



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 9      Issue: VI      Month of publication: June 2021**

**DOI: <https://doi.org/10.22214/ijraset.2021.35938>**

**[www.ijraset.com](http://www.ijraset.com)**

**Call:  08813907089**

**E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)**

# Processing of Digital Signals in (+1, -1) Binary System

P. Chaubey

Assistant Professor, Department of Physics, Govt. P.G. College, Tikamgarh (M.P.)

**Abstract:** Digital Signal Processing in (+1, -1) system has been studied in detail. Effects of number representation on quantization and truncation of rounding have been discussed. First order and second order filters have been investigated. Differential pulse code modulation method of coding is being studied. It is found that this system is well suited for VLSI design. The hardware realization of different components of DPCM is easy and straight. The position and negative signals can be processed in a unified way. It gives better performance in comparison of two's complement representation of conventional binary system. The effect of limit cycles, stability and round off noise in first and second order filter have been discussed too.

**Keywords:** (+1, -1) Binary system, Digital Signal Processing

## I. INTRODUCTION

The (+1, -1) number representation was proposed by Pekmestzi in 1989[1] for digital signal processing. On this representation low level of the signal is represented by -1, in place of zero in conventional binary. In digital signal processing the fractional signed numbers are required. Fixed point arithmetic numbers are employed. In fixed point arithmetic two's complement number suffers due to some drawbacks. To overcome the drawbacks, Pekmestzi pointed out that in digital signal processing, (+1, -1) binary system is very useful and suitable in VLSI circuit design. In digital signal processing, realization of digital filters of 1<sup>st</sup> and 2<sup>nd</sup> order either recursive or non recursive are important. In the design of digital filters finite word length including roundoff noise, coefficient quantization noise, limit cycles, and stability of filter design are essential quantization effects in digital filters have been extensively studied for different type of digital filters [2, 3]. Differential pulse code modulation is very useful in digitizing the signals. The use of DPCM in digital filter has been extensively studied by various workers [4, 5] in both recursive and non-recursive type. Linear predictive coding has been studied in (+1, -1) system [6].

In present paper we are mainly interested in digital signal processing using (+1,-1) binary system. Digital signal or sequences in this system is related to signum function which is related to unit step function. The quantization characteristic has been discussed. The paper contains rounding and truncation effects. The limit cycle effects and the effect on stability due to rounding and quantization on first and second order recursive digital filter has been studied. The results have been compared with two's complement representation of conventional binary system. Effect of prediction coefficient on mean square error is presented. Finally, differential pulse code modulation has been presented for digitizing signals.

## II. DIGITAL SEQUENCES

The unit sample sequence in (+1, -1) is defined by signum function

$$\text{Sgn}(n) = \begin{cases} +1 & n > 0 \\ -1 & n < 0 \end{cases}$$

The unit sample sequence is shown as in fig. (1a) and (1b).

A sequence of number x can be written as

$$x = \begin{cases} \{+x(n)\} & \text{for } n > 0 \\ \{-x(n)\} & \text{for } n < 0 \end{cases}$$

Where x(n) is the n<sup>th</sup> sample of sequence. The product and sum of sequence x and y are defined as

$$x, y = \{ \{ (\pm x(n) (\pm y(n))) \} \dots (3)$$

or x, y = {+x(n) y(n)} or {+x(n) - y(n)}

$$\text{or } \{-x(n) y(n)\} \text{ and or } \{-x(n) - y(n)\} \dots (4)$$

$$x + y = \{+x(n) + y(n), \text{ or } \{-x(n) + -y(n)\}$$

$$\text{or } \{+x(n) + - y(n) \text{ or } \{- x(n) + y(n)\} \quad \dots (5)$$

Multiplication of a sequence x by a number  $\alpha$  is defined as

$$\alpha \cdot x = \{\alpha (\pm x(n))\} \quad \dots (6)$$

Any arbitrary sequence in (+1, -1) system is defined by

$$x(n) = \sum_{k=-\infty}^{+\infty} x(k) \text{Sgn}(n-k) \quad \dots (7)$$

Where Sgn(n-k) is delayed unit sample.

The unit sample response in this system is defined by

$$h(n) = \begin{cases} a^n & \text{for } n > 0 \\ -a^n & \text{for } n < 0 \end{cases} \quad \dots (8)$$

or equivalently, by  $h(n) = a^n \text{Sgn}(n)$

The stability of linear time invariant system is computed by the sum.

$$S = \sum_{k=-\infty}^{+\infty} |h(k)| = \sum_{k=-\infty}^{+\infty} |\pm a|^k \quad \dots (9)$$

When  $|\pm a| < 1$ , the geometric series for S is given by

$$S = \frac{1}{1 - |\pm a|} \quad \dots (10)$$

The system is stable when  $|\pm a| < 1$ . Causal sequence is not possible in the system.

### A. Quantizer

Quantizer characteristic in (+ 1, -1) system is shown in Fig. (2). Here, b is the number of bits used for digital representation of sampled signals. The quantizer saturates output either at  $1-2^{-b}$  or  $-(1-2^{-b})$ .  $\Delta$  is the step size and is given by  $2^{-b+1}$ . This quantizer characteristic if compared with 2's complement quantizer characteristic, one finds that in 2's complement representation, b+1 bits left portion of decimal point represent the sign bit and right portion of decimal point represents the magnitude of the number but in (+1, -1) system sign bits are not required. The range of saturation is  $1-2^{-b+1}$  and -1 in conventional binary system and the step' size is  $2^{-b}$ .

Thus, the saturation points are not symmetrical in positive and negative side. The negative side contains more steps. At the same time the quantizer characteristic is symmetric in our system. The step size in our system is double that of two's complement quantizer. So, the number of steps are halved in our system.

Signal to noise ratio (SNR) for tlls system is given by  $(3 \times 2^{2b}) \sigma_f^2$ . Whereas in 2s complement, it is given by  $(12 \times 2^{2b}) \sigma_f^2$ . In (0, 1) system, if the sign of unquantized signal is unknown, the variance is given by  $2-211/3$ . For 1st order filter, noise variance due to

$$\text{quantization is given by } \sigma_f^2 = \frac{2^{-2b}}{3(1-c^2)} = \frac{2^{-2b}}{3 \times 2\epsilon} = \frac{2^{-2b}}{6\epsilon}$$

where, c is the amplitude.

If the pole is very close to the unit circle i.e.  $c = 1 - \epsilon$  where  $\epsilon \ll 1$

$$c^2 = (1-\epsilon)(1-\epsilon) = 1 + \epsilon^2 - 2\epsilon = 1 - \epsilon(2-\epsilon) \text{ so } c^2 \approx 2\epsilon, \text{ s is the distance of the pole from the unit circle.}$$

In conventional binary system of  $\sigma_f^2 = 2^{-2b}/1-c^2 \approx 2^{-2b}/24\epsilon$  and one extra bit for sign is required in two's complement representation.

In our system variance of quantization noise is  $2^{-2b}/3$  whereas in 2's complement representation, it is  $2^{-2b}/12$ .

### III. TRUNCATION AND ROUNDING:

The process of the truncation of numbers in this binary system can be explained by an example. Let us take fractional number (0.84375) represented by  $(.1\ 1\ 1\ \bar{1}\ 1)$  where complement of 1 is  $\bar{1}$  which is equal to -1. The rightmost bit in our example is 1 and assuming it to be an extra bit, it has to be truncated. After truncation the number becomes  $1\ 1\ 1\ \bar{1} = 13/16 = 0.8125$ . The error of truncation thereby is 1132. In case we have the last bit as 1 for a number say  $(.1\ 1\ 1\ 1\ \bar{1}) = 29/32 = 0.90625$  then on truncating, we obtain the number as  $30/32 = 0.9375$ . One observes that depending upon the number to be truncated, the value after truncation may be greater or smaller than the original number.

Next, we consider the process of rounding off a number in (+1, -1) representation. A number which is to be rounded, the last bit is removed and necessary quantity is added to the number. Using the process of rounding repeatedly one can round off any number of digits. In the case of (+1, -1) system, a full adder is necessary therefore we need three digits for addition. The rounding process can be achieved in the following manner.

Let us take an example: Let the number to be rounded off be  $(.1\ 1\ 1\ \bar{1}\ 1)$ . The value of this number is (0.84375). The third binary number we take as the complement of the last bit (whose value in our case is 1/32) and equals -1/32 which will be written as  $(\bar{1}\ 1\ 1\ 1\ 1)$ . Now we get on adding the three numbers

$$\begin{array}{r}
 \bar{1}\ 1\ 1\ 1\ 1 \quad \longrightarrow 3^{\text{rd}} \text{ number} \\
 \phantom{\bar{1}\ 1\ 1\ 1\ 1} \phantom{1} \quad \longrightarrow \text{Carry number} \\
 \hline
 .1\ 1\ 1\ \bar{1}\ 1 \quad \longrightarrow \text{Original number} \\
 \hline
 1.\ \bar{1}\ 1\ 1\ \bar{1}\ 1
 \end{array}$$

The result of the addition has an leftmost digit (1). The next to leftmost ( $\bar{1}$ ) is replaced by the bit (1) and one gets the sum as  $.1\ 1\ 1\ \bar{1}\ 1$  which 'is equal to 27/32. Since we have used an additional third number, we remove that value from the result and then we obtain  $27/32 - 1/32 = 26/32 = 13/16$ , which is equal to  $(.1\ 1\ 1\ \bar{1})$ . Thus, we find that truncation and rounding produce the same result.

In two's complement representation the truncation and rounding produces different results. The magnitude of the number after truncation is always less than the magnitude before truncation. In our system, the magnitude of the number after truncation is either smaller or greater than the number before truncation. The rounding and truncation produce the same result.

The truncation error in two's complement for positive and negative numbers are given by the range  $-2^{-b} \leq E_T \leq 0$  and  $0 \leq E_T \leq 2^{-b}$  respectively. The rounding error for positive as well as negative numbers is given by  $-\frac{1}{2}2^{-b} < E_R \leq \frac{1}{2}2^{-b}$

In our system the rounding or truncation error is given by the relation

$$-2^{-b} < E_R < 2^{-b}.$$

#### A. Limit Cycle Effects

The limit cycle behavior for first order filter can be realized as

$$p(n) == Q(n) + CP(n-1) \quad \dots\dots\dots (11)$$

Let us assume that C = 1/8 and the register length for storing the coefficient C, the input Q(n) and filter node variable P(n-1) are three bits respectively. Because of finite length registers, the product CP(n-1) is rounded or truncated to three bits before addition to Q(n). Let us consider the rounding or truncation of the product. The actual output R(n) satisfies the nonlinear difference equation.

$$R(n) = Q[CR.(n-1)] + Q(n) \quad \dots\dots\dots (12)$$

where Q [ ] represents the rounding or truncation operation.

Now C = 1/8 =  $.1\ \bar{1}\ \bar{1}$

The amplitude of unit sample = 7/8 =  $.1\ 1\ 1$

From equation (12) for n = 0

R(0) =  $.1\ 1\ 1$  and for n = 1

$$R(1) = C R(0) \text{ when the input i.e. } Q(n) = 0$$

$$= 7/64 = .1 \bar{1} \bar{1} \bar{1} \bar{1} \bar{1} \bar{1}$$

After truncation  $R(1) = .1 \bar{1} \bar{1} = 1/8$

Now, for  $n = 2$

$$R(2) = C R(1) = 1/64 = .1 \bar{1} \bar{1} \bar{1} \bar{1} \bar{1} \bar{1}$$

After truncation

$$R(2) = .1 \bar{1} \bar{1} = 1/8$$

Thus for  $n \geq 2$ , all the values of R are the same and equal to 1/8

For  $C = -1/8$  in similar fashion,

we get

$$R(1) = -1/8$$

$$R(2) = +1/8$$

$$R(3) = -1/8$$

Thus, periodic steady state solution between +1/8 and -1/8 is obtained. for  $C = -1/8$

Following the method of Jackson [14] the condition for dead bands is given by

$$|R(n-1)| \leq \frac{2^{-b}}{1-|C|} \quad \dots\dots(13)$$

For  $|C| = 1/8$ ,  $R(n-1) \leq 8/7 \times 2^{-b} \approx 2^{-b}$

Thus  $|C| = 1/8$  gives the correct value of dead band.

In two's complement representation, the register length is limited to four bits. The method of rounding determines the limit cycles and requires lengthy process. For  $C = 1/2$ , the correct value for the dead band is achieved. Thus, we find that limit cycle calculation in (+1, -1) representation is easier in comparison of two's complement representation. In (+1,-1) system, the dead bands are achieved using truncation and less number of bits and less number of n in comparison of conventional binary system.

For second order filter, the non linear difference equation is given by

$$R(n) = Q(n) + Q[A Q(n-1)] + Q[B Q(n-2)] - Q[CR(n-1)] - Q[DR(n-2)] \quad \dots\dots (14)$$

In case of limit cycle

$$R(n) = -Q[CR(n-1)] - Q[DR(n-2)] \quad \dots(15)$$

Using truncation relation

$$|Q [DR(n-2)] - DR (n-2)| \leq 2^{-b} \quad \dots(16)$$

Since  $Q[DR(n-2)] = R(n-2)$

$$\text{So } |R(n-2) - DR (n-2)| \leq 2^{-b}$$

$$\text{i.e. } |R(n-2)| \leq \frac{2^{-b}}{1-|D|} \quad \dots(17)$$

The value of D should be such that poles of the system are on the unit circle.

**B. Effect of Quantization Error on Stability**

Step size between quantized level is given by

$$\Delta = 2^{-b+1}$$

Where b is the number of bits for coefficient C.

For stable filter  $-1 < C < 1$

If  $\epsilon = 1 - C$  is the distance of the pole to the unit circle.

The smallest value of  $\epsilon = 2^{-b+1}$  when it is either truncated or rounded.

$$\text{Now } \log_{10} \epsilon = \log_{10}(2^{-b+1}) = - (b-1) \log_{10}2 \quad \dots\dots(18)$$

$$\frac{\log_{10} \varepsilon}{\log_{10} 2} = -(b-1) \quad \dots(19)$$

$$\text{So, } b = \frac{\log_{10}(1-C)}{\log_{10} 2} + 1 \quad \dots(20)$$

In case of two's complement representation in (0, 1) system for truncation.

$$b = -\frac{\log_{10}(1-C)}{\log_{10} 2} - 2 \text{ and for rounding, } b = -\frac{\log_{10}(1-C)}{\log_{10} 2} - 1 \quad \dots(21)$$

For second order filter, the difference eqn. is given by

$$Y_k = A_1 y_{k-1} + A_2 y_{k-2} + B_1 x_k + B_2 x_{k-1} \quad \dots(22)$$

Let the poles be  $p_1$  and  $p_2$ .  $A_1$ ,  $A_2$ ,  $B_1$  and  $B_2$  are coefficients with respect to pole position.

Sensitivity of coefficient is given by the relation.

$$\frac{\partial p_1}{\partial A_1} = \frac{p_1}{p_1 - p_2} \quad \frac{\partial p_2}{\partial A_1} = \frac{p_2}{p_2 - p_1}, \quad \frac{\partial p_2}{\partial A_2} = \frac{1}{p_2 - p_1}$$

$$\text{and } \frac{\partial p_1}{\partial A_2} = \frac{1}{p_1 - p_2} \quad \dots(23)$$

Total variation in pole position is given by

$$\Delta p_1 = \frac{1}{p_1 - p_2} (p_1 \Delta A_1 + \Delta A_2) \quad \dots(24)$$

and

$$\Delta p_2 = \frac{1}{p_2 - p_1} (p_2 \Delta A_1 + \Delta A_2) \quad \dots(25)$$

For stabilities  $-2 < A_1 < 2$  and  $-1 < A_2 < 1$ .

For truncation or rounding,

$$\Delta A_1 = \Delta_{1/2} = \frac{2^{-b+2}}{2} = 2^{-b+1} \text{ where } \Delta_1 = 2^{-b+2} \quad \dots(26)$$

$$\Delta A_2 = \Delta_{2/4} = \frac{2^{-b+1}}{2} = 2^{-b+1} \text{ where } \Delta_2 = 2^{-b+1} \quad \dots(27)$$

From (24), (25), (26) and (27) and also using

$p_1 = 0.98$  and  $p_2 = 0.94$ , we get

$$\begin{aligned} \Delta p_1 &= \frac{1}{0.04} (0.98 \times 2 + 0.5) 2^{-b} \\ &= \frac{1}{0.04} (1.96 + 0.5) 2^{-b} \\ &= \frac{2.46}{0.04} \times 2^{-b} \quad \dots(28) \end{aligned}$$

$$\text{or, } \Delta p_1 = 1 - p_1 = 0.02 = \frac{2.46}{0.04} \times 2^{-b} \quad \dots(29)$$

$$\text{so, } 2^{-b} = \frac{0.04 \times 0.02}{2.46} \quad \dots(30)$$

$$2^{-b} = \frac{2.46}{0.04 \times 0.02} = \frac{2.46}{0.0008} = \frac{24600}{8} = 3075 \quad \dots(31)$$

b = 12 bits for the minimum register length

In case of 2's complement number for rounding, we get

$$\Delta p_1 = \frac{1}{0.04} (0.98 + 0.5) 2^{-b} \quad \dots(32)$$

$$0.02 = \frac{1.48}{0.04} \times 2^{-b}$$

$$2^b = \frac{1.48}{0.0008} = \frac{14800}{8} = 1850 \quad \dots(33)$$

b = 11 bits for the minimum register length.

and for truncation, we get

$$\Delta A_1 = \frac{1}{0.04} (2 \times 0.98 \times 1.0) 2^{-b} \quad \dots(34)$$

$$0.02 = 74 \times 2^{-b} \text{ or } 2^b = 3700 \quad \dots(35)$$

b = 12 bits for minimum register length.

Thus, minimum register length in both the systems are approximately the same.

#### IV. MULTIPLIER

Let us take two fractional numbers A and B. In (+1, -1) system these numbers can be written as

$$A = \sum_{k=1}^m (a_k) 2^{-k}$$

$$B = \sum_{l=1}^n (b_l) 2^{-l}$$

where  $a_k, b_l = +1$  or  $-1$  and m and n are the number of bits. To clarify the multiplication process, we take an example.

Let  $A = (.1 \bar{1} \bar{1}) = 1/8$

and  $B = (. \bar{1} \bar{1} 1) = -5/8$

then  $A \times B = -5/64 = 0. \bar{1} 111 \bar{1} 1$ . The multiplication process can be understood from figure (3).

In figure (3), the value of  $C_3 C_2 C_1$  is ten as  $1 \bar{1} \bar{1}$  which is general will be  $C_n C_{n-1} \dots C_2 C_1$  as  $1 \bar{1} \dots \bar{1} \bar{1}$ . The values of  $S_3 S_2 S_1$  is complement of  $C_3 C_2 C_1$  and hence it is  $\bar{1} 1 1$  and in general, it will be  $S_n S_{n-1} \dots S_2 S_1$  as  $\bar{1} 1 \dots 1.1$ .

The input sum ( $S_3 S_2 S_1$ ) and input carry ( $C_3 C_2 C_1$ ) are introduced in the partial product because in this system only full addition is possible. Introduction of input sums and input carries do not affect the result because they are of complementary nature.

$C_1$  and  $S_1$  are added to the partial product marked as a rectangle to produce a sum S and a carry C in the figure (3). Since S has nothing else to be added, it is carried towards final total sum as  $T_1$ . The generated carry C and the input sum  $S_2$  are added to the partial product  $\bar{1}$  to give a generated carry and a generated sum. These generated carry and sum will be subsequently added to partial products. The last generated carry is however, taken as the generated sum for the diagonally adjacent partial product as shown by arrow.

A. The Prediction Filter

The first order prediction' filter is given by the expression

$$P(n) = Q(n) + C.P. (n-1) \dots\dots(36)$$

where Q(n) is real data and C.P. (n-1) is real data with additional bits.

Let us take an example

$$\text{Let } Q(n) = -3/16 = 0.\bar{1} 1 1 \bar{1}$$

$$\text{and } C P(n-1) = +59/128 = 0.1\bar{1} 1 1 \bar{1}\bar{1}$$

substituting these values in equation (19), we get

$$\begin{array}{rcl} Q(n) & = & 0.\bar{1} 1 1 \bar{1} \\ CP(n-1) & = & 0.1 \bar{1} 1 1 \bar{1} \bar{1} \\ \text{Sum} & = & 1 \bar{1} \bar{1} 1 \bar{1} \bar{1} \bar{1} 1 \end{array}$$

$$\text{So, } p(n) = 0.1 \bar{1} 1 \bar{1} = +5/16$$

$$\text{The rounding error} = \bar{1} \bar{1} 1 = -5/128$$

In CP(n-1) the bits to be truncated are  $1 \bar{1} 1 = 3/128$ . The MSB of truncated value is taken as a carry bit for the sum and the addition is done. In the result, if the carry bit and the MSB are opposite, the MSB is replaced by carry bit and the carry is ignored. This is due to the fact that when two fractional numbers are added, an additional bit is always generated. If the result is in the same range as the two numbers, the last two bits  $S_0$  and  $S_1$  will be of opposite sign and the sum is corrected by replacing  $S_0$  by  $S_1$  because of the relation

$$(S_0) + (S_1)2^{-1} = (S_0) 2^{-1} \quad \text{when } (S_0 = - (S_1)).$$

The carry bit, used for the sum does not produce any error in the result which can be explained as,

The truncated value  $1 \bar{1} 1 = 3/128$  can be written as  $1/16 - 5/128$ , which implies the addition of  $1/16$  as carry in the addition does not change the value of the rounding error.

Next, we consider the second order prediction filter. The second order filter is given by the expression

$$p(n) = Q(n) + AQ(n-1) + BQ(n-2) - CP(n-1) - DP(n-2) \dots\dots (37)$$

We take an example to illustrate the above filter,

$$\begin{array}{rcl} \text{Let } Q(n) & = & 0.1 \bar{1} \bar{1} \bar{1} = 1/16 \\ AQ(-1) & = & 0.1 \bar{1} \bar{1} 1 \quad 1 \bar{1} \bar{1} = 3/16 + 1/128 \\ BQ(n-1) & = & 0.1 \bar{1} 1 \bar{1} \quad 1 \bar{1} = 5/16 + 1/64 \\ CP(n-2) & = & 0.\bar{1} 1 1 1 \quad \bar{1} 1 \bar{1} = -1/16 - 3/128 \\ DP(n-2) & = & 0.\bar{1} 1 1 \bar{1} \quad \bar{1} 1 = -3/16 - 1/64 \end{array}$$

Using the same method as done in first order filter, we obtain



$$p(n) = 0.1 \bar{1} \bar{1} 1 = 5/16$$

and the truncated value = - 1/64 which equals the round off error.

In a similar way the third order filter can be written as

$$p(n) = Q(n) + AQ(n-1) + BQ(n-2) + CQ(n-3) - DP(n-1) - EP(n-1) - FP(n-3) \dots\dots\dots(38)$$

Higher order filters can be similarly generated.

Next, we discuss the evaluation of the prediction coefficients by using the Lavinson and Durbin algorithm.

The autocorrelation function for sample ( $x_n$ ) is given by the relation

$$\phi(n) = (1/N) \sum_{i=1}^{N-1} x_i x_{i+1} \text{ for } n = 0,1,2, \dots\dots\dots p \dots\dots\dots(39)$$

For minimum mean square error (MSE) with respect to the prediction coefficient ( $a_i$ ), we have the relation for the set of linear equations

$$\sum_{i=1}^p a_i \phi(i-j) = \phi(j) \text{ for } j = 1,2,3, \dots\dots\dots p \dots\dots\dots(40)$$

Equation (23) can be written in matrix form as

$$\phi_a = \phi \dots\dots\dots(41)$$

Where  $\phi$  is a  $p \times p$  matrix with elements

$\phi_{ij} = \phi(i-j)$ ,  $a$  is a  $(p \times 1)$  column vector of prediction coefficients and  $\phi$  is a  $(p \times 1)$  column vector with elements  $\phi(i)$ , for  $i = 0,1,2,3, \dots\dots\dots p$ .

The matrix  $\phi$  has equal diagonal elements.

In this system the samples  $\{x_n\}$  are always odd fraction numbers. For a 5 bit quantizer the input to the quantizer will be 2/32, 4/32, 6/32 ..... 30/32. Thus, there will be 15 samples. The values of the prediction coefficients from equation (8). The values of the prediction coefficients and the results are shown in figure (4).

From the figure (4), we observe that prediction coefficients sharply changes from a maximum value  $a_{11} = 0.4$  to the minimum value  $a_{22} = - 0.167$  when the order of the filter is changed from 1 to 2. However for  $p = 2$  to 9 the values of filter coefficients increase exponentially and finally for  $p = 9$  to 12, the filter coefficients show some variations. We find that the changes in the values of prediction coefficients with order of prediction filter show a damped oscillatory behavior.

The mean square error for first order filter is given by .

$$e_1 = \phi(0) (1 - a_{11}^2) \dots\dots\dots(42)$$

for  $m^{\text{th}}$  order filter, mean square error is given by

$$e_m = e_{m-1} (1 - a_{mm}^2) \text{ and}$$

$m = 2, 3, \dots\dots\dots p \dots\dots\dots < a_{mm}$  is the prediction coefficients of  $m^{\text{th}}$  order filter. From equation (22), we can easily compute the values of MSE for different order of filters. Variation of MSE with  $p$  has been shown in figure (5). The values of mean square error decreases with increase of  $p$ . Firstly, it decreases rapidly and then slowly and finally it becomes almost constant. Thus the necessary and sufficient condition  $\{(e_m \leq e_{m-1} \leq \dots\dots\dots e_1) \text{ and } (|a_{mm}| < 1)\}$  is satisfied in this system.

**V. DIFFERENTIAL PULSE CODE MODULATION (DPCM)**

The current sample for the source can be obtained by weighted linear combination of past samples. Hence, we can write

$$y_q(n) = \sum_{i=1}^r a_i y_{n-i} \tag{43}$$

where  $y_q(n)$  is the predicted value of  $y$ , which is the current sample.  $r$  is the number of samples and  $a_i$  represents the coefficients of the past samples. The equation (26) can be rewritten as

$$y_q(n) = a_1 y_{n-1} + a_2 y_{n-2} + a_3 y_{n-3} + \dots \tag{44}$$

$a_1 y_{n-1}$  represents first order prediction filter,  $a_2 y_{n-2}$  represents second order filter and so on. The prediction filter of different orders have been discussed earlier.

In order to explain the concept of the DPCM we take an example. Let the four samples be  $-2/16, +2/16, +4/16$  and  $+6/16$ . The quantized values of these are  $-3/16, +3/16, +5/16$ , and  $+7/16$  respectively. The predictor output of these samples is given by using equation (44).

$$y_q(n) = 3/16 + (-3/16) + 5/16 + 7/16 = 13/16 = 1 \ 1 \ 1 \ \bar{1}$$

whose truncated value =  $-3/128$ .

The current value of the incoming signal  $y_n = +8/16$ . The difference of the input to the quantizer and the predicted value is given by

$$e_n = y_q(n) - y_n + q_e(n) \tag{45}$$

where  $q_e(n)$  is the quantization error.

Substituting the values in the above equation (45), we get  $e_n = 6/16$  where  $q_e(n) = 1/16$ .

The quantized value of error is given by the relation

$$\begin{aligned} y'_n &= y_q - e_q(n) + q_e(n) \tag{46} \\ &= 13/16 - 7/16 + 1/16 \\ &= 7/16 \end{aligned}$$

If we add the quantization noise to the prediction filter input, we obtain the current signal. Thus

$$y'_n + q_r(n) = y_n = 8/16 \tag{47}$$

Hence one finds that the quantization error IS independent of the prediction filter employed.

The PCM transmitter and receiver are shown in figure 6(a) and 6(b) respectively.

The average power of the message sequence of length  $N$  is given by

$$P_m = (1/N) \sum_{n=0}^{N-1} y_n^2 \tag{48}$$

The average power of the quantizing error of message sequence is however given by

$$P_q = (1/N) \sum_{n=0}^{N-1} q_r^2(n) \tag{49}$$

and the average power of the error sequence is given by

$$P_e = (1/N) \sum_{n=0}^{N-1} e_n^2 \tag{50}$$

The output signal to quantization noise ratio

$$(SNR)_0 = P_m / P_q \quad \dots (51)$$

The signal to quantizing noise ratio

$$(SNR)_q = P_e / P_q \quad \dots (52)$$

and the prediction gain is given by

$$G_p = P_m / e \quad \dots (53)$$

The values of  $(SNR)_0$  and  $G_p$ , for different values of  $N$  are shown in table 1.

The results have been compared with (0, 1) system [15].

### VI. CONCLUSIONS

Signal can be processed in efficient way in (+ 1, -1) binary system. Digital signals can be represented by using signum function. Sequences are evenly spaced on time scale in positive as well as negative direction. The amplitudes of sequences are positive as well as negative. Sign is inherent in this system. The quantizer is symmetric and step size is twice. Signal to noise ratio is small. Rounding and truncation are equivalent. Therefore, the limit cycle calculation becomes very simple. Dead bands are achieved in simple way and also rapidly. The number of bits to represent the coefficients for stabilities is less than two's complement representation for 1st order filter.

The proposed system is well suited for differential pulse code modulation. Each component of DPCM can be constructed in modular form. Positive and negative signals can be handled in a unified way. The internal symmetry of the system is reflected from the result that the truncation and rounding errors are equal. The internal symmetry of system results in a highly modular circuits. The unified circuit design leads to reduction in circuit complexity and delay time. The modularity of the circuit is yet another important factor for VLSI design [15, 16]. The addition and multiplication for positive and negative signals can be performed with the help of a single basic cell. An array of binary adders is also possible. Carry-save type, the cellular array multipliers and the iterative multipliers can be constructed with the help of similar cells. The prediction gain in this system is very high and hence the average power of error sequence is small for a given base band signal with fixed average power of message sequence. Thus, a prediction filter with minimum error power can be designed. The values of prediction coefficients are positive as well as negative in this system. The prediction filter is very sensitive upto third order and the predicted value decreases very sharply. The predicted values increase after third order. It first increases rapidly up to 6th order and then increases slowly up to 10th order. This type of behavior of the filter is expected due to uniform quantization of positive and negative signals and ordering realization of filter. The behavior of filter after third order shows resemblance with conventional binary but the predicted values are different. Output signal to noise ratio increases with prediction filter coefficients.

Quantization noise ratio is found to be 0.25 for different number of samples. It is independent of filter coefficients. Thus, we conclude that the (+1, -1) is suitable for VLSI design and also for fast communication whenever positive and negative signals are transmitted simultaneously.

Table 1.  $(SNR)_0$  and  $G_p$  for different values of  $P$ .

P	$(SNR)_0$ (db)	$G_p$ (db)	$(SNR)_0$ (db) <sup>[14]</sup>
1	0.87	5.23	5.4
3	4.98	10.22	8.4
5	8.43	13.66	8.6
7	10.83	16.77	10.5



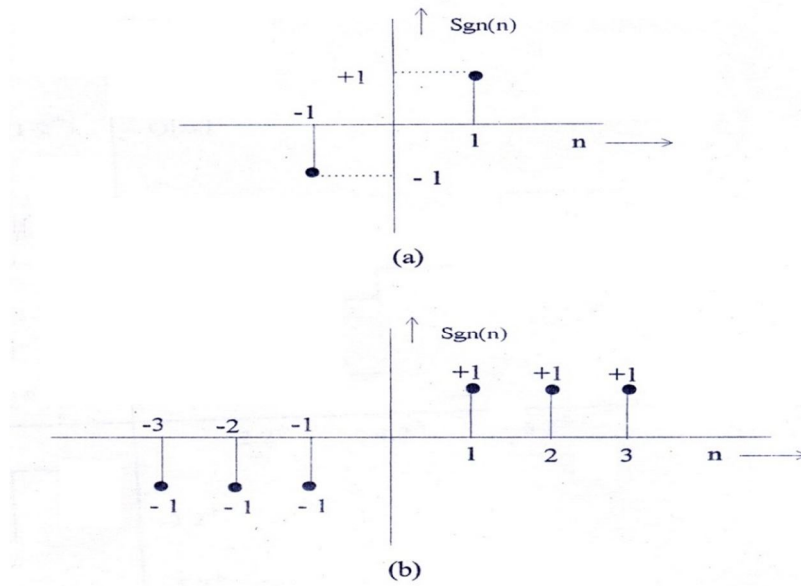


Fig. (1)

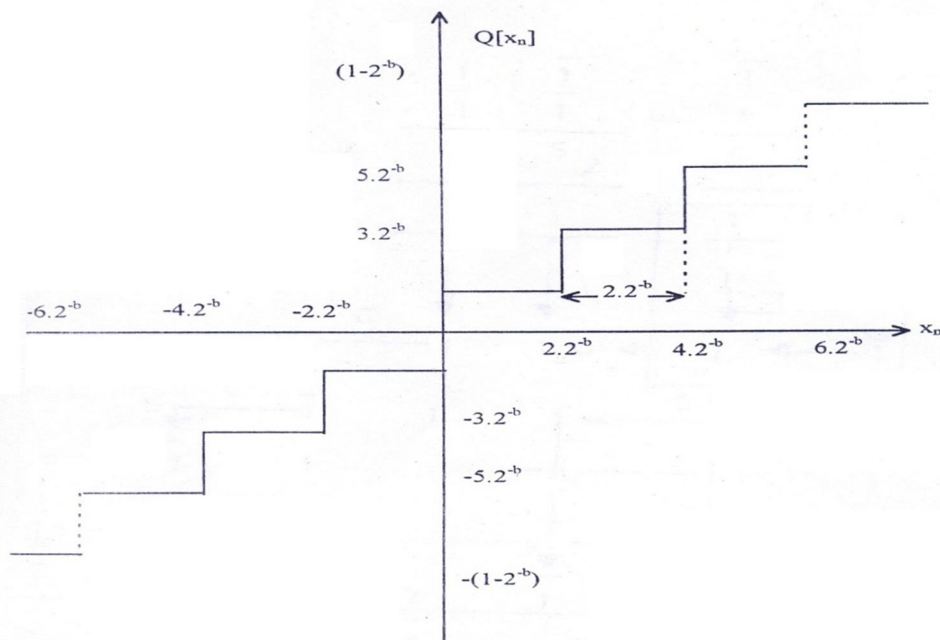
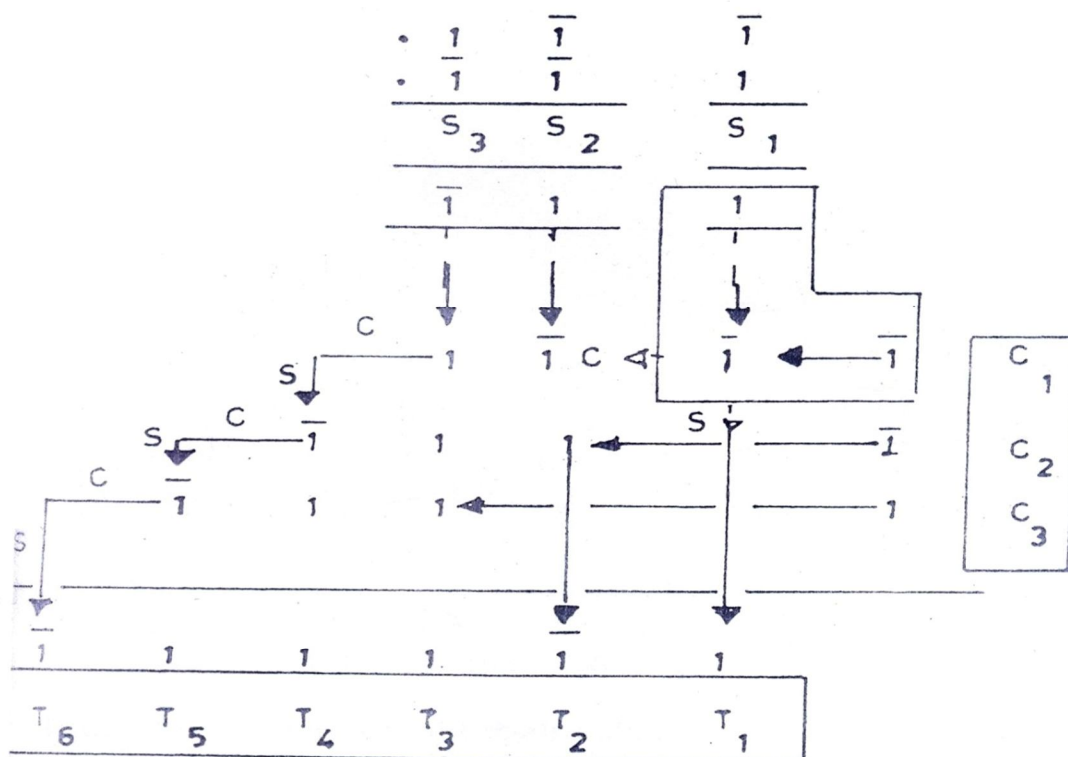
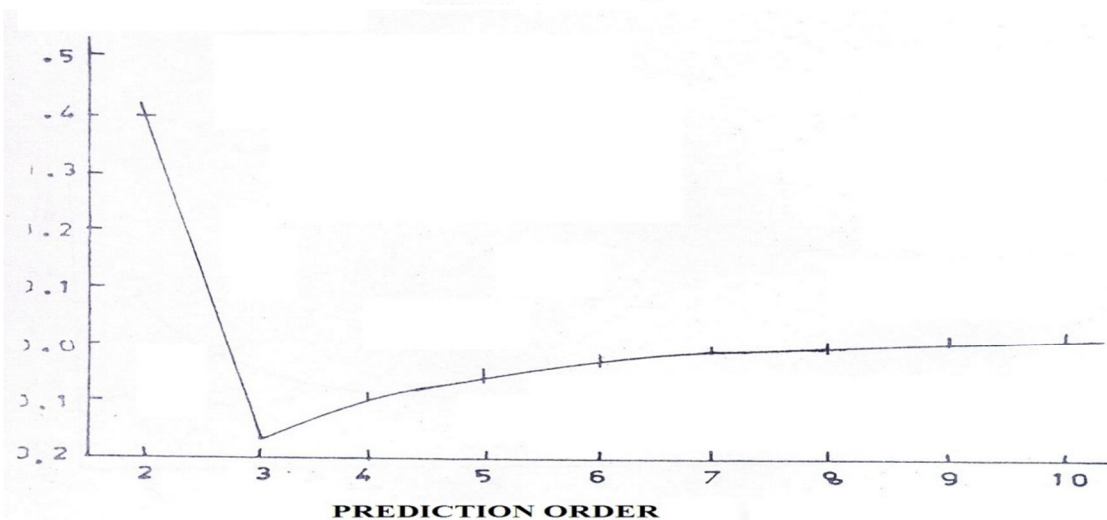


Fig. (2)



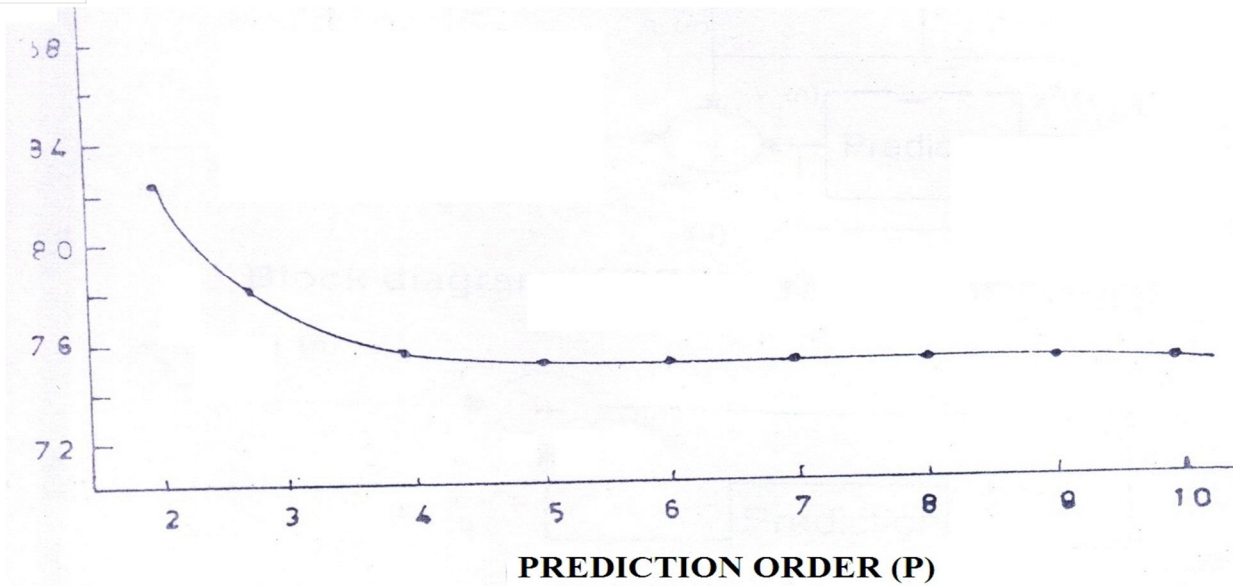
Three-bit multiplier

Fig. (3)



Prediction order against their values

Fig. (4)



Variation of MSE with orders of filter

Fig. (5)

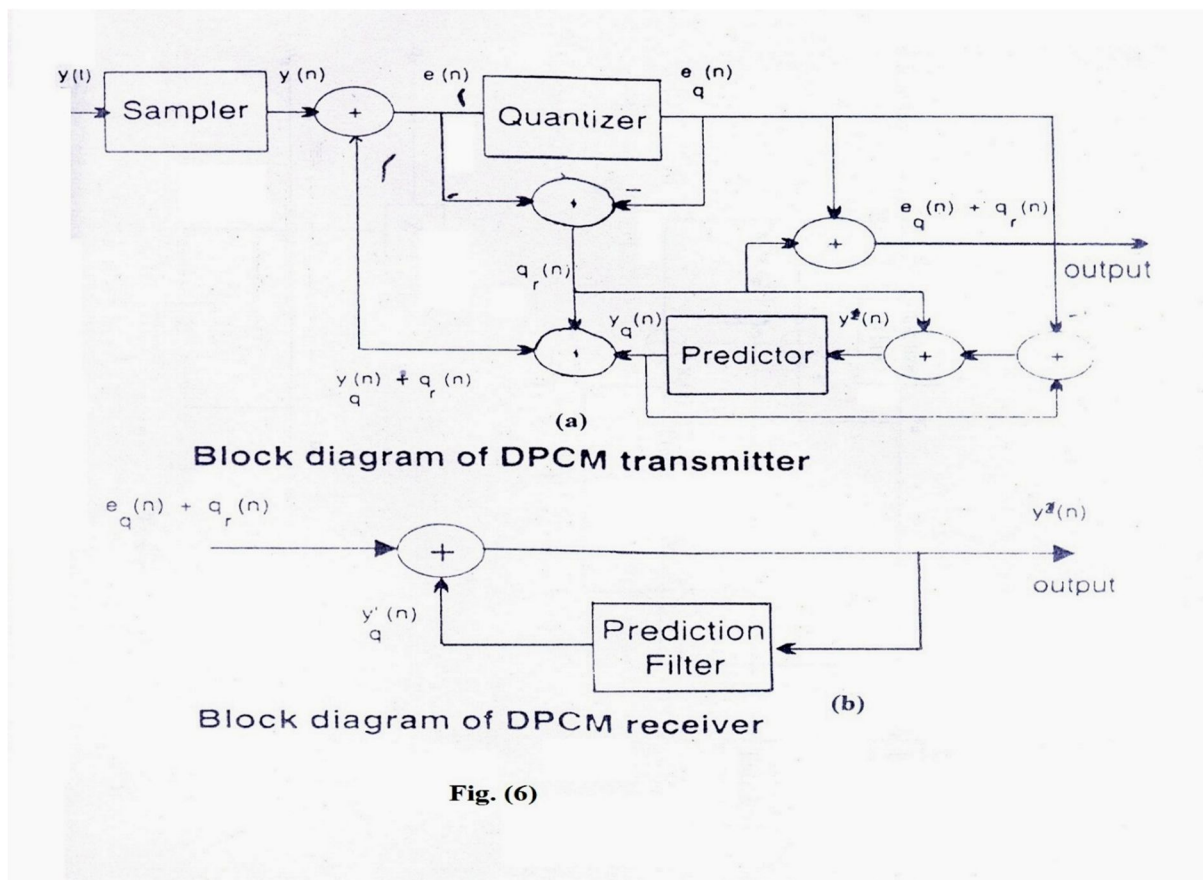


Fig. (6)

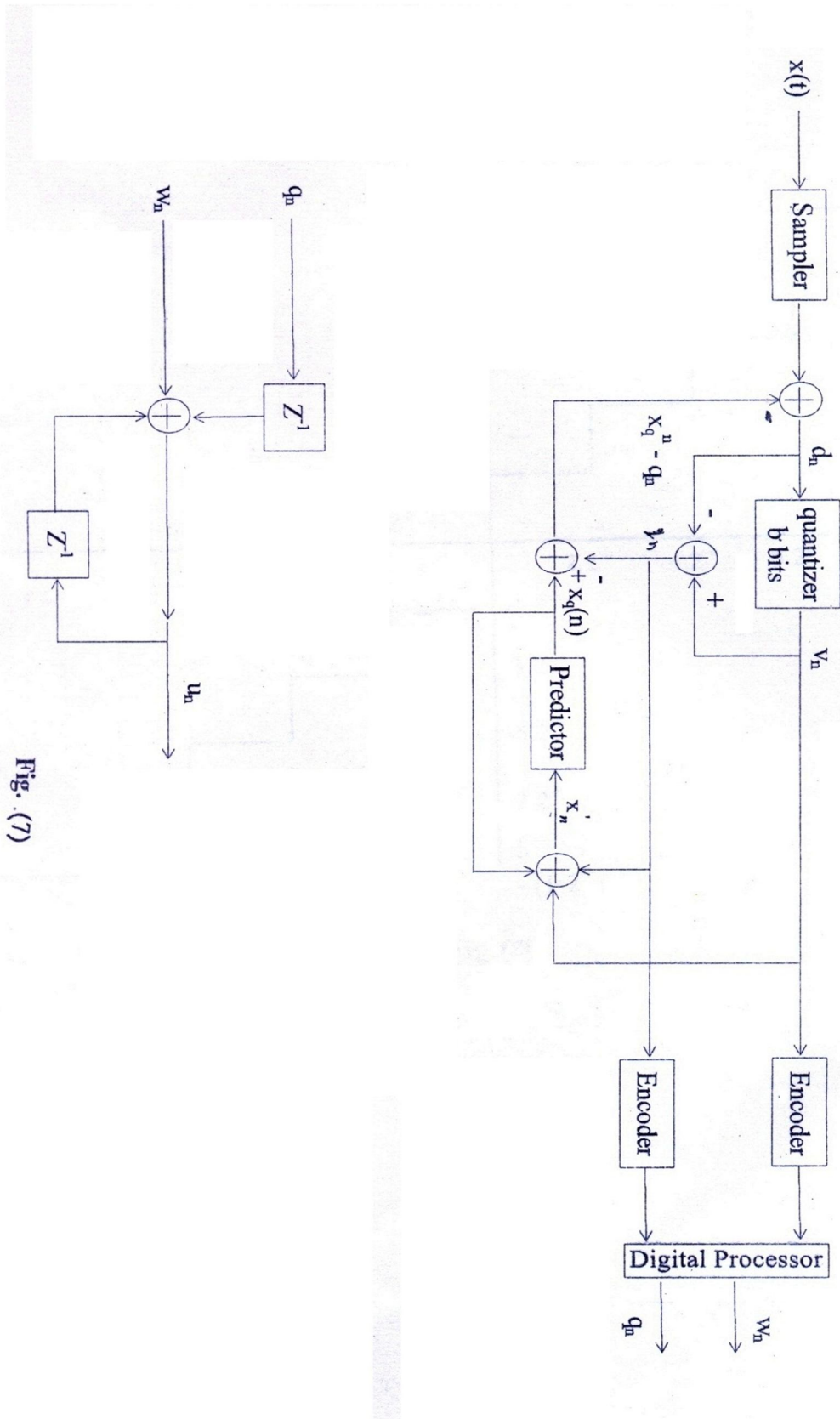
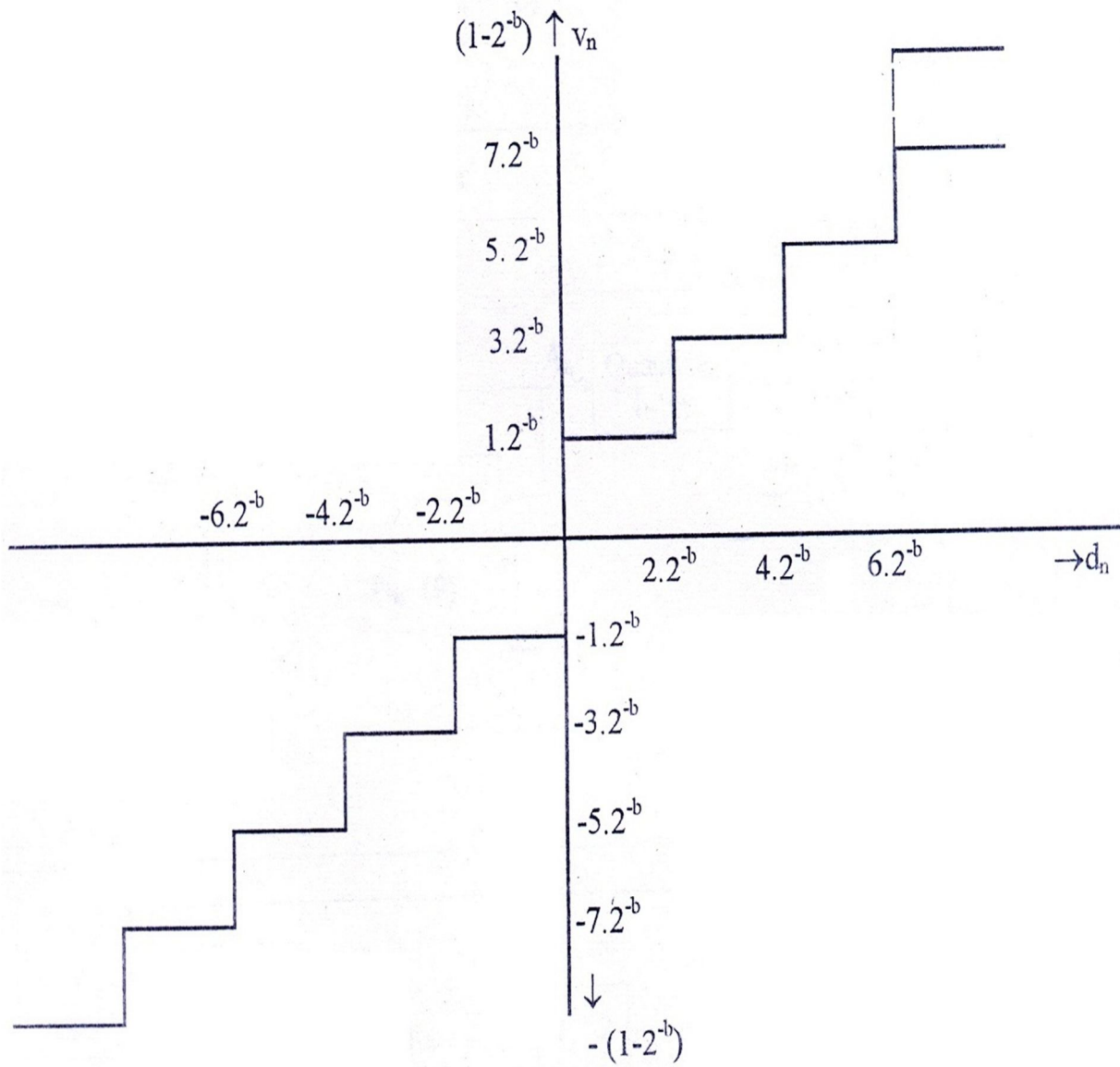


Fig. (7)





**Fig. (8)**

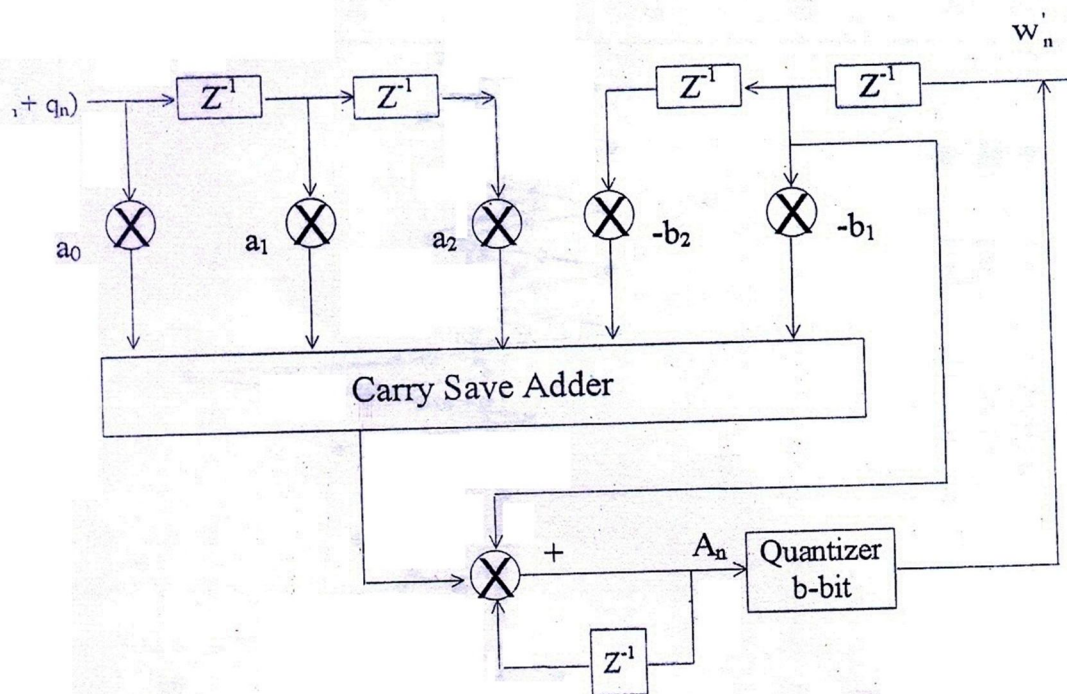


Fig. (9)

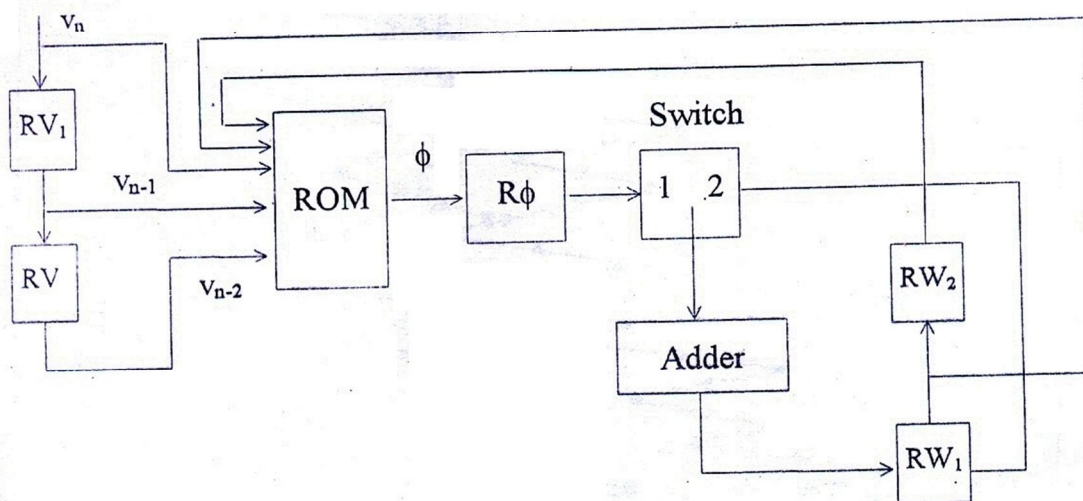


Fig. (9)

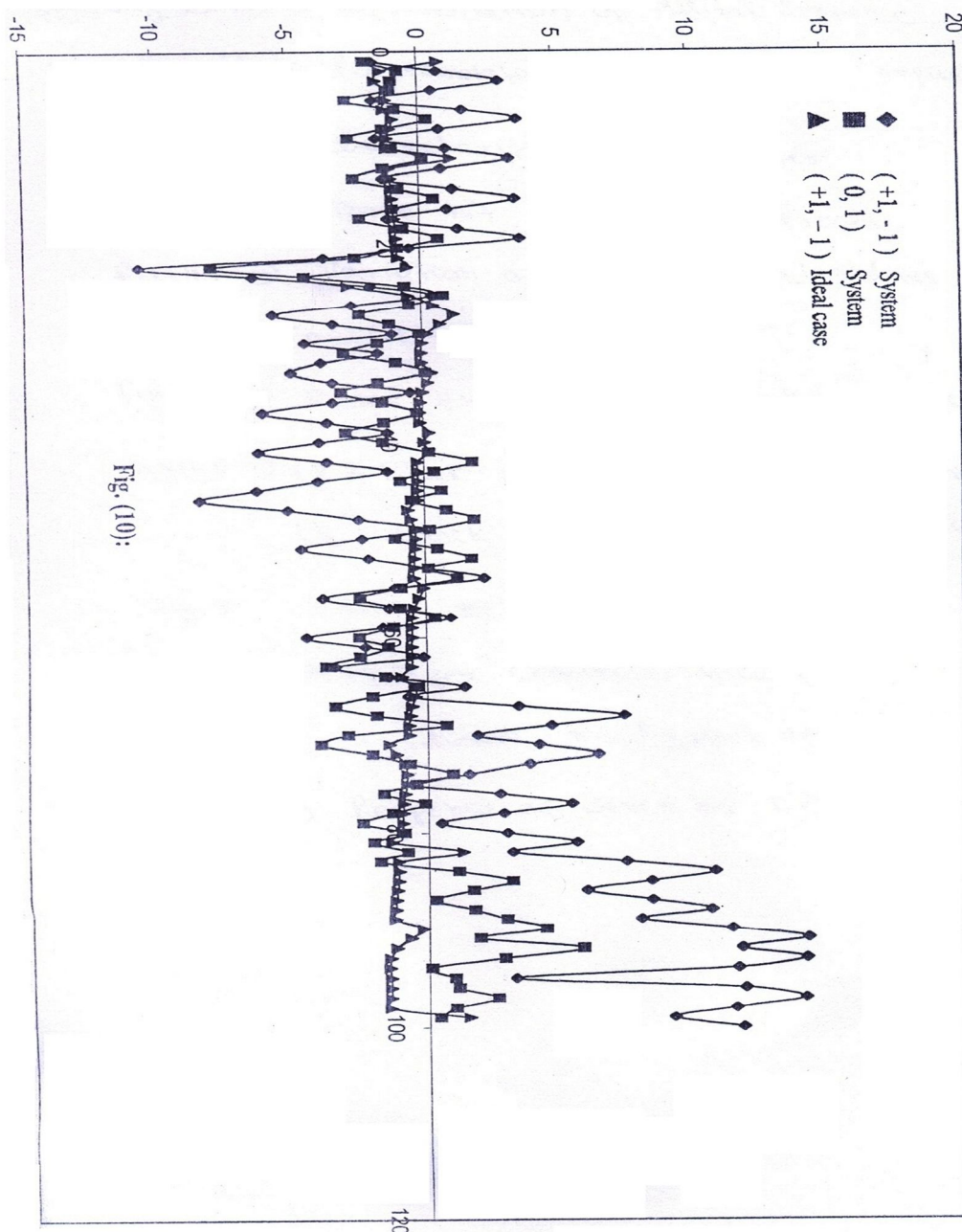


Fig. (10):

A. Captions to Figures

- 1) Fig. (1) (a) Representation of digital sequence  
(b) Representation of digital step sequence
- 2) Fig. (2) Characteristics of quantizer
- 3) Fig. (3) Three bit Multiplication Process
- 4) Fig. (4) Variation of the Predicted Values with Prediction Coefficients
- 5) Fig. (5) Mean square Error and the orders of filter Coefficients
- 6) Fig. (6) (a) Block diagram of DPCM Transmitter  
(b) Block diagram of DPCM receiver
- 7) Fig. (7) DPCM digital Processor System
- 8) Fig. (8) Quantizer Characteristics
- 9) Fig. (9) Hardware realization of Second order DPCM Filter
- 10) Fig. (10) Response of Seven bit DPCM High Pass Filter

## REFERENCES

- [1] K.Z. Pekmestzi, "New number representation of digital signal processing", *Int. Journal of Electronics*, Vol. 66, pp. 709-23, 1989.
- [2] B.-H. Juang and A.H. Gray Jr., "Multiple stage vector quantization for speech coding", in *Proc. ICASSP*, 1982, pp. 597-600.
- [3] J.D. Markel and A.H. Gray Jr., *Linear Prediction of Speech*. New York: Springer-Verlag, 1976.
- [4] B.W. Bomar, "New second order state space structures of realizing low round off noise digital filters", *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol. ASSP-33, pp. 106-10, 1985.
- [5] M. Kawamata and T. Higuchi, A limited approach to the optimal synthesis of fixed point state space digital filters, *IEEE Trans. Acoust, Speech, Signal processing* Vol. ASSP-33 (1985) pp. 911-920.
- [6] L.B. Jackson, Analysis of Limit cycles, due to multiplication Rounding in Recursive Digital Filters. *Proc. 7<sup>th</sup> Annual Allerton Conf. Circuit System Theory*, 1969, 66-79.
- [7] E. Avenhaus, "On the design of Digital filters with coefficient of Limited work Length. *IEEE Trans. Audioelectro acouties*, AU-20, 1972, 206-212.
- [8] T.A.C.M. Classen, W.F.G. Meeklenbrauker and J.B.H. Peek, Quantisation. Noise analysers for fixed Prime Digital Filters using Magnitude transaction for Quantization, *IEEE. Trans. Circur and systems*, CAS-22, 1975, 887-895.
- [9] R.H. Chang and W.G. Bleis, "Register modeling in the unification of quantization analysis in the HR digital filters, *Proceeding a silomare Conf.*, Oct. 1992, pp. 278-82.
- [10] M. Iwatsuke; M. Kaunmata and T. Higuchi, "Statistical sensitivity and minimum sensitivity structure with fewer coefficients in discrete time linear system" *IEEE Trans. Circuit System*, Vol. 37, pp. 72-80, 1990.
- [11] Tran-Thong and B. Liu, "A recursive digital filter using DPCM", *IEEE Trans. Commun.*, Vol. 24, pp 2-11, 1974
- [12] L. Goldstein and B. Liu, "An ADPCM realization of no recursive digital filter" *IEEE Trans. Acoustics, Speech and Signal Processing*", Vol. 24, pp 312-20, 1975.
- [13] *Principles of Digital Communications*, Das, J. Mullick, S.J. and Chatterjee, P.K. 1986, pp. 183.
- [14] K.Z. Pekmestzi and G.D. Papadopoulos, "A two's complement cellular assay multiplier", *The Radio and Electronic Engineer*, 51, pp. 94-96, 1981.
- [15] D.J. Goodman and L.J. Greenstei "Quantizing noise of  $\Delta M/PCM$  encoder" *Bell System. Tech. Journal*, Vol. 52, pp. 183-204, 1973.
- [16] N.S. Jayant, "Digital coding of speech waveforms, PCM, DPCM and DM quantizer, *Proc. IEEE*, 62, pp. 611, 1974.



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)