VLB ARRAY MEMO No. 427

THE 8/5 CODE FOR THREE-LEVEL QUANTIZATION: TAPE FORMAT CONSIDERATIONS

Larry R. D'Addario
16 January 1985

INTRODUCTION

A fairly efficient encoding of three-level quantized noise, from the point of view of minimizing storage requirements for achieving a given SNR, is to store each group of 5 samples in an 8-bit word. The details of this are given in VLBA Memo No. 332. To achieve the same SNR (in continuum) with 2-level quantization would require about 1% more storage, and with 4-level it would require about 8% more. Moreover, from the point of view of conserving bandwidth, observations achieving the same SNR will require 62% more signal bandwidth with 2-level compared to 3-level (this assumes rectangular bandpasses and Nyquist rate sampling).

Thus it seems clear that, for continuum observations limited by storage, 3-level quantization with 8/5 encoding is the system of choice. For observations which are not storage-limited, the finest quantization consistent with reasonable system complexity is desired. Since 4-level may lead to too much correlator complexity for the VLBA, 3-level is again the choice; the SNR per unit time would be 21% worse with 2-level.

The question remains whether the 8/5 code can be efficiently utilized with the tape storage system planned for the VLBA. First, is the equipment required for encoding and decoding excessively complex? Second, can the tape system capacity be well matched to the encoded data rates at VLBA sampling rates?

ENCODERS AND DECODERS

Figure 1 is a logic diagram of a digitizer and encoder. They will work up to sample rates of 16 Ms/sec, and the cost of the encoder is just 3 MSI chips and a small ROM. (There is some extra cost in timing generators, but this can be neglected since these would be shared by many encoders .) A decoder could be constructed very similarly.

The incremental cost cannot be more than a few hundred dollars per station, compared with using a 2-bit/sample code.

TAPE AND SAMPLING PARAMETERS

In what follows, I shall assume that the tape recording equipment is subject to the following limitations:

1. Linear density - 33,000 bits per inch maximum
2. Tape velocity - 270 inches/sec maximum
3. Number of simultaneous tracks - 32, plus 3 spares (or possibly 33 or 34 with fewer spares)

4. Number of passes at different head positions - 32 maximum

This implies a maximum tape bandwidth of 32 x 270 ips x 33000 bpi = 288 Mb/s.

The VLBA will use sampling rates of $2^{-k}*16E6$ samples/sec, for k = 0 to 7. After 8/5 coding, the bit rates are 1.6 times the sampling rates, or $2^{-k}*25.6E6$ bits/sec. The number of channels of such data which will fit within the tape bandwidth is thus $11.25*2^k$; or

f_s = 16 Ms/s => 11.25 channels
8 => 22.5
.
.
.
.125 => 1440.

Here, and throughout this memo, I have neglected the effect of overhead bits, including headers, synchronization words, and error checking bits; the number of these is a function of the track format, which is a separate issue. Thus, when exact results are required, most of the rates and speeds given here should be increased to account for the overhead bits.

For maximum utilization of the tape, we would like to operate always at the maximum available linear bit density, and we would like to always use all available tracks. To achieve the latter, note that it is acceptable to use only an integal fraction of the tracks at one time and to use the others on successive passes; but it is not feasible to use different numbers of tracks on different passes. On the other hand, tape utilization is not impaired by running at less than the maximum possible tape velocity. The latter is required only if the maximum total bandwidth must also be achieved.

POSSIBLE DATA ORGANIZATIONS

The data rate written onto a track need not be the same as the data rate of a channel, since multiplexing and demultiplexing of the channel data is possible. To keep the implementation as simple as possible, I consider only schemes which mux or demux the data by small integral ratios. Table I then illustrates some modes that result in a constant track data rate. The mux/demux ratios are all powers of two. To achieve a linear bit density of 33,000 bpi at 6.4 Mb/s requires a tape velocity of 194 in/sec. At sample rates of 2 Ms/s or less, one could use more tracks per pass and correspondingly less multiplexing and lower tape velocities to maintain the same density. For the highest three sampling rates, the total Nyquist-rate bandwidth is 64 MHz and the total data rate to the tape is 204.8 Mb/s.

TABLE I: Constant-Speed Modes
(Rates in units of 10^6/sec)

Sample Rate	Encoded Date Rate	Mux/ Demux	No. of Channels	No. of Tracks	Track Data Rate
16	25.6	1/4	8	32	6.4
8	12.8	1/2	16	32	6.4
4	6.4	1/1	32	32	6.4
2	3.2	2/1	32	16	6.4
1	1.6	4/1	32	8	6.4
0.5	0.8	8/1	32	4	6.4

Implementing this set of modes has some advantages. Since the maximum tape velocity is never used at record time, some speedup at playback time is always possible. Additional speedup is possible at the lower sampling rates. The sensitivity achieved at the maximum total bandwidth of 64 MHz is 27% better than the same bandwidth with 2-level quantization, at a tape consumption rate that is between those of 1 bit/sample and 2 bit/sample codes (135 < 194 < 270 ips). However, it does not utilize the maximum possible bandwidth of the recorders, which would require running at the maximum tape velocity.

For those experiments demanding the maximum possible sensitivity, we must ask how much of the recorder bandwidth is actually usable. This may be less than the recorder's capacity because of constraints on tape changing times and tape shipping costs. If the full tape bandwidth is used (270 ips, 32 passes) then for 12 hours operation the tape reel must be 30,400 feet long. This tape length is very unlikely to be feasible. It seems clear that the 24-hour unattended operation specification cannot be met at the full tape bandwidth without three or more transports per station, which is expensive. Even then, the tape shipping costs may be prohibitive. So something less than the full bandwidth must be used; this means running at something less than maximum tape velocity, if maximum tape utilization (bit density) is to be maintained.

Using the modes of Table I, the maximum bandwidth (at 194 ips) would achieve 12 hours of operation with 22,500 feet of tape. Thus, whether this is a reasonable set of modes depends on whether such a tape length is achievable and on whether the shipping costs would then be acceptable.

Having considered these examples, we can now see the general result: we start with the amount of tape that will fit on a transport, and this implies (given the other parameters) the usable tape bandwidth; from the above discussion, we expect it to be less than the maximum tape bandwidth. This also implies a particular tape velocity. We must then format our data so that it just fits onto the tape at this rate. Some examples are given in Table II. It should be noted that if bit densities higher than 33,000 bpi are achieved, then the data rates (but not the velocities) are increased proportionately.

TABLE II: Track Rates vs. Tape Length
(For 12 hour operation on one reel with 32 passes)

Length of Reel	Maximum Velocity	Data Rate for 33,000 bpi at max velocity
9,600 ft	85.33 ips	2.82 Mb/s
18,200	161.8	5.34
22,500	200	6.60

I have previously advocated, in private notes and discussions, a very general formatter implementation that would be capable of producing data streams for all available tracks at nearly any specified track rate from an appropriate number of channels at any convenient channel rate. This would allow full utilization of the usable tape bandwidth, and would be easily adaptable to improvements in tape technology or to adoption of other technologies in the future. However, the consensus seems to be that the extra complexity of digital hardware that this would require is prohibitive (probable cost is a few K$ per station, plus extra design time and possibly reduced reliability). Therefore, the remainder of this memo considers a simplified system in which each channel's data stream is mulitplexed and/or demultiplexed by small integers in producing the track data streams. This turns out to have sufficient flexibility for most purposes.

Consider the block diagram of Figure 2. Here each of the N_{ch} channel data streams at rate f_d is first multiplexed by M, so that M of them fit onto one line; then each of these is demultiplexed by D, so that each is split into D lines. The output data rate is then $f_t = (M/D)f_d$ for each of the $N_t = (D/M)N_{ch}$ tracks. Several reasonable cases are considered in Table III. The choice of cases is based on the assumption that the tape velocity and the tape length may be chosen freely, up to the maximum limits (i.e., any value is available to the designer; once the system is built, only a small number of descrete values will be available to the user). Thus, the tape velocities of Table III were chosen to give f_d/f_t ratios which are small integers, and the required tape lengths were then calculated.

TABLE III: Data Configurations For 8/5 Code

f_d Mb/s	11.2kft,97ips,3.2Mb/s				16.4kft,145ips,4.8Mb/s				22.5kft,194ips,6.4Mb/s			
	M	D	N_{ch}	N_t	M	D	N_{ch}	N_t	M	D	N_{ch}	N_t
25.6	1	8	4	32	3	16	6	32	1	4	8	32
12.8	1	4	8	32	3	8	12	32	1	2	16	32
6.4	1	2	16	32	3	4	24	32	1	1	32	32
3.2	1	1	32	32	3	2	24	16	2	1	32	16
1.6	2	1	32	16	3	1	24	8	4	1	32	8
0.8	4	1	32	8	6	1	24	4	8	1	32	4

The third set of columns of Table III is essentially the same as Table I. All cases are based on fixed tape velocities, but one could always increase D or decrease M by any desired tape speedup factor. The middle columns illustrate the effect of having a tape length which implies a somewhat odd M/D ratio. In this example, all 32 tape tracks are efficiently utilized but the natural number of channels is 6 times a power of 2; I have shown a maximum of 24, although it could have been 48, compared to the usually-assumed value of 32 for the VLBA. If this is inconvenient, other choices of M,D lead to N_{ch} as large as 33. Note, however, that any inconvenience is due not to the the "oddness" of f_d, but rather to an inconvenient tape length combined with the choice of 32 for the number of simultaneous tracks. To illustrate this, Table IV shows how data rates which are powers of 2 times 16 MHz (appropriate to 1 or 2 bit/sample codes) could be used with a 18,200 foot tape.

TABLE IV: 16 MHz Data Rate Example

f_d	18.2kft,162ips,5.33Mb/s			
Mb/s	M	D	N_{ch}	N_t
16	1	3	10	30
8	2	3	20	30
4	4	3	40	30
2	8	3	40	15
1	16	3	32	6
.5	32	3	32	6

OTHER ADVANTAGES OF THE 8/5 CODE

Of the 256 possible codes using 8 bits, only $3^5 = 243$ are required to represent 5 samples. Some of the extra 13 codes can be put to good use. The all-zero and all-one codes need not be used at all; this limits the maximum number of consecutive zeros or ones in valid data to 14, which limits the bandwidth over which the playback signals must be equalized. Furthermore, the codes which represent efficient sync patterns can be reserved for that purpose and not used to represent valid data; in this way, sync patterns will not occur in valid data, and this allows use of a simple design for the sync detectors. (For example, an optimum 32-bit sync pattern is $69969669 in hexadecimal; this can be produced by reserving just two of the spare codes, $69 and $96.)

SUMMARY AND RECOMMENDATIONS

It has been assumed that the requirement for 12 hours operation per tape reel (or 24 hours, if possible) is a hard specification; that the number of tracks per pass is fixed at about 32; and that the maximum tape velocity is fixed at 270 ips. The linear bit density and the size of a tape reel may be subject to some improvement in the future. It has been shown that these parameters determine the usable tape bandwidth, and that this will require an average tape velocity less than the maximum. The tape velocity then determines the track data rate, which implies a multiplexing ratio for putting the available channel data rates onto the tracks. By suitable choice of the tape velocity and slight adjustment of the tape length (downward from the maximum), the multiplexing factors can be kept to ratios of small integers.

One could figure out a reasonable set of velocities and multiplexing factors based on what we now know about the tape technology, but this may change and we may wish to support other technologies. It costs very little to allow a wide range of multiplexing factors; I suggest providing for programmable values of M and D as follows:

$M = 1, 2, 3, 4, 6,$ or 8;
$D = 1, 2, 3, 4, 5, 6, 8,$ or 16.

This also allows a small number of record velocities to be implemented without constraining the speedup factor required at playback time.

The actual record velocities implemented should be determined by the achievable tape length, as described above. However, there should be one additional velocity near the maximum of 270 ips, regardless of whether it differs from the others by a power of two, and regardless of the tape consumption rate. This will allow the maximum tape bandwidth to be available for short periods of time if needed. This would allow, for example, 11 channels of 25.6 Mb/s data to be recorded (using $M = 1$, $D = 3$, and $N_t = 33$). This represents 88 MHz bandwidth with Nyquist sampling and 8/5 3-level encoding, compared with 64 MHz with a 2-bit/sample code. Of course, one could also record 128 MHz of bandwidth with a 1-bit/sample code, but this would give about 2% less sensitivity.

The above recommendations apply regardless of the quantization or encoding chosen; they come only from considering the parameters of the tape technology.

I also recommend implementation of the 8/5 code for 3-level data, considering that (1) the cost of the encoders and decoders is negligible; (2) it is the optimum practical code for tape-limited observations; and (3) much less RF bandwidth is consumed in tape-limited observations than with 2-level quantization.

Having done this, there seems to be no need at all for 2-level quantization; however, maintaining 2-level support requires no significant additional hardware and it could be kept. But it does introduce complexity and some inefficiency, so it probably should be dropped in the absence of a strong argument to the contrary. Whether 4-level quantization should be supported in the recorders depends on whether the correlator can use it; if not, there is no point in including it in the recorders. Supporting all three quantizations, including 8/5 encoding of 3-level data, is possible at reasonable hardware cost; but the overall system simplification that results from supporting only a single code may be significant. Since 3-level is optimum and should be supported, it is the addition of 2-level and/or 4-level codes which should be thought of as extra-cost options.

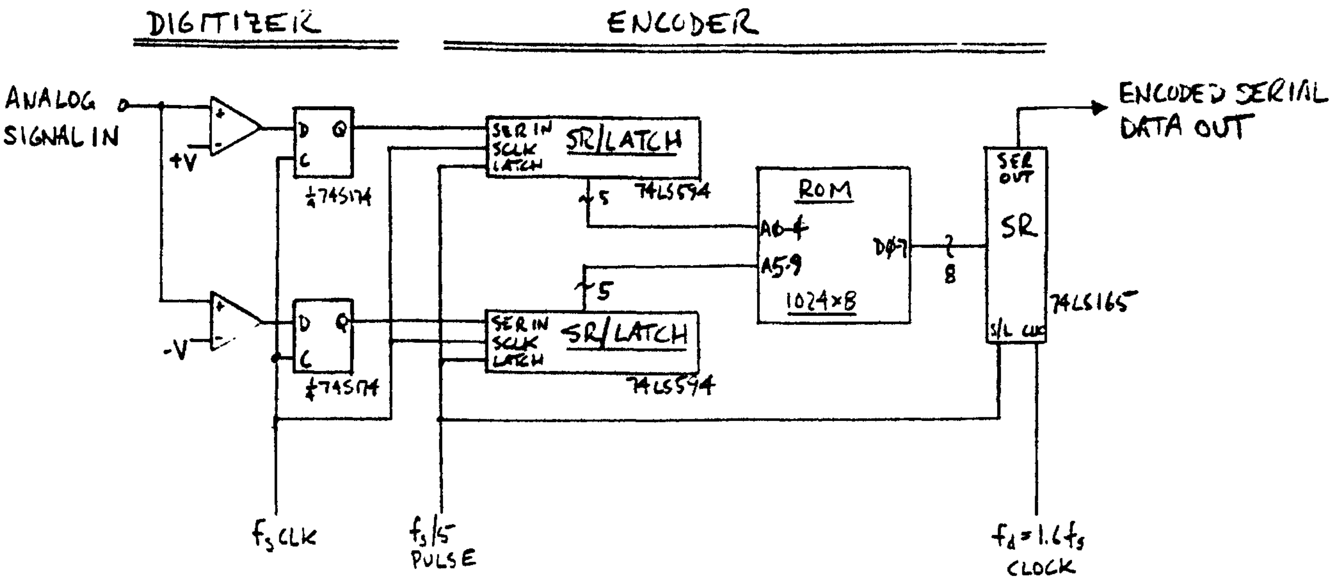

FIGURE 1: DIGITIZER AND ENCODER CIRCUITS

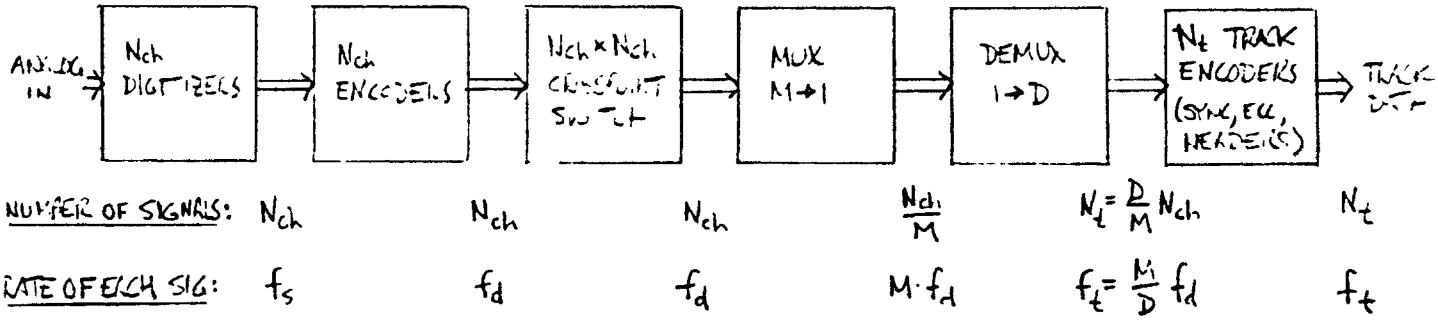

FIGURE 2: RECORDING SYSTEM BLOCK DIAGRAM