THIS DOCUMENT HAS BEEN CONVERTED FROM THE ORIGINAL PUBLICATION:

Meiners, E. P., Lenz, L. L., Dalby, A. E. and Hornsby, J. M., 1972, Recommended standards for digital tape formats: Geophysics, 37, no. 01, 36-44.

RECOMMENDED STANDARDS FOR DIGITAL TAPE FORMATS[1]

E. P. MEINERS[2], L. L. LENZ[3], A. E. DALBY[4], AND J. M. HORNSBY[5]

This paper is the result of work done by a subcommittee of the SEG Committee on Digital Recording Standards. It recommends an additional tape format for 9-track field recording of seismic data and an optional use of 1600 bpi, as well as 800 bpi, for all field formats.

INTRODUCTION

During early 1966, it became apparent that unless something was done, the number and variety of 9-track tape formats used in field recording would become unwieldy. What was needed was an industry-wide agreement to limit the number to a preferred few. To help resolve this problem, the SEG set up a Committee on Digital Recording Standards. This committee, under the chairmanship of E. J. Northwood, culminated its activities in a report (Northwood et al, 1967) which recommended two field tape formats, labeled Format A and Format B. These formats, which have been generally accepted, applied primarily to binary-gain recording. Standardization efforts with regard to 21-track formats were not considered necessary, as it was felt that the danger of proliferation was not great in this area.

During early 1971, it became apparent that owing to a number of instrument developments it was desirable again to consider the matter of standards. Among these developments were the following: (1) Instantaneous floating point amplifiers, (2) 1600 bpi recording density, and (3) possible need for larger data words to meet the requirements of field compositing. A number of proposals were made by instrument manufacturers for new formats. These were subsequently reviewed by the present reconstituted committee. The following recommendations by the Subcommittee on Field Tape Formats is the result of considerations of these various proposals.

RECOMMENDATION

The subcommittee recommends the adoption of an IBM compatible, four bytes per sample, full-word floating point format to be designated Format C. The SEG Format A cannot accommodate the new, true floating point recording systems; and it was the sense of the subcommittee that the industry would best be served if *a single* new format could be adopted. It is further recommended that both 800 bpi and 1600 bpi be acceptable for Formats A, B, and C.

RECORDING SPECIFICATIONS

The following is a description of general recording specifications for Format C. Reference should be made to Figures 1 and 2.

1. Either the NRZI method at 800 bpi density or the phase encoded (PE) method at 1600 bpi may be used for recording. (Due to tape speed limitations of present field transports, Format C will probably be limited to 1 msec sampling of 30 channels, 2 msec sampling of 62 channels, or 4 msec sampling of 126 channels for 1600 bpi, and 2 msec sampling

---

Manuscript received by the Editor August 25, 1971.

[2] Chairman, Committee on Digital Recording Standards (Chevron Oil Company, Geophysical Division, Houston, Texas 77036).

[3] Chairman, Subcommittee on Field Tape Formats (Esso Production Research Company, Houston, Texas 77001).

[4] Member, Subcommittee on Field Tape Formats (The Superior Oil Company, Houston, Texas 77001).

[5] Member, Subcommittee on Field Tape Formats (Western Geophysical Company of America, Houston, Texas 77001).

of 30 channels or 4 msec sampling of 62 channels for 800 bpi.)

2. The seismic record file is divided into a header block, containing identification information pertaining to the seismic record, and a data block, containing data values of the seismic channels. The first header block on a reel begins at least 3.0 inches past the trailing edge of the loadpoint reflective marker. For NRZI (800 bpi) recording, tape is erased from at least 1.7 inches before the trailing edge of the loadpoint marker to the beginning of the first header block. For PE recording, a PE identification burst is written beginning at least 1.7 inches before the trailing edge of the loadpoint marker and continuing past the trailing edge of the marker. The identification burst consists of 1600 flux reversals per inch in track P with all other tracks dc erased. The preamble of the first header block is separated from the identification burst by at least 0.5 inch of erased tape.

3. The insertion of a standard IBM inter-record gap between the header and seismic data block is optional.

4. The seismic data are to be written in gapless mode.

5. When recorded by the phase encoded method (1600 bpi), the header block, if it is separated by an inter-record gap, must be preceded by a preamble, consisting of 40 bytes of all zeros and one byte of all ones, and followed by a postamble, consisting of one byte of all ones and 40 bytes of all zeros. The seismic data block preamble may be written immediately after the minimum inter-record gap following the header block. Zero data (bit P a one and bits 0-7 all zero) are then written until time break and the first start-of-scan is written. The zero data section *must* contain an integer multiple of four bytes. The seismic data block will be followed by a postamble and end-of-file. If there is no inter-record gap between the header and the seismic data blocks, the preamble will precede the header block, and zero data (bit P a one and bits 0-7 all zero) are then written until time break and the first start-of-scan is written. Again the zero data section *must* contain an integer multiple of four bytes. The postamble and end-of-file will follow the seismic data block.

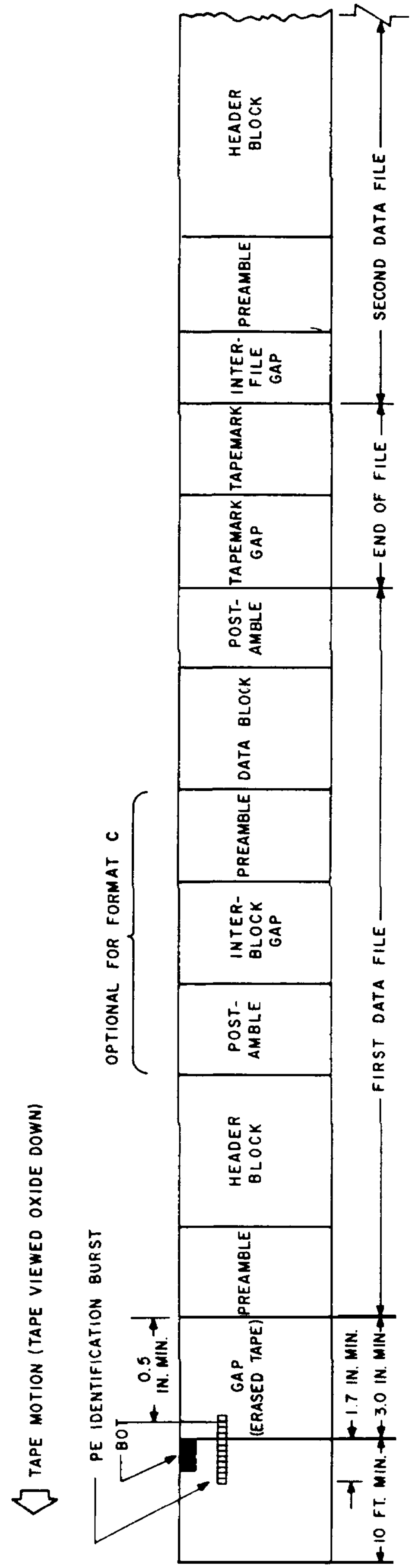

FIG. 1. General recording specifications for Format C 1600 BPI phase encoded recording. Note: Order and location of tracks on tape, direction of magnetization, and all other applicable specifications in accordance with IBM Form GA 22-6862.

6. When recorded by the NRZI method (800 bpi), the header block, if it is separated from the data block by an inter-record gap, must be followed by the redundancy checks (CRC and LRC) before the inter-record gap. If the entire seismic record file is written in a gap-less mode (no inter-record gap between header and seismic data), the redundancy checks and end-of-file will follow the seismic data block. Zero data (bit P a one and bits 0-7 all zero) are then written until time break, and the first start-of-scan is written. The zero data section *must* contain an integer multiple of four bytes. The seismic data block is followed by redundancy checks, inter-record gap, and finally an end-of-file.
7. Data values are written in eight bit bytes with vertical (byte) parity odd.
8. An end-of-file mark is recorded after each seismic record file with *at least* two end-of-file marks after the last seismic record file on the reel. The end-of-file mark consists of the 800 bpi NRZI tapemark or the 1600 bpi tape-mark character, as appropriate, preceded by a gap of at least 0.5 inch of erased tape.
9. Tape specifications, location and content of redundancy check characters, track dimensions and numbering, and all other applicable specifications shall be in accordance with IBM Form GA 22-6862 entitled "IBM 2400-Series Magnetic Tape Units Original Equipment Manufacturers' Information."

DESCRIPTION OF HEADER BLOCK

The header block for SEG Format C is shown in Figure 3. It has 24 bytes of standard information identical to the first 24 bytes of the header for SEG Format A and Format B, with one exception: the LSD (bits 4-7) of byte 24 is assigned a gain constant which is common to all seismic channels. This may avoid the necessity of recording the initial gain on an individual channel basis if the initial gain is the same for all seismic channels. Data are recorded in packed BCD. Each byte accommodates two binary coded decimal (BCD) digits. The most significant digit (MSD) position is in bits 0, 1, 2, and 3, with respective decimal values of 8, 4, 2, and 1. The least significant (LSD) position is in bit numbers 4, 5, 6, and 7 with respective decimal values of 8, 4, 2, and 1.

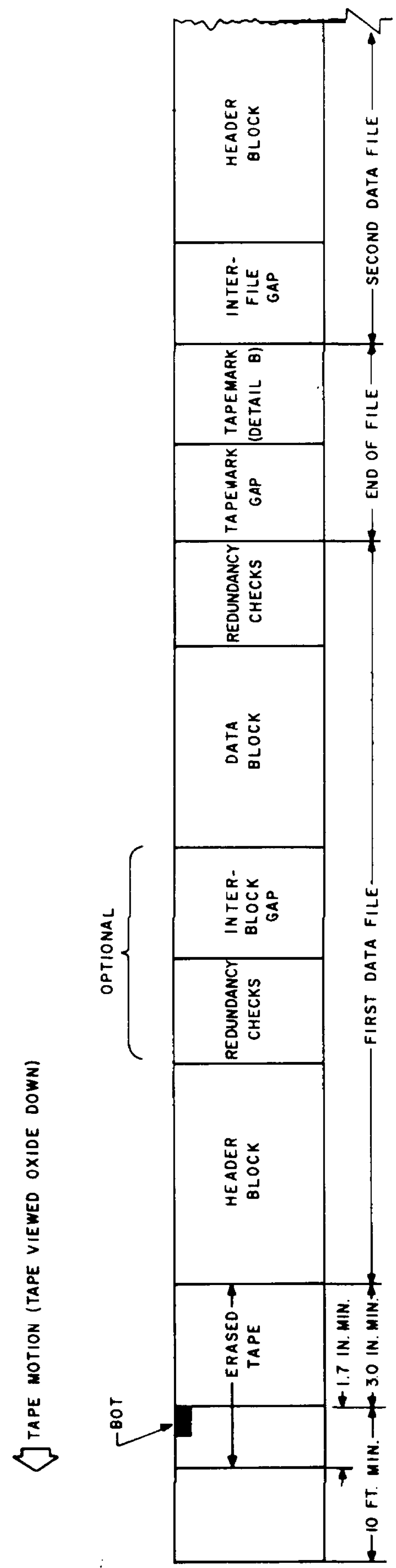

FIG. 2. General recording specifications for Format C 800 BPI NRZI recording. Note: Order and location of tracks on tape, direction of magnetization, and all other applicable specifications in accordance with IBM Form GA 22-6862.

Numbered subscripts on code letters in the format drawings refer to decimal digit position; e.g., Y_1 = high order position of format code and Y_4=low order position of format code. Format code number 0273 is written as follows: Y_1=0; Y_2=2; Y_3=7; Y_4=3.

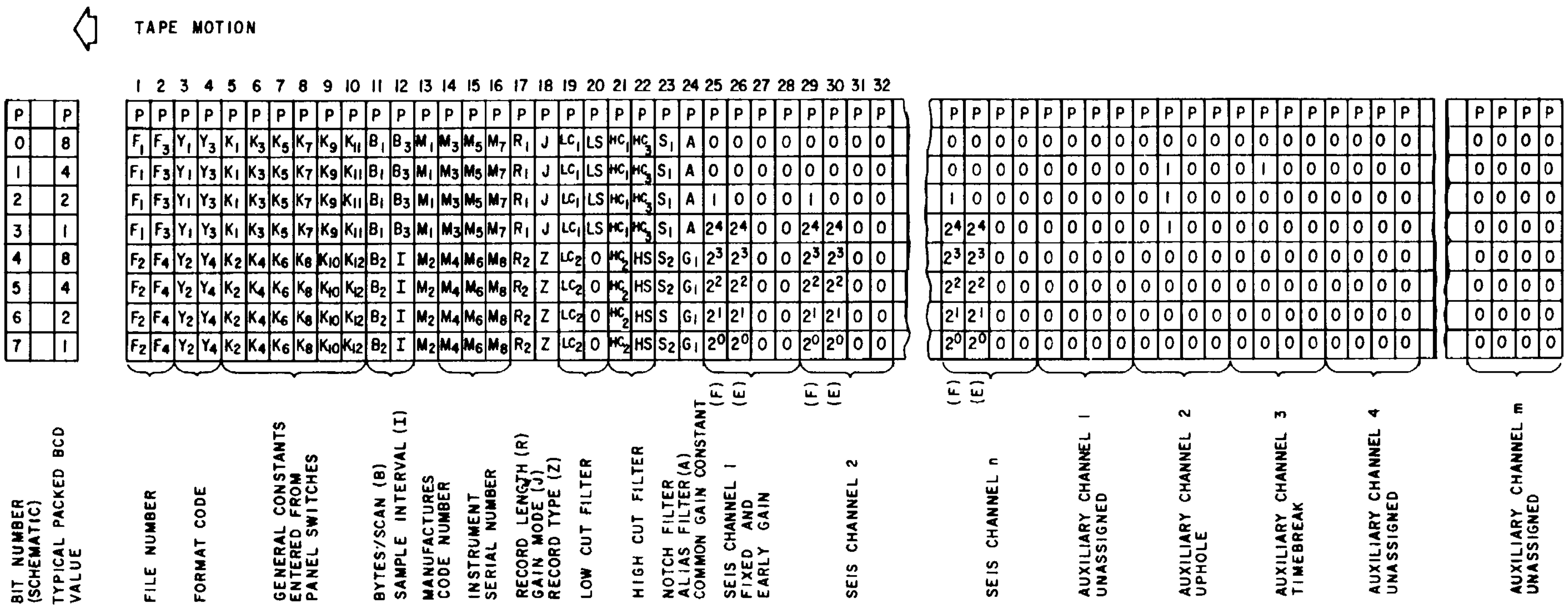

FIG. 3. Header block for Format C. Note: Additional externally supplied digital data may be recorded in the header block following byte 24 or end of the optional gain data.

A table describing the header block entries follows:

Byte No.	Letters	Description
1	F_1, F_2	4-digit file number.
2	F_3, F_4	
3	Y_1, Y_2	4-digit format code assigned by SEG upon filing of format.
4	Y_3, Y_4	
5	K_1, K_2	12-digit data identification constants (e.g., date, line no., reel No., etc.).
6	K_3, K_4	
7	K_5, K_6	
8	K_7, K_8	
9	K_9, K_{10}	
10	K_{11}, K_{12}	
11	B_1, B_2	Number of bytes per data scan.
12 MSD	B_3	
12 LSD	I	Sample interval in integral number of milliseconds.
13	M_1, M_2	2-digit manufacturers code.
14	M_3, M_4	6-digit equipment serial number.
15	M_5, M_6	
16	M_7, M_8	
17	R_1, R_2	Length of record in seconds. Code 00 indicates continuous recording.
18 MSD	J	Amplifier gain control mode: 8 = binary gain, 4 = programmed gain, 2 = ganged AGC, 1 = individual AGC, 9= floating point gain control.
18 LSD	Z	Type of record: 8 = shot, 4 = shot bridle, 2 = test, 1 =other.
19	LC_1, LC_2	2-digit low-cut filter setting.
20 MSD	LS	Low-cut filter slope setting in dB/octave, BCD number representing slope (to calculate actual slope, multiply number by 6 dB/octave).
20 LSD		All zeros.
21	HC_1, HC_2	First two digits of 3-digit highcut filter setting.
22 MSD	HC_3	Third digit of 3-digit high-cut filter setting.
22 LSD	HS	High-cut filter slope setting in dB/octave, BCD number representing slope (to calculate actual slope, multiply number by 6 dB/octave).
23	S_1,S_2	2-digit special filter setting (rejection or other).
24 MSD	A	Indicates sample rate of alias filter.
24 LSD	G	Common gain constant.

If it is necessary to record the initial gain of the seismic channels on an individual basis, this will begin with byte *25* and use one four byte word for each seismic channel. To maintain compatibility with the previous SEG header block, bits 3-7 of the first byte of each word are used for the fixed gain portion of either binary gain or true floating point amplifiers. This information is in binary form with the most significant bit in bit number 3 and the least significant bit in bit number 7. Bits 3-7 of the second byte are used for the initial setting of the variable gain portion of binary gain amplifiers if they are used. For true floating point amplifiers which have no initial gain settings, the second byte may be zero filled or the amplifier gain at the time of writing the header block may be recorded. For the unused bytes, zeros are recorded in bits 0-7 and one is recorded in bit P. Zeros are also recorded in unused bit positions of the first two bytes.

Again to maintain compatibility with the previous SEG header block, for each initial gain word, bits 0, 1, and 2 of the first byte contain the channel identifier code used to identify usage of each channel recorded as follows:

Bit no.			Type channel
0	**1**	**2**	
0	0	0	Unused channel.
1	0	0	Water-break channel.
0	1	0	Time-break channel.
0	0	1	Seismic channel.
1	0	1	Time counter.
0	1	1	Uphole channel.
1	1	1	Other.

Should information not defined in the standard header block be required, these data may be recorded beginning in byte 25, if individual initial gain settings are not recorded, or beginning in the next byte following the initial gain settings, when initial gains are recorded. In implementing this extension, an integer multiple of four bytes must be added; and unused bits should be zero filled.

To maintain integrity of the start-of-scan code, no data may be recorded in the header in a form which might produce three successive bytes of all ones.

DESCRIPTION OF DATA BLOCK

A detailed drawing of the data block is shown in Figure 4. It begins with an eight byte synchronization group which is recorded preceding seismic channel 1 at the start of each data scan. The first synchronization group recorded is coincident wi th the time-break.

To maintain compatibility with SEG Format A, the start-of-scan indication is provided by three successive bytes of all ones recorded in bits 0-7 of the first three bytes, and all zeros recorded in bits 0-7 of the fourth byte of the sync group. A recommended 15 bit binary time counter (counting time in milliseconds from the time-break) may be recorded in the fifth and sixth bytes of the sync group (T_1-T_{15}). The seventh and eighth bytes of the sync group contain all zeros in bits 0-7. There will be a single start-of-scan indication per scan regardless of the number of channels being recorded.

Each data value of a seismic channel is recorded in four successive bytes, using IBM compatible floating point notation as defined in IBM Form GA 22-6821, entitled "IBM System/360 Principles of Operation." The four bytes form a 32 bit word consisting of a sign bits Q_S, a seven bit exponent, or characteristic, Q_C, and a 24 bit fraction Q_F.

Q_S	Q_C		Q_F	
0	1	7	8	31

Q_S indicates signal polarity and is a one for negative. Q_C signifies a power of 16 expressed in excess 64 binary notation, allowing both negative and positive powers of 16 to be represented by a true number. Q_F is a six hexadecimal digit (24 binary bit) number with a radix point to the left of the most significant digit. The significance of the fraction is

$$2^{-1} + 2^{-2} + 2^{-3} ... +2^{-23} + 2^{-24}.$$

The data value represented by a floating point number is

$$Q_S \times 16^{(Q_C-64)} \times QF.$$

Consider a typical field recorder which uses a 15 bit analog-to-digital converter to measure channel output. The converter output is a sign bit and 14 binary magnitude bits, with the least significant magnitude bit weighted 2^{-14} (or 0.00006103515625) times the digitizer reference voltage, and the most significant bit weighted 2^{-1} (or 0.5) times the reference voltage. When the converter measures a positive full scale signal, ($1.0-2^{-14}$) times the reference voltage, its output is

0	1111 1111 1111 11
SIGN	MAGNITUDE

Thus, if the recorded data value is considered as a multiplier of the digitizer reference voltage, the full scale value recorded would be

0	**1000000**	**1111 1111 1111 1100 0000 0000**
Q_S	Q_C	Q_F

For the special case where the reference voltage can be expressed as a power of 16, the reference voltage can be incorporated in *Qc*. Multiplying the data value by 4096, or 16^3, yields the signal amplitude in millivolts if the reference voltage is 4096 mv. Thus, the recorded representation in millivolts in Format C would be

0	**1000011**	**1111 1111 1111 1100 0000 0000**
Q_S	Q_C	Q_F

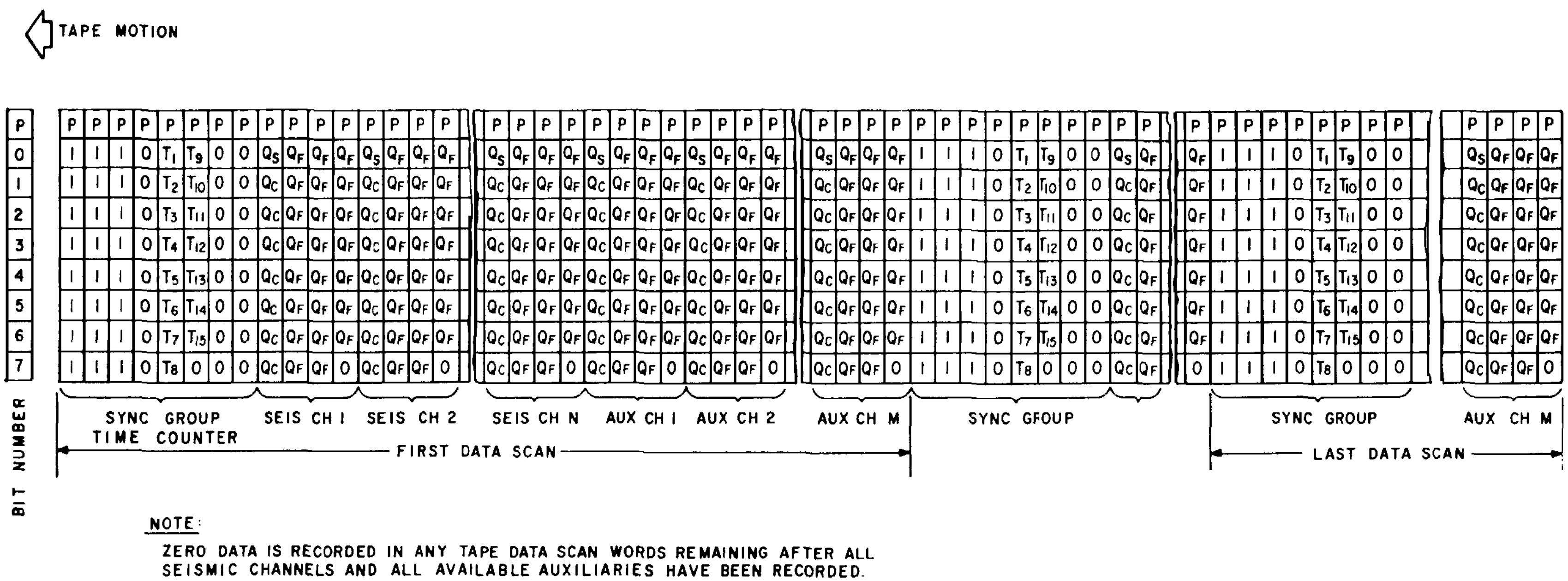

FIG. 4. Data block for Format C. Note: Zero data are recorded in any tape data scan words remaining after all seismic channels and all available auxiliaries have been recorded.

Note that the fraction is a true number regardless of signal polarity. A negative full-scale signal would be recorded in one's complement code in SEG Formats A and B as

1	0000 0000 0000 00
SIGN	MAGNITUDE

The corresponding representation in Format C is

1	**1000000**	**1111 1111 1111 1100 0000 0000**
Q_S	Q_C	Q_F

In this and all following examples, the value recorded in millivolts, for a reference voltage of 4096 my, can be obtained by adding 3 (binary 0000011) to *Qc.* Also note that the fraction is normalized; that is, the most significant digit (four most significant bits) *cannot* be zero unless the entire fraction is zero.

If the converter input is equal to the least significant bit, the converter output is

0	0000 0000 0000 01
SIGN	MAGNITUDE

The recorded representation is:

0	**0111101**	**0100 0000 0000 0000 0000 0000**
Q_S	Q_C	Q_F

which is obtained by left-shifting the magnitude 12 places (multiplying by 16^3) and subtracting 3 from the characteristic (dividing by 16^3).

For simplicity, the previous examples have dealt with channel output level. In a variable gain recording system, of course, the channel input, rather than the output, is the data of interest. In Format C, the recorded data value is the quotient of the channel output divided by the channel gain, or the input signal referenced to the digitizer reference voltage. The Format C representation of signals producing positive full scale outputs at various channel gains is shown below:

6dB	0 1000000 0111 1111 1111 1110 0000 0000
18dB	0 1000000 0001 1111 1111 1111 1000 0000
24dB	0 0111111 1111 1111 1111 1100 0000 0000
72dB	0 0111101 1111 1111 1111 1100 0000 0000
96dB	0 0111100 1111 1111 1111 1100 0000 0000

Either the variable portion of the channel gain alone or the total channel gain (fixed plus variable) may be used as the divisor. The total gain should always be used when all channels do not operate at the same fixed gain. When the total gain is used, recording of the fixed and early gains in the header block is optional for Format C.

For the special case where the channel output is zero, the recorded representation is always

0 0000000 0000 0000 0000 0000 0000 0000

regardless of channel gain.

Auxiliary channels are also recorded in four bytes each, with the bit configuration specialized to the types of data being recorded. Where the auxiliary channel data are a digitized analog signal, data are recorded in the same manner as seismic channels. If the auxiliary channel contains digital data, the digital data are left justified, so that the most significant bit is in bit 0 of the first byte of the word.

Although the fractional portion of the floating point notation can accommodate 24 bits significance, the examples above show that with no field summing, and with a 15 bit converter, there will be at least seven significant bits unused. Unused bits should be filled with zeros. To maintain integrity of the start-of-scan, bit 7 of the fourth byte of each seismic or auxiliary channel word must always be zero; and in the floating point format the largest negative number that can safely be written is $-5.9 \times 10^{+75}$. If an auxiliary channel is recorded using some notation other than the floating point notation, care must be taken that it cannot produce three successive bytes of all ones.

The assignment of channels as seismic or auxiliary channels is completely flexible, provided the seismic channels are sampled sequentially, beginning with channel 1 and followed by the auxiliary channels.

Suggested channel assignments for the 30 channel svstem are as follows:

Channel	Assignment
1-24	Seismic channels.
25	Auxiliary.
26	Auxiliary or equipment monitor.
27	Auxiliary.
28	Uphole or water breaks.
29	Time break.
30	Timing (400 Hz from an independent source).

EXPANSION OF FORMAT C TO MORE THAN 30 CHANNELS

For a 30 channel system, Format C utilizes 128 bytes per data scan. Expansion of the format beyond 30 channels will be in integer multiples of 128 bytes per data scan.

RECOMMENDED STANDARDS FOR 1600 BPI RECORDING OF SEG FORMATS A & B

The SEG Subcommittee on Field Tape Formats regards 1600 bpi phase encoded recording as essentially independent of any particular data format. For this reason, the committee recommends that both SEG Formats A and B allow either 800 bpi NRZI or 1600 bpi phase encoded recording, and that these formats remain unchanged in all other respects. Figures 1 and 2 illustrate general specifications of both recording modes.

The following is a description of 1600 bpi recording tape format:

1. APE identification burst is written beginning at least 1.7 inches before the trailing edge of the loadpoint reflective marker and continuing past the trailing edge of the marker. The identification burst consists of 1600 flux reversals per inch in track P with all other tracks dc erased.
2. A seismic record file is divided into a header block, containing constant identification information pertaining to the seismic record, and a seismic data block(s), containing data values of the seismic channels. A gap of at least 0.6 inch of erased tape separates the header block from the seismic data block(s). The first header block on tape begins at least 3.0 inches past the trailing edge of the loadpoint reflective marker.
3. Each block, header or data, must be preceded by a preamble consisting of 40 bytes of zero-bits in all tracks and one byte of all one-bits, and followed by a postamble consisting of one byte of all one-bits and 40 bytes of zero-bits in all tracks. The preamble of the first header block is separated from the identification burst by at least 0.5 inch of erased tape. The data block preamble must be written immediately after the minimum inter-record gap following the header block. Zero data (bit P a one and bits 0-7 all zero) may be written until the first start-of-scan is written.
4. An end-of-file mark is written after each seismic record file, with at least two end of file marks after the last seismic record file on the tape. The end-of-file mark is preceded by at least 0.5 inch of erased tape.
5. Tape specifications will be in accordance with Form GA *22-6862.*

ACKNOWLEDGMENTS

The various proposals were originally presented at a meeting attended by representatives from a number of oil companies, contractors, and instrument manufacturers. A preliminary version of the present recommendation was sent to each of these for comment. We gratefully acknowledge the support and helpful suggestions received.

REFERENCE

Northwood, E. J., Weisinger, R. C., and Bradley, J. J., 1967, Recommended standards for digital tape formats: Geophysics, v. *32,* p. 1073-1084.