \\ y_{-n} \end{pmatrix} \qquad \text{(B1)}$$

Theorem B.1: For all $x_k$ and $x_{Sk}$ in equation (10), $G \leq \sum_{i=1}^{n} |b_{-i}|$. Moreover, if $y_{-k} = x_k = \bar{x}_{Sk}$, $G = \sum_{i=1}^{n} |b_{-i}|$, where $\bar{x}_{Sk}$ denotes the complement of $x_{Sk}$.

The elements of the vector defined by equation (B1) are

$$b_{-1} = \frac{1}{2} y_{-1} - \left( \frac{1}{4} y_{-2} + \frac{1}{8} y_{-3} + \cdots + \frac{1}{2^n} y_{-n} \right)$$

$$b_{-2} = \frac{1}{2} y_{-2} - \left( \frac{1}{4} y_{-3} + \frac{1}{8} y_{-4} + \cdots + \frac{1}{2^{n-1}} y_{-n} \right)$$

$$\vdots$$

$$b_{-k} = \frac{1}{2} y_{-k} - \left( \frac{1}{4} y_{-k-1} + \frac{1}{8} y_{-k-2} + \cdots + \frac{1}{2^{n-k+1}} y_{-n} \right)$$

$$\vdots$$

$$b_{-n} = \frac{1}{2} y_{-n}$$

Since $1/2 > 1/4 + \cdots + 1/2^n$,

$$b_{-k} \leq 0 \quad \text{if} \quad y_{-k} = 0$$

$$> 0 \quad \text{if} \quad y_{-k} = 1 \tag{B2}$$

It is next noted that each element $(x_k - x_{Sk})$ may have only three possible values; that is, 0,1, or -1. This may be verified by direct computation:

| $x_k$ | $x_{Sk}$ | $x_k - x_{Sk}$ |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | -1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

(B3)

Since $|x_k - x_{Sk}| \leq 1$ then

$$G = (x_1 - x_{S1})b_{-1} + (x_2 - x_{S2})b_{-2} + \cdots + (x_n - x_{Sn})b_{-n} \leq \sum_{i=1}^{n} |b_{-i}| \tag{B4}$$

A sufficient condition for $G$ to attain the upper bound of equation (B4), that is, $\sum_{i=1}^{n} |b_{-i}|$, is that

$$(x_k - x_{Sk}) = -1 \quad \text{if} \quad b_{-k} \leq 0$$

$$= +1 \quad \text{if} \quad b_{-k} > 0 \tag{B5}$$

Combining equations (B2), (B3), and (B5) results in

| $y_{-k}$ | $b_{-k}$ | $x_k - x_{Sk}$ | $x_k$ | $x_{Sk}$ |
|---|---|---|---|---|
| 0 | $\leq 0$ | -1 | 0 | 1 |
| 1 | $>0$ | +1 | 1 | 0 |

Therefore, $y_{-k} = x_k = \bar{x}_{Sk}$ is a sufficient condition for $G = \sum_{i=1}^{n} |b_{-i}|$.

As a consequence of Theorem B.1 the maximum of $G$, denoted by $G_{max}$, is such that $G_{max} = \max_{y} \sum_{i=1}^{n} |b_{-i}|$. The procedure to be followed is to find the value $y$ where the maximum of $\sum_{i=1}^{n} |b_{-i}|$ is attained and then evaluating this function. In order to aid this analysis the notation $b_{-i}(y)$ is introduced, where $y = y_{-1}y_{-2} \cdots y_{-n}$ and $b_{-i}(y)$ denotes the value $b_{-i}$ for the vector

$(y_{-1}, y_{-2} \cdots y_{-n})$. For an n-stage BRM, all possible $b_{-i}(y)$ values may be obtained by multiplying M with all possible values of y. This particular matrix is called $B_n$.

$$B_n = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & -\frac{1}{8} & \cdots & -\frac{1}{2^n} \\ 0 & \frac{1}{2} & -\frac{1}{4} & & -\frac{1}{2^{n-1}} \\ \vdots & & \ddots & & \\ & & & \frac{1}{2} & -\frac{1}{4} \\ 0 \cdots & & & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0\,0\cdots1\,1\cdots1 \\ 0\,0 \quad\; 0\,0 \quad\; 1 \\ \vdots \quad\; \vdots \quad\; \vdots \\ 0\,1\cdots0\,0\cdots1 \\ 1\,0\cdots0\,1\cdots1 \end{pmatrix} = \begin{pmatrix} b_{-1}(1) & b_{-1}(2) & \cdots & b_{-1}(2^n-1) \\ b_{-2}(1) & b_{-2}(2) & \cdots & b_{-2}(2^n-1) \\ \vdots & \vdots & & \vdots \\ b_{-n}(1) & b_{-n}(2) & \cdots & b_{-n}(2^n-1) \end{pmatrix} \quad \text{(B6)}$$

A few examples will help clarify equation (B6)

Two-stage BRM:

$$B_2 = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

Three-stage BRM:

$$B_3 = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & -\frac{1}{8} \\ 0 & \frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0\,0\,0\,1\,1\,1\,1 \\ 0\,1\,1\,0\,0\,1\,1 \\ 1\,0\,1\,0\,1\,0\,1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{8} & -\frac{1}{4} & -\frac{3}{8} & \frac{1}{2} & \frac{3}{8} & \frac{1}{4} & \frac{1}{8} \\ -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

Four-stage BRM:

$$B_4 = \begin{pmatrix} -\frac{1}{16} & -\frac{2}{16} & -\frac{3}{16} & -\frac{4}{16} & -\frac{5}{16} & -\frac{6}{16} & -\frac{7}{16} & \frac{1}{2} & \frac{7}{16} & \frac{6}{16} & \frac{5}{16} & \frac{4}{16} & \frac{3}{16} & \frac{2}{16} & \frac{1}{16} \\ -\frac{1}{8} & -\frac{1}{4} & -\frac{3}{8} & \frac{1}{2} & \frac{3}{8} & \frac{1}{4} & \frac{1}{8} & 0 & & & & & & & \\ & & & 0 & -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & & & & B_3 & & & \\ & B_2 & & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & & & & & & & \end{pmatrix}$$

Theorem B.2: For all values of $y$, $|b_{-i}(y)| = |b_{-i}(2^n - y)|$.

This result follows by induction on the number of stages $k$. It was shown as an example for the $k = 2$ case. Assume it is true for the $k = n - 1$ case Then the $k = n$ case is

$$B_n = \begin{pmatrix} -\frac{1}{2^n} & -\frac{2}{2^n} & \cdots & -\frac{2^{n-1}-1}{2^n} & \frac{1}{2} & \frac{2^{n-1}-1}{2^n} & \cdots & \frac{1}{2^n} \\ & & B_{n-1} & & \begin{matrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{matrix} & & B_{n-1} & \end{pmatrix}$$

where this case is partitioned to show its structure. This theorem is obviously true for the first row. The $b_{-i}(y)$ element for the $n - 1$ case is now the $b_{-i-1}(y)$ and the $b_{-i-1}(2^n + y)$ elements of the $n$ case; and the $b_{-i}(2^n - y)$ element of the $n - 1$ case is now the $b_{-i-1}(2^n - y)$ and the $b_{-i-1}(2^n + 2^n - y)$ elements of the $B_n$ case. By the induction hypothesis $|b_{-1}(y)| = |b_{-i}(2^n - y)|$ for the $n - 1$ case. Therefore, these elements for the $n^{th}$ case yield

$$|b_{-i-1}(y)| = |b_{-i-1}(2^{n+1} - y)| \tag{B7}$$

and

$$|b_{-i-1}(2^n - y)| = |b_{-i-1}(2^n + y)| \tag{B8}$$

Substituting $u = 2^n - y$ into equation (B8) gives

$$|b_{-i-1}(u)| = |b_{-i-1}(2^{n+1} - u)|$$

which completes the proof.

Lemma B.1: There exists a $y^*$ in the domain $0\,1\,0\,\ldots\,0\,0 \le y^* \le 0\,1\,1\,1\,\ldots\,1\,1$ so that

$$\sum_i |b_{-i}(y^*)| \ge \sum_i |b_{-i}(y)|$$

This theorem states that $G_{max}$ is attained in the domain $0\,1\,0\,\ldots\,0\,0 \le y \le 0\,1\,1\,\ldots\,1\,1$. As a result of Theorem B.2, the search for a point where $G_{max}$ is attained can be immediately restricted to the $y$ domain $0 < y \le 1\,0\,0\,\ldots\,0\,0$. Consider $B_n$ for these values of $y$:

$$B_n = \begin{pmatrix} -\frac{1}{2^n} & -\frac{2}{2^n} & \cdots & -\frac{1}{4} & \cdots & \frac{1}{2} & \cdots \\ & & & \frac{1}{2} & & 0 & \\ & B_L & & 0 & B_R & 0 & \\ & & & \vdots & & \vdots & \\ & & & 0 & & 0 & \cdots \end{pmatrix}$$

The vector $\begin{pmatrix} -1/4 \\ 1/2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ results from the $y$ vector $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. The vector $\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ can be immediately ruled out. The structure $B_L \begin{pmatrix} 1/2 \\ 0 \\ 0 \\ \vdots \end{pmatrix} B_R$ in the preceding matrix is $B_{n-1}$. Because of Theorem B.2, the absolute values of the elements in $B_L$ are identical to the absolute values of the elements $B_R$. Moreover, since the elements of first row, that is, $|b_{-1}(y)|$, increase as $y$ increases, then for each column sum to the left of $\begin{pmatrix} -1/4 \\ 1/2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ there is a column sum to the right that exceeds it. Therefore, $G_{max}$ must lie to the right.

As a result of Theorem B.2 and Lemma B.1, there must be at least two values of $y$ where $G_{max}$ is attained. For an n-stage BRM, the $y$ value corresponding to $G_{max}$ on the left of $y = 1000 \ldots 00$ will be called $L_n$; the one on the right of $y = 100 \ldots 00$ will be called $R_n$.

Lemma B.2: For an n-stage BRM

$$OR_{n-1} \leq L_n < 0\ 1\ 1 \ldots 1\ 1$$

This result follows from the proof of Lemma B.1. Since the first row; that is,

$|b_{-1}(y)|$, is increasing, $G_{max}$ must lie between the right maximum of $B_{n-1}$ and the rightmost value of Lemma B.1.

Lemma B.3: For an n-stage BRM

$$0\,1\,0\,0\,\ldots\,0 \leq L_n \leq 0\,1\,L_{n-2}$$

Consider $B_n$ for the values of y of Lemma B.1; that is, $0\,1\,0\,0\,\ldots\,0 \leq y \leq 0\,1\,1\,\ldots\,1$:

$$B_n = \begin{pmatrix} -\frac{1}{4} & -\frac{2^{n-2}+1}{2^n} & -\frac{2^{n-2}+2}{2^n} & \cdots & \frac{1}{2} \\ \frac{1}{2} & \frac{2^{n-2}-1}{2^{n-1}} & \frac{2^{n-2}-2}{2^{n-1}} & \cdots & 0 \\ & & B_{n-2} & & \vdots \\ & & & & 0 \end{pmatrix}$$

The column by column sums of the absolute value of the elements of the first two rows are

$$\frac{3}{4},\ \frac{2^{n-1}+2^{n-2}-1}{2^n},\ \frac{2^{n-1}+2^{n-2}-2}{2^n},\ \ldots$$

Therefore, this is a decreasing sequence. Since by Theorem B.2 and Lemma B.1, $B_{n-2}$ attains a maximum for at least two values, then $G_{max}$ must lie between the leftmost value of Lemma B.1 and $0\,1\,L_{n-2}$.

Theorem B.3: For an n-stage BRM

$$0R_{n-1} \leq L_n \leq 0\,1\,L_{n-2}$$

This theorem is the combination of Lemma B.2 and Lemma B.3.

Theorem B.4: $R_n$ and $L_n$ are unique and $0R_{n-1} = L_n = 0\,1\,L_{n-2}$. This theorem follows immediately from the proofs of Lemma B.2 and Lemma B.3 by using the principle of strong induction as the method of proof; that is, assume it is true for $k < n$ and prove for the case $k = n$. The values of y where $G_{max}$ is attained can be obtained by using Theorem B.4. These values are listed in table VIII.

The BRM counter value and starting value corresponding to the y values listed in table VIII can be obtained by Theorem B.1. These values are listed in table VI (p. 10).

An equation for $G_{max}$ as a function of n may be obtained by a procedure similar to that used to obtain $E_{max}$. In particular, if the pattern established

TABLE VIII. - VALUES OF y WHERE $G_{max}$ IS ATTAINED

| | $011_{n-2}$ | $OR_{n-1}$ | $L_n$ | $R_n$ |
|---|---|---|---|---|
| n | $y_{-1} \cdots y_{-6}$ | $y_{-1} \cdots y_{-6}$ | $y_{-1} \cdots y_{-6}$ | $y_{-1} \cdots y_{-6}$ |
| 1 | | | 1 | 1 |
| 2 | | | 01 | 11 |
| 3 | 011 | 011 | 011 | 101 |
| 4 | 0101 | 0101 | 0101 | 1011 |
| 5 | 01011 | 01011 | 01011 | 10101 |
| 6 | 010101 | 010101 | 010101 | 101011 |

in table VIII is used a difference equation may be written for n even, and a difference equation may be written for n odd. Combining the solutions of these difference equations gives $G_{max}$ as a function of n:

$$G_{max}(n) = \frac{1}{9} + \frac{n}{3} - \frac{(-1)^n}{9 \cdot 2^n} \tag{B9}$$

Noting the similarity between the rightmost term of equation (11) to that of equation (10), establishes immediately a minimum error bound when the BRM counter starts with an arbitrary value:

$$H_{min}(n) \geq -\frac{10}{9} - \frac{n}{3} + \frac{(-1)^n}{9 \cdot 2^n} \tag{B10}$$

A tight error bound may be obtained as follows: Expanding equation (11) gives the equation for H in the form

$$H = -x_{SR}y_{-R} - \left[(C_1 - C_{S1})b_{-1} + (C_2 - C_{S2})b_{-2} + \cdots + (C_n - C_{Sn})b_{-n}\right] \tag{B11}$$

In the initial value of the BRM counter, $x_{SR}$ identifies the rightmost 1. It may be simply argued that, if a binary number has a rightmost 1 in position R, its 2's complement also has a 1 in that position and, moreover, has zeros at all positions j, where $j < R$. Therefore, $x_{SR} = C_{SR}$ and $C_{Sj} = 0$ for all $j < R$. Equation (B11) may be expressed as

$$-H = C_{SR}y_{-R} + \left(C_1b_{-1} + C_2b_{-2} + \cdots + C_{R-1}b_{-R+1}\right) + (C_R - C_{SR})b_{-R} + \left[(C_{R+1} - C_{SR+1})b_{-R-1} + \cdots + (C_n - C_{Sn})b_{-n}\right] \tag{B12}$$

or alternately as

$$-H = C_1b_{-1} + C_2b_{-2} + \cdots + C_Rb_{-R} + (C_{R+1} - C_{SR+1})b_{-R-1} + \cdots + (C_n - C_{Sn})b_{-n} + C_{SR}\left(\frac{1}{2}y_{-R} + \frac{1}{4}y_{-R-1} + \cdots +\right) \tag{B13}$$

Therefore

$$\leq |C_1 b_{-1}| + |C_2 b_{-2}| + \cdots + |C_R b_{-R}| + |(C_{R+1} - C_{SR+1}) b_{-R-1}| + \cdots + |(C_n - C_{Sn}) b_{-n}| + C_{SR}\left(\frac{1}{2} y_{-R} + \frac{1}{4} y_{-R-1} + \cdots\right) \qquad \text{(B14)}$$

The upper bound of $-H$ is attained when equality is attained in equation (B14). It can be demonstrated by an argument similar to that used in Theorem B.1 that the conditions for equality are

$$y_{-j} = C_j \qquad \text{for all } j$$

and

$$C_{Sj} = \overline{C}_j \qquad \text{for all } j > R$$

Therefore, equation (B14) may be written as

$$-H = (C_1, C_2, C_3 \cdots C_n) M \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ \cdot \\ \cdot \\ \cdot \\ C_n \end{pmatrix} - (0, 0, 0 \cdots 0, \overline{C}_{R+1}, \overline{C}_{R+2} \cdots \overline{C}_n) M \begin{pmatrix} C_1 \\ C_2 \\ \cdot \\ \cdot \\ \cdot \\ C_n \end{pmatrix} + C_{SR}\left(\frac{1}{2} C_R + \frac{1}{4} C_{R+1} + \cdots\right) \qquad \text{(B15)}$$

The last two terms in equation (B15) are always nonnegative. Moreover, the sum of these two terms is nondecreasing as $R$ decreases. Since by Theorem A.1, $C_1 = 1$ is a condition for maximizing the first term of equation (B15), it must also be a condition for maximizing equation (B15) itself.

Equation (B11) may be rewritten with this condition as follows:

$$-H = y_{-1} + (C_2 - C_{S2}, C_3 - C_{S3} \cdots C_n - C_{Sn}) M \begin{pmatrix} y_{-2} \\ \cdot \\ \cdot \\ \cdot \\ y_{-n} \end{pmatrix} \qquad \text{(B16)}$$

Therefore, the equation for $H_{min}$ can be immediately written as:

$$H_{min}(n) = -1 - G_{max}(n-1) = -\frac{10}{9} - \frac{n-1}{3} - \frac{(-1)^n}{9 \cdot 2^{n-1}} \qquad \text{(B17)}$$

Based on this analysis $C_{S1}$, $x_{S1}$, $C_1$, $x_1$, and $y_{-1}$ are equal to 1 for $H_{min}$. The remaining stages are determined to maximize $G$ for an $(n-1)$-stage

TABLE IX. - C, y, AND $C_S$ WHERE

$H_{min}$ IS ATTAINED

| n | $y_{-1} \cdots y_{-6}$ | $C_6 \cdots C_1$ | $C_{S6} \cdots C_{S1}$ |
|---|---|---|---|
| 2 | 11 | 11 | 01 |
| 3 | 101 | 101 | 011 |
| 4 | 1011 | 1101 | 0011 |
| 5 | 10101 | 10101 | 01011 |
| 6 | 101011 | 110101 | 001011 |

| n | $y_{-1} \cdots y_{-6}$ | $C_6 \cdots C_1$ | $C_{S6} \cdots C_{S1}$ |
|---|---|---|---|
| 2 | 11 | 11 | 01 |
| 3 | 111 | 111 | 001 |
| 4 | 1101 | 1011 | 0101 |
| 5 | 11011 | 11011 | 00101 |
| 6 | 110101 | 101011 | 010101 |

BRM. The values of y, C, and $C_S$, where $H_{min}$ is attained, are listed in table IX. The BRM counter value and the initial value for these $H_{min}$ values are the 2's complement of the preceding numbers. These values are tabulated in table VII.

## REFERENCES


1. Gordon, B. M.: Adapting Digital Techniques for Automatic Controls, I. Elec. Mfg., vol. 54, no. 5, Nov. 1954, pp. 136-143; 332.

2. Gordon, B. M.: Adapting Digital Techniques for Automatic Controls, II. Elec. Mfg., vol. 54, no. 6, Dec. 1954, pp. 120-125; 298; 300.

3. Mergler, H. W.: Digital Control Systems Engineering. Vols. I-II. Case Inst. of Tech., 1961.

4. Moshos, George J.: Design of Real Time Computers Utilizing Counting Techniques. NASA TN D-3042, 1965.

*"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."*

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

# NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Technical information generated in connection with a NASA contract or grant and released under NASA auspices.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

TECHNICAL REPRINTS: Information derived from NASA activities and initially published in the form of journal articles.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities but not necessarily reporting the results of individual NASA-programmed scientific efforts. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

***Details on the availability of these publications may be obtained from:***

SCIENTIFIC AND TECHNICAL INFORMATION DIVISION

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

**Washington, D.C. 20546**