VSC-TR-81-16

DIGITAL SEISMIC DATA REPRESENTATIONS: GAIN-RANGING, PRECISION, SYSTEM RESPONSE, SEISMIC AND COMPUTER CONSIDERATIONS.

Robert R. Blandford and Paul Kovacs
Seismic Data Analysis Center
Teledyne Geotech
314 Montgomery Street
Alexandria, Virginia 22314

30 DEC 1981

Monitored By:
VELA Seismological Center
312 Montgomery Street
Alexandria, VA 22314

Sponsored by

The Defense Advanced Research Projects Agency (DARPA)

DARPA Order No. 2551

Disclaimer: Neither the Defense Advanced Research Projects Agency nor the Air Force Technical Applications Center will be responsible for information contained herein which has been supplied by other organizations or contractors, and this document is subject to later revision as may be necessary. The views and conclusions presented are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, the Air Force Technical Applications Center, or the US Government.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER VSC-TR-81-16	2. GOVT ACCESSION NO. AD-A110 690	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DIGITAL SEISMIC DATA REPRESENTATIONS: GAIN-RANGING, PRECISION, SYSTEM RESPONSE, SEISMIC AND COMPUTER CONSIDERATIONS		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER SDAC-TR-80-3
7. AUTHOR(s) Robert Blandford, Paul Kovacs		8. CONTRACT OR GRANT NUMBER(s) ARPA Order-2551 F08606-79-C-0007
9. PERFORMING ORGANIZATION NAME AND ADDRESS Teledyne Geotech, 314 Montgomery Street, Alexandria, Virginia 22314		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS VT/0709
11. CONTROLLING OFFICE NAME AND ADDRESS VELA Seismological Center, 312 Montgomery Street, Alexandria, Virginia 22314		12. REPORT DATE 31 March 1980
		13. NUMBER OF PAGES 46
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Defense Advanced Research Projects Agency, 1400 Wilson Boulevard, Arlington, Virginia 22209		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

AFTAC-VELA-T

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Gain-ranging
Seismic data
Dynamic range
Data storage
Precision

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Because of trends toward increasingly precise analog-to-digital converters, we recommend that the data storage standard for the Seismic Data Analysis Center be a 16 bit word that is gain-ranged or becomes a 32-bit word only when necessary. There is no problem in converting any existing input format to this format. The amount of additional archival storage media used will be insignificant and rapid access can be assured. The primary disadvantage is the complexity of decoding and converting the archival record to a process desirable word.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified 408258

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DIGITAL SEISMIC DATA REPRESENTATIONS: GAIN-RANGING, DYNAMIC RANGE, PRECISION, SYSTEM RESPONSE, SEISMIC AND COMPUTER CONSIDERATIONS

SEISMIC DATA ANALYSIS CENTER REPORT NO.: SDAC-TR-80-3

AFTAC Project Authorization No.:	VELA/0709
Project Title:	Seismic Data Analysis Center
ARPA Order No.:	2551
Name of Contractor:	TELEDYNE GEOTECH
Contract No.:	F08606-79-C-0007
Date of Contract:	1 October 1979
Amount of Contract:	$279,927
Contract Expiration Date:	30 September 1980
Project Manager:	Robert R. Blandford (703) 836-3882

P.O. Box 334, Alexandria, Virginia 22314

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

ABSTRACT

Because of trends toward increasingly precise analog-to-digital converters, we recommend that the data storage standard for the Seismic Data Analysis Center be a 16-bit word that is gain-ranged or becomes a 32-bit word only when necessary. There is no problem in converting any existing input format to this format. The amount of additional archival storage media used will be insignificant and rapid access can be assured. The primary disadvantage is the complexity of decoding and converting the archival record to a process desirable word.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Digital seismic data were first gathered in large quantities at the LASA array, beginning in 1965. Since then, digital data has been gathered at a large number of sites and recorded in a number of digital formats. It is the goal of this report to survey the various formats and to recommend a format for storage which will, if possible, be easy to handle on computers, allow for efficient use of storage media by minimizing "empty" space in the data word, and which will not lead to loss of either precision or the dynamic range which has been captured by the field system.

The next section of this study gives a short introduction to the subject of gain-ranging. Next, we survey a few of the gain-ranging formats used heretofore and discuss some of the weaknesses which have been apparent in our experience. We also mention some of the confusion which has arisen in connection with the various formats. This last discussion is important because it points up the importance of simplicity and straightforwardness in the data format, both for ease of programming and for ease of understanding by the many workers using the format.

We follow this with a section on seismological constraints. Examples of seismological constraints are the desire of seismologists for a system response of particular shape, e.g., flat to displacement or of the standard narrow band character, which is good for detection of teleseismic signals. Other constraints are the desire to detect weak high-frequency signals; the desire not to clip on large signals; and the desire to measure the high-frequency energy from large signals. Finally, we give a recommended data storage format and the rationale used for the recommendation.

The first appendix gives a recommended format for recording seismic data; the second contains the statement of work.

GAIN-RANGING

The term "gain-ranging" in data collection systems is analogous to "floating point" in digital processing and computer systems. In both systems, it simply means that some of the bits in a register, memory location, or computer word are used to scale the remaining bits. The number of bits used to do the scaling, the value of the scaling bits, where they are positioned in the computer word, and the number of bits in the word, vary in both systems according to the application, the computer supplier, and architecture or design of the system. Digital computers, however, have floating point operations that add, subtract, multiply, and divide, the data values. In general, the format of the floating-point number in computers is more complex than gain-ranging numbers used in data collection systems and, as we shall see, like gain-ranging techniques, the format and methods used for digital computers vary among manufacturers and computers.

Gain-ranging is performed at the time analog values are converted into digital counts. As an example, consider an analog to digital converter that has 'n' bits as output and requires a voltage of v "to turn on" all of the bits. The resolution or the required voltage to "turn on" a single bit would be $v/(2^n-1)$. As the voltage increases, reaches, and goes beyond v, the condition is detected and the voltage is divided by two while the gain code is increased by one. The value of the voltage has been doubled by halving the gain prior to the digital conversion and the bit value is consequently doubled. The same bit position doubles for each successive increasing increment in the gain code.

Figure 1 depicts the bit patterns for the output of a sample 12 bit analog to digital converter for various input voltages. It also shows the equivalent expression for a four bit gain code and how the binary information could be placed in a 16 bit word. To simplify this presentation, only positive voltages are shown as input and only positive values are indicated in the bit patterns. The most significant bit, the one shown on the left of the output value of the digitizer and the mantissa, is the sign bit; it remains off for this

STEP	INPUT TO GAIN RANGING SYSTEM	VOLTAGE TO A/D	GAIN CODE	GAIN VALUE	12 BIT DIGITIZED VALUE	16 BIT WORD: GAIN	16 BIT WORD: MANTISSA
1	0	0	00 00	=0	00 00 00 00 00 00	00 00	00 00 00 00 00 00
2	b	b	00 00	=0	00 00 00 00 00 01	00 00	00 00 00 00 00 01
3	$\frac{V}{2}$	$\frac{V}{2}$	00 00	=0	01 00 00 00 00 00	00 00	01 00 00 00 00 00
4	V	V	00 00	=0	01 11 11 11 11 11	00 00	01 11 11 11 11 11
5	V+b	$\frac{(V+b)}{2}$	00 01	=1	0 10 00 00 00 00 0	00 01	01 00 00 00 00 00
6	2V	V	00 01	=1	0 11 11 11 11 11 1	00 01	01 11 11 11 11 11
7	2(V+b)	$\frac{(2V+2b)}{4} = \frac{(V+b)}{2}$	00 10	=2	0 1 00 00 00 00 00	00 10	01 00 00 00 00 00
8	4V	V	00 10	=2	0 1 11 11 11 11 11	00 10	01 11 11 11 11
9	4(V+b)	$\frac{(4V+4b)}{8} = \frac{(V+b)}{2}$	00 11	=3	0 1 0 00 00 00 00 0 0	00 11	01 00 00 00 00 00
10	8V	V	00 11	=3	0 1 1 11 11 11 11 1 1	00 11	01 11 11 11 11 11
•	•	•		•	•		•
•	•	•	•		•		•
•	•	•	•		•		•
•	•	•		•	•		•
•	•	•	•		•		•
•	•	•	•		•		•
33	16384(V+b)	$\frac{(16384V+16384b)}{32768} = \frac{(V+b)}{2}$	11 11	=15	01 00 00 00 00 00	11 11	01 00 00 00 00 00
34	32768V	V	11 11	=15	01 11 11 11 11 11	11 11	01 11 11 11 11 11

xx xx x| xx xx xx xx xx xx |x|x|x| xx xx xx xx x|x|x| 32 BIT REGISTER

GAIN CODE=15 GAIN CODE=0 GAIN CODE=1 GAIN CODE=2 GAIN CODE=3

NOTE
V= MAXIMUM VOLTAGE OF A/D
b= REQUIRED VOLTAGE TO TURN ON ONE BIT AT GAIN CODE ZERO

Figure 1. Bit Patterns for Various Gain Values

presentation. Note from the figure that step 1 represents the zero input and output condition of the system. The digitized values for the output of the converter, the gain code, and the resulting 16 bit computer word are all zero. Step 2 shows that input voltage to the system which sets only the least significant bit, changing only the output of the digitizer and mantissa, and leaving the gain code unchanged. Step 3 shows the binary output for half of the voltage range at the gain value of zero. Step 4 represents the voltage into the A/D which causes all of the bits to be set (except the sign bit) without affecting the gain code. The resulting 16 bit computer word also shows that the four bit gain portion is zero and the mantissa is set to the same value as the output of the converter. Throughout this example, the resulting computer word is indicated by simply joining the gain code with the A/D converter output and placing the gain code in the upper part of the 16 bit word.

Step 5 shows the condition which causes the first increment of the gain code and the next successive value of the mantissa. When this or any shift occurs to the next higher gain value, each bit in the mantissa requires twice the voltage input to the gain ranging system to turn it on as in the preceding gain code. Step 6 and all subsequent even numbered steps show the highest voltage for a given gain code to set all of the bits in the mantissa. Step 7 sets the gain code to two with the lowest mantissa value. At the bottom of the figure is a 32 bit register, showing where the 12 bits could be placed when the gain-ranged values are converted into integer quantities.

The sensor in the seismograph system is an analog device which produces a voltage in response to earth motion. The important parameters in digitizing these signals are the sampling rate, which is related to the frequency response, and the number of bits in the digital word, which determines the resolution and dynamic range.

Suppose that the A/D converter has a precision of 11 bits zero-to-peak. That is, when presented with a suitable voltage, (for example, 0.1 volt) it can accurately measure to $.1/(2^{11}-1) = 0.1/2047 =$ .0488 mv 0.05 mv. The output of the A/D would be one, if the voltage was between 0.05 and 0.1 mv, and would run up to 2047 counts for a voltage just below 0.1 v. In this case, the precision is 0.05 mv and the dynamic range is 20 $\log_{10}$ 2048 = 66.2 dB. (Note that one gets almost exactly 6 dB per bit, so that the easy way to compute dynamic range is to multiply the number of bits by six; in this case 11 x 6 = 66.)

Sixteen bits (15 bits plus sign) would give a dynamic range of 15 x 6 = 90 dB. If, however, the original 11 bits were enough to adequately sample the background noise, then we can obtain more than 90 dB of dynamic range by using the remaining four bits in the technique of gain-ranging. The analog circuitry in the A/D converter is then set so that if the voltage is below 0.1 v, the situation is the same as if there were no gain-ranging and all four bits in the gain code are zero. If the voltage is between 0.1 and 0.2 v, then the least significant count occurs from 0.1 to 0.2 mv and the first bit in the gain code is set to one. As the maximum voltage increases, the gain code increases, until all the four bits are set to one (representing the number 15 in the base 10). This then may be used to represent a gain in amplitude of 2^{15} or 90 dB. Thus, the overall dynamic range of this system is 66 + 90 = 156 dB; 66 dB more than can be obtained from 16 bits without gain-ranging.

It is, of course, just a convention that the next gain level is up a factor of two when the voltage exceeds the maximum in the lowest gain level. Other factors are possible, and in fact, factors of four have been used in both the LASA Hi-rate data and in the NSS system. The trade-off for using factors larger than two is that in the lower ranges of the data values there are less than 12 bits precision.

Note that with this large dynamic range, it is possible to set the lowest gain code or maximum sensitivity such that 2048 counts lies just above the ambient noise voltage as seen through the system. Then, the least count will be 66 dB down, and for most systems the analog noise (and small signals) will be adequately quantized at any frequency of interest. Then, the gain-ranging system will provide 90 dB more to avoid clipping on the largest events. It is details of this sort which must be considered when deciding if a particular format will be adequate for any particular system.

SURVEY OF GAIN-RANGING FORMATS

About the only standard used by the various seismic systems is the number of bits given to a data sample. No doubt this choice is based on the available hardware rather than on the requirements of the investigators. In any case, 16 bit samples are used by nearly everyone and nearly everyone uses them differently.

Twelve Bit

The one current exception to the 16 bit format is that some systems still have 12 bit converters and send the data in 12 bit two's complement form to conserve storage media and reduce transmission time.

In recovering data recorded in this format, it is noted that substantial computer time is required to unblock or format the 12 bit samples into either a 16 or 32 bit integer or a 32 bit floating point format. The process involves considerable record keeping in the code and numerous shift and pack operations, more than is suspected when the problem is given casual consideration.

Fourteen Bit

Fourteen bit ADC's are not uncommon and are used by the NORSAR and the old LASA. Most systems, however, store and transmit this information in a 16 bit two's complement integer word. Endless confusion persists because the last two bits are used for status. In some cases, this has caused an error in the scaling by a factor of four, and if the values of the bits are used as data, then when the status bits are on, the mix up causes an error in the data values.

Sixteen Bit

The most common format currently in use is the 16 bit integer two's complement. Although only one agency actually has a 16 bit ADC, the word length, in either full or half word integer, is readily handled by most digital processors (IBM, DEC 11's, TI Station Processor, HP, etc.).

LASA High Rate

The SP data during the LASA era was handled in a unique manner. The system had a 14 bit ADC, but only sent 10 bits to the SAAC (now SRC) for recording and processing. The 10 bits were formatted in a floating point of eight bits of two's complement for the integer mantissa, with a two bit base-four exponent. One wonders why the 14 bit ADC was implemented; presumably, the data at the Montana Data Center was recorded with this precision and could be obtained, if necessary.

The actual format of the 10 bit data value is: bits 2^0 and 2^1 are gain values, bits 2^2 through 2^8 had the mantissa and 2^9 was the sign bit. The bits selected from the 14 bit ADC to form the 10 bit data sample are shown below.

Mantissa	Gain Code	Scale Factor
bits 2^0 through 2^7	00	4^0
bits 2^2 through 2^9	01	4^1
bits 2^4 through 2^{11}	10	4^2
bits 2^6 through 2^{13}	11	4^3

Note that during the transition from one gain code to another, a two bit shift occurs, causing one bit less resolution of the sampled value for the lower half of the data counts for the data samples having non-zero gain codes. This is similar to the method used by IBM for floating point arithmetic, except a four bit shift occurs when the floating point exponent changes.

Geotech Gain Range (12/4)

This format is commonly referred to as the 12/4. It first appeared in the data from ALPA, which was a seismic array built by Geotech. It is the most commonly used gain-range format of all recorded gain-range data. This format was used as the standard for recording LP data. The format is that bits 2^0 through 2^{10} are the mantissa, 2^{11} the sign bit and bits 2^{12} through 2^{15} are used for the gain code. Numeric values are determined by:

$$V = D \times 2^{(S-G)},$$

where: V - fixed point data value,
D - digitized 12 bit sample,
S - scale factor, and
G - gain code.

Most systems have assigned 10 to the scale factor S and defined the range of the gain G to be from zero to 10, so that the smallest value is represented in hex by A000 and the largest positive value by 07FF. This scheme does not permit a legal value of hex 0000, so that no data are distinguishable from zero values. Since G ranges from zero to 10 and S is assigned a value of 10, gain codes of hex B, C, D, E, and F are illegal as data values and can be used for status information without adding to the transmission bandwidth or storage requirements.

ILPA and the LP data from the Station Processor have assigned zero to the value of S.

<u>Sandia Gain Range (14/2)</u>

The prototype for the National Seismic Stations (Sandia was the prime contractor) used a format known as 14/2. In this format, the gain code is two bits and takes on the following values:

Gain Bit Pattern	Gain Code	Scale Value
00	0	$2^0 = 1$
01	3	$2^3 = 8$
10	5	$2^5 = 32$
11	7	$2^7 = 128$

This method of gain ranging is similar to the IBM LASA 10 bit method in that it does not shift by one during gain code changes; in fact, the shift from one gain code to the next is three bits at the high limit of the data values and two bits at the lower values.

AFTAC (13/3)

The preliminary specification for the AFTAC systems is a 13/3 format. The data word format is that bits 2^0 through 2^{11} are the two's complement mantissa, 2^{12} the sign bit, and 2^{13} through 2^{15} the gain code, which has a range of zero through seven. The gain code is used as a power of four, similar to the IBM format. The values obtainable are:

Gain Bit Pattern	Gain Code	Scale Value
000	0	$2^0 = 4^0 = 1$
001	1	$2^2 = 4^1 = 4$
010	2	$2^4 = 4^2 = 16$
011	3	$2^6 = 4^3 = 64$
100	4	$2^8 = 4^4 = 256$
101	5	$2^{10} = 4^5 = 1024$
110	6	$2^{12} = 4^6 = 4096$
111	7	$2^{14} = 4^7 = 16384$

This scheme has the greatest dynamic range of the formats considered so far in this report and retains 12 bits of resolution of the sampled value.

Figure 2 is a table that summarizes and shows the dynamic range and bits of resolution of these various formats. The word format used by the various systems and agencies is also shown.

	12 BIT	14 BIT	16 BIT	LASA 10 BIT	Geotech 12/4	Sandia 14/2	Proposed 13/3
Dynamic Range(dB)	72	84	96	84	132	132	162
#BITS	11	13	15	6-7	11	10-13	11-12
LASA	LP			SP			
NORSAR	LP/SP						
Sta. Prcr.	SP				LP		
NSS						ALL	
SRO/ USGS			SP		LP		
ALPA					LP		
ILPA			SP		LP		
ALQ			SP				
SDAC			SP		LP		
DWWSSN			ALL				
HFS		LP					
SDCS	SP						
A/L	ANY						

Note 1: Dynamic range is negative to positive or peak to peak.

Note 2: LASA, NORSAR, and HFS LP are transmitted in 16 bits.

Note 3: The Geotech 12/4 has a maximum gain of 2^{10}.

Note 4: The Sta. Prcr. is used at ALK, KSRS, and PWY.

Figure 2. Summary of Sample formats

Let's consider a selected few of these formats and compare them to each other.

Figure 3 gives a graphic representation of the precision and dynamic range of the different systems. The vertical axis is a logarithmic one, and columns I-III give amplitude in decimal, bit-number, and dB, respectively. Columns IV-IX are each for a different digital format. Column IV is the simplest possible one, 11 bits plus sign with no gain-ranging, giving a total of 66 dB zero-to-peak. This is the capability of the SDAC A/D system, and of many other systems, such as the Alaskan SP systems.

The number one above the line at zero dB indicates that this is the bottom of the first (and only) gain-ranging level. The number one below the line at 66 dB indicates that this is the top of the gain level, and the difference in dB (66 minus zero) shows the precision available within the gain-range.

Column V is 11 bits plus sign, with four additional bits given over to gain-ranging. (This format was discussed earlier.) The fact that the total is 16 bits is significant. Almost all computers have hardware features that make the manipulation of 16-bit words highly efficient. Large amounts of special code are required to handle 17 bits (for example) as a unit. Sixteen bits are easily represented in the hexidecimal system. From all considerations except that of storage space, it is immensely preferable to handle data in units of 16 bits (two bytes).

We see that as the gain levels are increased, there is a constant 66 dB level of precision. While the maximum possible dynamic range for Column V is 156 dB, we have stopped the explicit depiction of gain levels at 120 dB. This is to make vivid the fact that many systems made little use of the upper gain levels.

Column VI gives the NSS gain levels. This system is remarkably efficient. At every gain level, it gives precision equal to or superior to that of Column V. The numbers in circles between Columns V

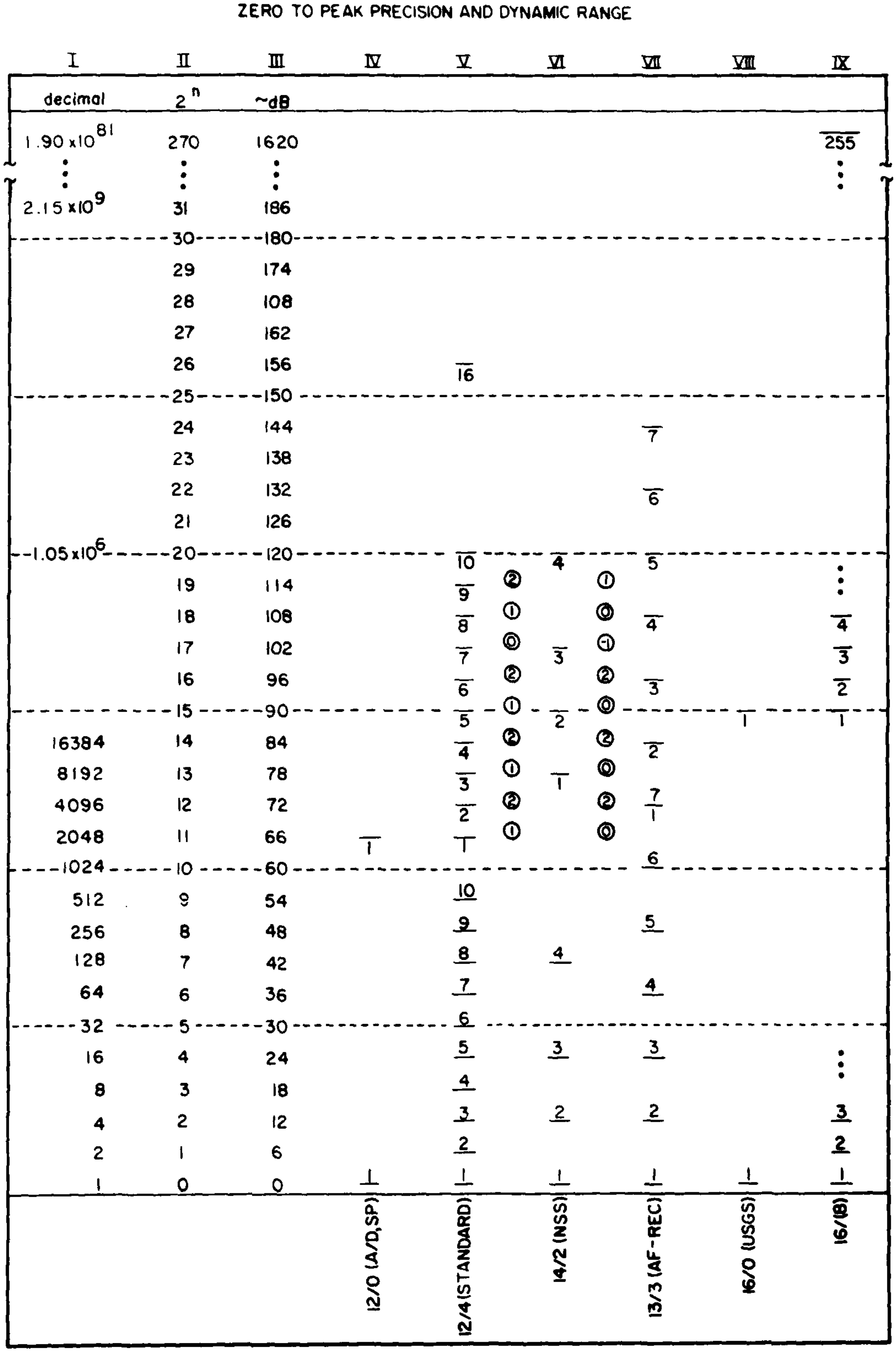

Figure 3. Representation of Dynamic Range and Precision of a Number of Selected Gain-Ranging Formats

and VI show the number of bits of precision by which Column VI is superior to Column V. Of course, Column VI does not have the dynamic range of Column V.

A similar remark can be made about the relationship between columns VI and VII. Column VII is the suggested 13/3 format. We see that, in general, below 120 dB it has less precision than Column VI. There is one exception, and improvement of one bit near the top of gain level four for Column VII.

Column VIII shows the straight 16-bit, no gain-ranging option used at a number of SP sites. We see that, compared to the gain-ranging formats, one must sacrifice either dynamic range or precision in the low-voltage range.

In Column IX we see the format recommended in the abstract and discussed in detail in the final section. We see, of course, that it easily surpasses all other formats in dynamic range and, except for the 16 bit A/D of the USGS, in precision.

At this point, a few remarks seem to be in order as to our experience with these formats.

Our experience with the 12/0 format as implemented in Alaska is that it is set at too small a number of counts/volt so that small signals appear as a "staircase" functions. Despite this, the signals often clip on large signals. This latter fault is compensated for by the existence of low-gain channels. Spectra of these channels shows that the quantization noise is above earth noise for f > 3 Hz and is only 20 dB down from the one Hz noise level. Thus, small, high-frequency regional events might not be detected because of the quantization noise.

Our experience with the 12/4 format has mostly been in connection with LP data. We have also used it as our internal SDAC format for some years. A great deal of confusion has arisen with respect to this format because of the convention adopted that the data value in counts is to be computed according to the formula given earlier of V =

$D[2^{(S-G)}]$ where V is the data value in counts, D is the integer out of the A/D, G is the value entered in the gain code, and S is the gain code for a scale factor. S is commonly set equal to 10 (in decimal notation). Confusion then arises because $2^{10} = 1024$ 1000. Analysts then are sometimes led to conclude erroneously that the calibration factors sent from the field in counts/millimicron must instead have been in counts/micron.

A correct analysis of the situation is as follows: The largest number of counts occurs when the system has the smallest gain corresponding to G = 0. Since the gain code is in the first four bits of the 16-bit word, then in hexidecimal notation the first four bits has the value zero. The smallest number of counts occurs at the highest gain and corresponds to G = 10 in decimal notation or G = A in hexidecimal. Thus for the smallest number, the first four bits has the value A in hexidecimal.

If one erroneously multiplies D by $2^{(-G)}$ instead of by $2^{(10-G)}$, then one may still come out with nearly the proper earth motion, if one assumes that the calibration is in counts per micron instead of counts per millimicron. However, we have observed that in practice the problem does not end there. The confusion over the proper shift leads to inadequate communication. Then, in ways hard to understand, there results incorrect selection of the significant bits when transformations are made from 16-bits without gain-ranging to 32-bit without gain-ranging and back to 16-bits without gain-ranging, as is done in the NEP system. Thus a more straightforward format is clearly desirable.

The NSS format has been found by Noponen (1979) to work well in conjunction with the four bandpasses selected by Sandia within which the voltage is proportional to earth velocity. For an Ms = 7.6 event at a distance of 20 degrees, the short-period channel did not clip, and the mid-period channel clipped only slightly. Since the long-period data channel could be recovered from the mid-period frequencies of interest, there was then almost no loss of data from this event. However, the fact that clipping did occur because of the upper limit of the gain-ranging formula before the signal became noticably non-linear suggests that a higher dynamic range would be desirable. Thus, we may look forward to field formats which require more than 16 bits since, as we shall see in the following section, all of the 14 bits of precision in the NSS can be of use in the problem of detecting a weak regional signal in the presence of a large teleseismic signal.

The existence of 16-bit A/D systems in several operational systems clearly points to a time when, to satisfy seismological constraints, these A/D's will be coupled to a gain-ranging system, requiring that the data be stored in a word larger than 16-bits, at least during the arrival of large signals. The dynamic range available from a straight 16-bit number is illustrated in Column VIII of Figure 3. We see that its dynamic range is well below that attained by the NSS system, and without gain-ranging it can not be satisfactory for an observatory designed to monitor a test ban treaty, unless unusual pains were taken to prewhiten the noise spectrum, or to break the full spectrum up into small frequency bands.

In Column IX of the same figure, we see the precision and dynamic range available to the SDAC format discussed in the abstract, and in more detail in the final section. We see, of course, that it has far greater capacity than these other systems.

SEISMOLOGICAL CONSTRAINTS

Each seismic data collection system has a response as a function of frequency such that earth displacement at a particular frequency generates a particular voltage. Some systems are flat to acceleration input, some to velocity (e.g., the NSS system) and some to displacement. However, most systems (e.g., the LRSM, WWSSN and AFTAC systems) have a more complicated response, one which peaks near a signal-to-noise maximum in order to detect signals.

The conventional procedure followed in digitizing these signals in the field is to set the least significant bit or count, of the A/D converter, to some small fraction of the voltage generated by the noise level at the detection frequency of maximum interest for detection. For example, AFTAC data typically has about 10 counts peak-to-peak for a frequency band 0.5 to 2.0 Hz. An advantage of this is that when large signals come in they do not clip. A disadvantage is that 10 Hz earth noise generates a very small voltage. If a small signal comes in from regional distances, it may therefore be well above the earth and analog system noise, but not be detectable in the digital data because the least count is for a larger voltage than the signal can generate.

Figure 4 is from Durham (1979) and shows the design of the low-noise model displacement spectrum which he used to deduce the passbands and gain levels for the NSS system. Upon this figure, we have superimposed a rough estimate of the spectrum of the P wave from the March 14, 1979, Ms = 7.6, Mexican earthquake as recorded at CPO. It is often said by seismologists that they would like to have recorded an undistorted displacement spectrum for analysis. To adequately record the noise displacement spectrum from the peak near eight seconds to the value at about 16 Hz would seem to require more than 86 dB, depending on the noise resolution required. If we adopt Durham's criterion of five counts in the noise, then this represents an additional 14 dB for a total of 100. This is, however, an overestimate since, as discussed by Durham, we are generally interested in adequate sampling of something like a half-octave band. The root-mean-square

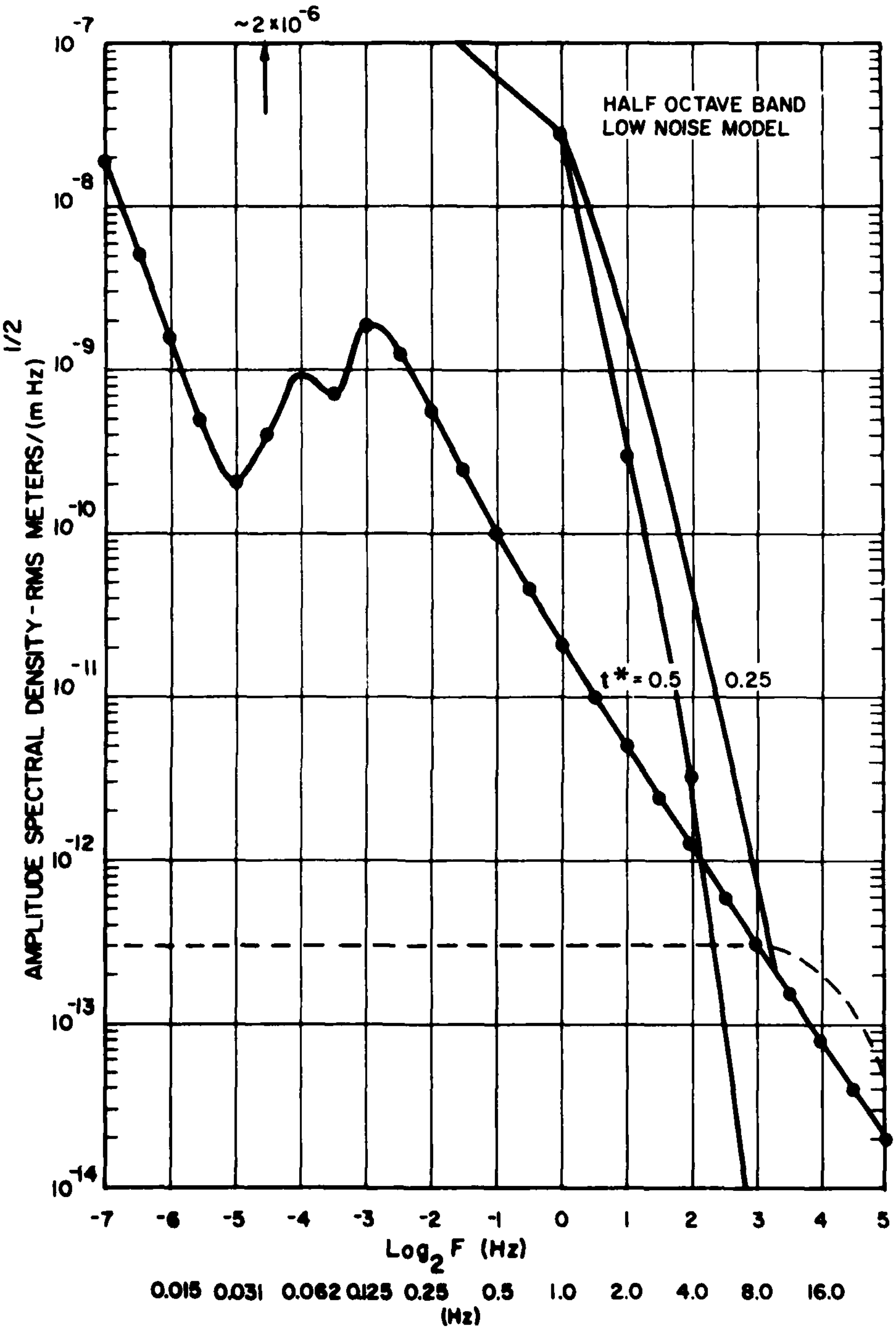

Figure 4. Ambient Displacement Noise Spectrum used as the basis of the NSS System Design by Durham (1979). Also shown is an estimate of the amplitude spectrum of the M_s = 7.6 Mexican earthquake on March 14, 1979 as observed at CPO; and a spectrum which realistically represents what would be observed at a distance of a few degrees from a small event.

amplitude in a half-octave band is proportional to the square-root of the frequency. Between eight seconds and 16 Hz, we have seven octaves and therefore, a relative increase in amplitude of (7/2) x 6 = 21 dB. Thus, to adequately sample the noise, there must be an A/D precision of 100 - 21 = 79 dB. We note, however, that when a larger signal comes in, such as that of March 14, 1979, that the signal level around eight seconds can rise as much as 60 dB, for a total range of 139 dB. This is much more than the precision available from an A/D and so, some other approach to the problem is required.

If gain-ranging is selected, then the large amplitude at low frequencies can be sampled, but the signal at high frequencies cannot be estimated because it will lie below the quantization level of the A/D. Another seismic consequence is that it will be impossible, in the presence of large signals, to detect small high frequency regional signals such as that indicated in Figure 4. At this point, we may choose to split the frequency window into several bands; or we may attempt to discover a representation other than displacement which will yield a smaller range. Selection of velocity instead of displacement yields six dB/octave for a total of 7 x 6 = 42 dB. This leaves a total required range for the A/D of 139 - 42 = 97 dB, still seven dB beyond the capability for a 16 bit A/D.

By considerations of this type, Durham (1979) arrived at a set of bandpasses and gain levels using a velocity response which resulted in a system with far better precision and dynamic range than those which existed previously.

The March 14, 1979 earthquake must be regarded as a severe test of this system. As mentioned before, the one Hz P wave and 25 second LR wave were recorded on scale. In addition, we can calculate that the eight Hz spectral density could be accurately estimated on the SP channel. The short-period maximum amplitude at one Hz was about 10% of clipping. Thus the system, in this frequency band, was not even in its highest gain level. The least count then was down 13 x 6 = 78 dB. From Figure 4 the eight Hz displacement was possibly below the earth noise which was down by 100 dB. However, we must remember that we have a velocity response and must allow for the 1/2 octave windows. In this way, we regain (3/2) x 6 + 3 x 6 = 27 dB; so, the net difference in amplitude between one and eight Hz is 100 - 27 = 73 dB. Thus, even in this case, the precision of this system would result in the detection at eight Hz of a small regional signal whose one Hz amplitude was 100 dB below that of the masking earthquake.

While some seismologists, mostly those with an academic background, may desire responses flat to displacement, workers in the VELA Uniform program will have a bias toward responses of the LRSM or similar systems which experience has shown to be generally satisfactory for the detection of teleseismic short-period P and long-period Rayleigh waves. The phases are of use in applying discrimination techniques. In accordance with the theorem that the optimum detection filters are proportional to $S/(N^2)$ where S is the signal amplitude spectrum and N is the noise amplitude spectrum, these responses generally fall off very fast at high frequencies (where absorption limits the signals) and outside the noise minima. This rapid fall-off requires a high precision A/D to adequately sample frequencies on the flanks of the response. Thus, the requirement for a response which leads to good detection and discrimination ability can be in conflict with the desire to adequately sample the entire spectrum in order to perform more sophisticated discrimination.

An overall system design might well include systems of digital filters, which enable the data to be recorded in the most suitable mode from the point of view of dynamic range and precision, and yet be presented to the user in the way most suitable to each task.

In this section we have not derived the optimum system response, merely outlined some consideration necessary in order to determine one. A point which should have been gained by the reader is that determination of the proper gain code is not simply a matter of bits and logarithms but also requires careful consideration of the goals of the project, and careful tailoring of the analog system response. It cannot yet be said that the NSS system is optimum but it is certainly many steps ahead of its predecessors.

RECOMMENDED FORMAT FOR DATA STORAGE

To be a candidate for archival storage a format should have sufficient precision and dynamic range that it will not be necessary to degrade data in order to store it. Secondly, since data in this format will be processed by many users, with both software and hardware, it would be good if the format were easy to remember and to manipulate. Finally, the format should be compact; that is, not use much more of a storage medium than is absolutely necessary to store the information.

There are so many different systems and data formats established that it seems impractical to attempt to standardize the output sample value. Moreover, the sample format selected reflects the dynamic range of the instrument and system digitizing the data. Consequently, the concept of standardizing the data at the headquarters or data center seems appropriate and increases the flexibility to change the data sources without requirements for the sender to comply with sample standards. Making this conversion early in the communication stream minimzes software complexity and data inconsistency. In our case, using the ILI's or a front end data concentrator (the CCP) to do this appears to be a sound approach.

Were we to use the 12/4 or 13/3 format, we would have to throw away precision from NSS data already stored and yet to come. While the loss of two bits here and there in the upper gain ranges does not seem particularly severe from a seismological point of view, it is certainly not an auspicious beginning. Furthermore, the existence of a 16-bit A/D device argues that in the near future even more bits would have to be thrown away.

Recommending a floating point format would be machine dependent, recommending any of the described gain-range formats or any 16-bit gained word would cause a loss of resolution for some of the systems and contribute to the processing time to unconditionally degain range for processing.

What is recommended is a variable length sample word that will encompass all of the formats that have been presented. Essentially, the stored word will be a 16-bit integer without a gain code unless it is needed to represent the data value. When a gain code is required, it is applied to each sample of all of the data during the time interval of the logical record being recorded. The gain code would only be required for signals well above the background noise level; consequently it would only be used a small perecentage of the time. It is also recommended that an eight bit gain code be used. Eight bits or a byte will facilitate computer manipulation of higher languages by allowing byte addressing and byte indexing. Determining whether or not the gain code is set would be part of the status information normally maintained for the data in the data stream. The existence of the gain code can be defined with a single bit.

The gain code recommended is an eight bit byte containing the integer quantity of the number of bits to left shift a 16 bit data sample when it is placed in the least significant 16 bits of an integer computer word.

A 32 bit two's complement integer value can easily be obtained from these conventions. If a data buffer does not contain gain ranged data, it is moved as an integer short (2 bytes or 16 bits) variables to an integer long (4 bytes or 32 bits) variable, with the sign extension automatically occurring in the hardware. Whether the gain codes precede or follow the data values is arbitrary; however, it is recommended that all of the gain codes be grouped together and separated from the block of data samples they are to scale in order to facilitate manipulation of the gain codes by byte instructions and maintain the data samples on 16 bit word boundaries.

The selection of this 16/(8) data sample format has a greater dynamic range than any of the formats reported herein. No precision is lost for any of the data words being sent by current and planned data sites. Gain codes are expected to be used a very small percent of the time, thereby showing a negligible increasing in storage costs.

The degree of difficulty in recovering the data from archive files at the record decoding level is great. Each record may be variable in length to account for a changing station/channel configuration as well as the presence or absence of the gain codes. Other checks concerning correct time values, time reversals, and I/O errors also make data recovery complicated. Consequently, the complexity of checking for, decoding, and formatting the gain ranged values does not significantly contribute to the overall processes of recovering data.

The uses of the sampled data may require that they be converted into floating point form for processes such as filtering or invoking detection algorithms using special array processor hardware devices. Other uses require integer data such as data displays on graphics system. Consequently, there is no recommendation of sample formats during the processing of the data. Each application would need to be analyzed for its most effective format.

We do recommend a method for recording the data on an intermediate storage device or disk. Our experience shows that data should be stored in a manner to minimize retrieval time, because most applications are related to some form of interactive processing. When retrievals are made they are usually done for a given time segment for a given station; often several data channels for each station are desired, such as all three components, SP, LP or perhaps even high and low gain. Because signals occur at different times at each of the stations, we recommend that the data be demultiplexed by station to enable rapid retrieval for a given time period.

When the section of the disk containing the proper station is determined, it remains only to find the disk area or record containing the desired start time. This access can be optimized if the time of the data is used to compute the disk address by way of a mapping scheme based on the data record length and disk blocking factor. This technique can be applied only if the data records are of fixed length. Because the gain codes of the 16/(8) format occasionally cause the record length to vary, they can be placed in a separate file or disk area. Their presence is determined by the status information header maintained for each data record for each second or time interval of data.

An alternate possibility is that if more than 16 bits are required (that instead of the 8 bit gain code) an entire 32 bit integer is stored in the separate file or disk.

Although this scheme would require additional disk access whenever the gain codes or 32 bit words are accessed, the conservation of disk space and reduction in time when accessing the data in the normal 16 bit mode is expected to more than compensate for the I/O occurring during the time the data are scaled according to the gain code.

A complete description of the format for the recorded data record is given in Appendix I.

ACRONYMS AND ABBREVIATIONS

A/D - Analog to Digital

ADC - Analog to Digital Converter

AFTAC - Air Force Technical Applications Center

ALK - Alaskan Array

dB - decibel

HFS - Hagfors Seismic Observatory

HP - Hewlett Packard

Hz - Hertz

I/O - Input Output

ILPA - Iranian Long Period Array

KSRS - Korean Seismic Research System

LP - Long Period

LASA - Large Aperature Seismic Array

NEP - Network Event Processor

NORSAR - Norwegian Seismic Array

NSS - National Seismic System

PWY - Pinedale Wyoming

SDAC - Seismic Data Analysis Center

SP - Short Period

Sta. Pcr. - Station Processor

TI - Texas Instruments

WWSSN - World Wide Seismic Station Network

ACKNOWLEDGEMENTS

The authors acknowledge valuable comments from W. Whyte, A. Hill, J. Martin, W. Dean, R. Baumstark, and J. Woolson.

REFERENCES

Noponen, I., E. B. McCoy and H. S. Sproules (1979). Evaluation of the National Seismic Station, SDAC-TR-79-8, Teledyne Geotech, Alexandria, Virginia.

Durham, H. B. (1979). NSS Seismic data system design goals and rationale, February 14, 1979, Sandia Laboratories, Albuquerque, New Mexico.

Staff ISRSPS General System Specification (SAAC) Ref. 101S Chapter 6. IBM Cooperation 31 January 1974.

APPENDIX I

DATA RECORD FORMAT

FOR

REAL TIME RECORDING

RECOMMENDED FORMAT FOR THE DATA RECORD

Introduction

Several kinds of record types are used while recording continuous data. Such things as the status of the various systems are recorded as well as the geophysical data. Geophysical data consist of several record types one of which is a record describing the configuration of the seismic stations. This information consists of static values (ground motion per count, site location, beam delays, beam deployment, etc.) that are recorded at the beginning of the tape and as changes occur. Other data may be segmented waveforms, detection logs, or event lists. This document concerns the recording format of only the continuous waveform data.

The continuous data represents at least 95% of the output file. To minimize recording media only the minimum of information is included in the output. Exceptions to this are the time and station identification fields. Experience has shown that this additional repetition is worthwhile for data searches and system development.

Each record contains data sampled for a time duration of one second. In general all of the times in the record will contain the same value thus placing the burden of time alignment in the recording system or any other programs used to create the data file.

The following information is organized by giving an overview of the record, then a brief description of each of the major sections within the record and finally a description of each of the data fields for each section of the record.

Overview of DA Record

```
| HEADER | STATION 1 | DATA | STATION 2 | DATA |.....| STATION n | DATA |
```

General Layout of the Record

HEADER Data

```
| Rcd. Type | Date/Time | No. Sta. | Sta. 1 Id. | Ptr. to Sta. 1 |
| Sta. 2 Id. | Ptr. to Sta. 2 |...    ...| Sta. n Id | # Byte n |
```

STATION Data

```
| Sta. Id. | Date/time |Ptrs. to Data | # Data Types | # Smpls/data type |
| # Chnls/data type | ASCII Data field |
```

DATA Description

```
| Status Chnls 1-n | Chnl. 1 ID | Data Chnl 1 | (Gain) |
                   | Chnl. 2 ID | Data Chnl 2 | (Gain) |
                                          .
                                          .
                                          .
                   | Chnl. n ID | Data Chnl n | (Gain) |
```

DETAIL RECORD DEFINITION

HEADER Description

Conventions used in developing this description were:

a) Lower case letters used as subscripts represent the limit of the index.

b) The following codes are used to define the data types:
   I*2 - two byte integer quantity,
   I*4 - four byte integer quantity,
   C*n - quantity having n bytes.

c) All offsets are relative to byte zero of the beginning of the physical record. For descriptions in sections other than for the header record the byte positions or offsets are shown relative to the section itself and the offsets are computed by the noted algorithms.

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 1	C*2	RECTYP	Record type, always 'DA'.
2 - 5	I*4	DAYRCD	Day of data relative to the year 1800.
6 - 9	I*4	TIMRCD	Time in 600ths of seconds from 0000 GMT to the whole second.
10 - 11	I*2	NOSTA	Number of stations included in physical record.
12 - 16	C*5	STAID(1)	Station identification for first station.
17			Pad.
18 - 19	I*2	PTRSTA(1)	Pointer to station description and data for first station.
20 - 24	C*5	STAID(2)	Station identification for second station.
25			Pad.
26 - 27	I*2	PTRSTA(2)	Pointer to station description and data for second station.
.	.	.	.
.	.	.	.
	C*5	STAID (NOSTA)	Station identification for the last (NOSTA) station. Begins in byte position 12+8*(NOSTA-1).
	I*2	PTRSTA (NOSTA)	Pointer to station description and data for last station in bytes. Begins in byte position 18+8*(NOSTA-1).

STATION Description

The following byte offsets are relative to PTRSTA(1).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 4	C*5	SITEID(1)	Station identification for first station.
6 - 9	I*4	DAY(1)	Day of data for site one relative to the year 1800 of the first station.
10 - 13	I*4	TIME(1)	Time in 600ths of seconds since 0000 GMT for the first station.
14 - 15	I*2	DATATP(1)	Number of data types for the first station, SP, LP, MP, etc. Each data type must have same number of samples but same number of samples can be used for more than one data type.
16 - 17	I*2	PTRDAT(1,1)	Pointer to beginning of data for first station and first data type.
18 - 19	I*2	PTRDAT(1,2)	Pointer to beginning of data for first station and second data type.
.	.	.	.
.	.	.	.
	I*2	PTRDAT(1,g)	Pointer to beginning of data for first station and last data type. NOTE: g=DATATP(1)

The following byte offsets are relative to PTRSTA(1)+16+2*DATATP(1).

0 - 1	I*2	SMPSEC(1,1)	Number of samples per second for first station and first data type.
2 - 3	I*2	SMPSEC(1,2)	Number of samples per second for first station and second data type.
.	.	.	.
.	.	.	.
	I*2	SMPSEC(1,g)	Number of samples per second for first station and last data type. NOTE: g=DATATP(1)

The following byte offsets are relative to PTRSTA(1)+16+4*DATATP(1).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 1	I*2	NUMCHN(1,1)	Number of channels for first station and first data type.
2 - 3	I*2	NUMCHN(1,2)	Number of channels for first station and second data type.
.	.	.	.
.	.	.	.
	I*2	NUMCHN(1,g)	Number of channels for first station and last data type. NOTE: g=DATATP(1)
	C*n	CMNTS(1,h)	Area reserved for unformatted data, normally used for comments or text. This field is padded to have an even number of bytes. NOTE: h=PTRDAT(1,1)-(PTRSTA(1)+16+6*DATATP(1))

DATA Description

The following byte offsets are relative to PTRDAT(1,1).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0	C*1	STATCH (1,1,1)	Status for first station, first data type, and first channel where the upper four bits of the eight bit byte are used as: Bit Definition 2**0 Data missing, 0 - no, 1 - yes; where the data are normally expected but is physically omitted from the record. The channel ID will remain even if the data are absent. 2**1 Data invalid, 0 - no, 1 - yes; where there are data present but it is not meaningful seismic information such as a calibration or faulty instrument. 2**2 Gain ranged, 0 - no, 1 - yes; where gain codes are written at the end of the data values of each group of samples per second. 2**3 Unused, set to 0.
0	C*1	STATCH (1,1,1)	Status for channel two where the lower four bits are defined above.
1	C*1	STATCH (1,1,2)	Status for channel three where the upper four bits are defined above.
1	C*1	STATCH (1,1,2)	Status for channel four where the lower four bits are defined above.
.	.	.	.
.	.	.	.
	C*1	STATCH (1,1,f)	Status for last channel where the four bits are defined above. NOTE: f=(NUMCHN(1,1)+1)/2

All index computations are based on the availability of data and gain codes. For the few examples given it was assumed that data were present and indicated as such in bit 2**0 of the status field.

The following byte offsets are relative to
PTRDAT(1,1)+(NUMCHN(1,1)+1)/2.

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 3	C*4	CHNLID (1,1,1)	Channel identification for station one, data type one, and channel one.
4 - 5	I*2	DATAPT (1,1,1,1)	Station one - s[1], data type one - d[1], channel one - c[1], and sample point one - p[1].
6 - 7	I*2	DATAPT (1,1,1,2)	Station one - s[1], data type one - d[1], channel one - c[1], and sample point two - p[2].
.	.	.	.
.	.	.	.
	I*2	DATAPT (1,1,1,e)	s[1],d[1],c[1],p[e] e=SMPSEC(1,1)

If bit 2**2 is set in the first four bits of STATCH(1,1,1) then the following SMPSEC(1,1) bytes contain gain codes.

The following byte offsets are relative to
PTRDAT(1,1)+(NUMCHN(1,1)+1)/2+2*SMPSEC(1,1),
if gain indicator set for first channel then
+SMPSEC(1,1).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 3	C*4	CHNLID (1,1,2)	Channel identification for station one, data type one, and channel two.
4 - 5	I*2	DATAPT (1,1,2,1)	s[1],d[1],c[2],p[1]
6 - 7	I*2	DATAPT (1,1,2,2)	s[1],d[1],c[2],p[2]
.	.	.	.
.	.	.	.
	I*2	DATAPT (1,1,2,e)	s[1],d[1],c[2],p[e] NOTE: e=SMPSEC(1,1)

If bit 2**2 is set in the last four bits of STATCH(1,1,1) then the following SMPSEC(1,1) bytes contain gain codes.

. . . .

. . . .

. . . .
. . . .

The following byte offsets are relative to
PTRDAT(1,1)+(NUMCHN(1,1)+1)/2+2*f*SMPSEC(1,1),
where f=NUMCHN(1,1),
if gain indicator set for previous channels then add
the number of times it occurred times SMPSEC(1,1).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 3	C*4	CHNLID (1,1,f)	Channel identification for station one, data type one, and channel f.
4 - 5	I*2	DATAPT (1,1,f,1)	s[1],d[1],c[f],p[1]
6 - 7	I*2	DATAPT (1,1,f,2)	s[1],d[1],c[f],p[2]
.	.	.	.
.	.	.	.
	I*2	DATAPT (1,1,f,e)	s[1],d[1],c[f],p[e] NOTE: e=SMPSEC(1,1) f=NUMCHN(1,1)

If bit 2**2 is set in the appropriate four bits of STATCH(1,1,f)
then the following SMPSEC(1,1) bytes contain gain codes.

The following byte offsets are relative to
PTRDAT(1,2).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0	C*1	STATCH (1,2,1)	Status for first station, second data type, and first channel where the four bits were defined previously.
0	C*1	STATCH (1,2,1)	Status for channel two.
1	C*1	STATCH (1,2,2)	Status for channel three.
1	C*1	STATCH (1,2,2)	Status for channel four.
.	.	.	.
.	.	.	.
	C*1	STATCH (1,2,f)	Status for last channel. NOTE: f=(NUMCHN(1,2)+1)/2

The following byte offsets are relative to
PTRDAT(1,2)+(NUMCHN(1,2)+1)/2.

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 3	C*4	CHNLID (1,2,1)	Channel identification for station one, data type two, and channel one.
4 - 5	I*2	DATAPT (1,2,1,1)	s[1],d[2],c[1],p[1]
6 - 7	I*2	DATAPT (1,2,1,2)	s[1],d[2],c[1],p[2]
.	.	.	.
.	.	.	.
	I*2	DATAPT (1,2,1,e)	s[1],d[2],c[1],p[e] e=SMPSEC(1,2)

If bit 2**2 is set in the first four bits of STATCH(1,2,1) then the following SMPSEC(1,2) bytes contain gain codes.

The following byte offsets are relative to
PTRDAT(1,2)+(NUMCHN(1,2)+1)/2+2*SMPSEC(1,2),
if gain indicator set for first channel then
+SMPSEC(1,2).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 3	C*4	CHNLID (1,2,2)	Channel identification for station one, data type two, and channel two.
4 - 5	I*2	DATAPT (1,2,2,1)	s[1],d[2],c[2],p[1]
6 - 7	I*2	DATAPT (1,2,2,2)	s[1],d[2],c[2],p[2]
.	.	.	.
.	.	.	.
	I*2	DATAPT (1,2,2,e)	s[1],d[2],c[2],p[e] NOTE: e=SMPSEC(1,2)

If bit 2**2 is set in the last four bits of STATCH(1,2,1) then the following SMPSEC(1,2) bytes contain gain codes.

.	.	.	.
.	.	.	.

. . . .
. . . .

The following byte offsets are relative to
PTRDAT(1,2)+(NUMCHN(1,2)+1)/2+2*f*SMPSEC(1,2),
where f=NUMCHN(1,2),
if gain indicator set for previous channels then add
the number of times it occurred times SMPSEC(1,2).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 3	C*4	CHNLID (1,2,f)	Channel identification for station one, data type two. and channel f.
4 - 5	I*2	DATAPT (1,2,f,1)	s[1],d[2],c[f],p[1]
6 - 7	I*2	DATAPT (1,2,f,2)	s[1],d[2],c[f],p[2]
.	.	.	.
.	.	.	.
	I*2	DATAPT (1,2,f,e)	s[1],d[2],c[f],p[e] NOTE: e=SMPSEC(1,2) f=NUMCHN(1,2)

If bit 2**2 is set in the appropriate four bits of STATCH(1,2,f)
then the following SMPSEC(1,2) bytes contain gain codes.

. . . .
. . . .

. . . .
. . . .

The following byte offsets are relative to
PTRDAT(1,b), where b=DATATP(1).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0	C*1	STATCH (1,b,1)	Status for first station, last data type, and first channel where the four bits were defined previously.
0	C*1	STATCH (1,b,1)	Status for channel two.
1	C*1	STATCH (1,b,2)	Status for channel three.
1	C*1	STATCH (1,b,2)	Status for channel four.
.	.	.	.
.	.	.	.
	C*1	STATCH (1,b,f)	Status for last channel. NOTE: f=(NUMCHN(1,b)+1)/2 b=DATATP(1)

The following byte offsets are relative to
PTRDAT(1,b)+(NUMCHN(1,b)+1)/2, where b=DATATP(1).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 3	C*4	CHNLID (1,b,1)	Channel identification for station one, data type b, and channel one.
4 - 5	I*2	DATAPT (1,b,1,1)	s[1],d[b],c[1],p[1]
6 - 7	I*2	DATAPT (1,b,1,2)	s[1],d[b],c[1],p[2]
.	.	.	.
.	.	.	.
	I*2	DATAPT (1,b,1,e)	s[1],d[b],c[1],p[e] e=SMPSEC(1,b)

If bit 2**2 is set in the first four bits of STATCH(1,b,1) then the following SMPSEC(1,b) bytes contain gain codes.

The following byte offsets are relative to
PTRDAT(1,b)+(NUMCHN(1,b)+1)/2+2*SMPSEC(1,b),
where b=DATATP(1)
if gain indicator set for first channel then
+SMPSEC(1,b).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 3	C*4	CHNLID (1,b,2)	Channel identification for station one, data type b, and channel two.
4 - 5	I*2	DATAPT (1,b,2,1)	s[1],d[b],c[2],p[1]
6 - 7	I*2	DATAPT (1,b,2,2)	s[1],d[b],c[2],p[2]
.	.	.	.
.	.	.	.
	I*2	DATAPT (1,b,2,e)	s[1],d[b],c[2],p[e] NOTE: e=SMPSEC(1,b)

If bit 2**2 is set in the last four bits of STATCH(1,b,2) then the following SMPSEC(1,b) bytes contain gain codes.

. . . .

. . . .

.   .   .   .
.   .   .   .

The following byte offsets are relative to
PTRDAT(1,b)+(NUMCHN(1,b)+1)/2+2*f*SMPSEC(1,b),
where f=NUMCHN(1,b), and b=DATATP(1)
if gain indicator set for previous channels then add
the number of times it occurred times SMPSEC(1,b).

BYTE	TYPE	MNEMONIC	DESCRIPTION
0 - 3	C*4	CHNLID (1,b,f)	Channel identification for station one, data type b, and channel f.
4 - 5	I*2	DATAPT (1,b,f,1)	s[1],d[b],c[f],p[1]
6 - 7	I*2	DATAPT (1,b,f,2)	s[1],d[b],c[f],p[2]
.	.	.	.
.	.	.	.
	I*2	DATAPT (1,b,f.e)	s[1],d[b],c[f],p[e] NOTE: e=SMPSEC(1,b) f=NUMCHN(1,b) b=DATATP(1)

If bit 2**2 is set in the appropriate four bits of STATCH(1,b,f)
then the following SMPSEC(1,b) bytes contain gain codes.

The remaining information in the record consists of data from the second to the NOSTA stations. The structure and format for these data are repetitious to the data described for station one.

APPENDIX II

STATEMENT OF WORK

STATEMENT OF WORK

Special Seismic Verification Study, PA VT/0709

HQ ESMC/PMP (Mr. W. T. Yearty)

1. References:

a. DARPA/NMRO letter to AFTAC/VSC, subject: Special Seismic Verification Studies, 22 Feb 80.

b. Contract F08606-79-C-0007, SOW Task 4.6

2. Reference 1a requests that a special study on seismic data sample representation technqiues be undertaken. Reference 1b makes provision for such studies.

3. The contractor should be directed to perform a special study addressing seismic data sample representations. Specific requiremens are as follows:

a. Identify all digital seismic data sample representations either currently in use or proposed for use by recognized members of the seismic community.

(1) Determine the advantages and disadvantages of the dynamic range and solution characteristics of each representation with respect to such issues as signal and noise detectability, signal separation, and signal dynamic range.

(2) Evaluate the effort required to convert each representation into a form suitable for processing on standard computers and, in the case of the NSS gain range formula, discuss any special hardware or software requirements for the conversion.

(3) Compare the different data representation techniques with each other in the areas discussed above.

b. Recommend a standard representation based upon the results of para 3a. This recommendation should also take into account loss of resolution in specific intervals as a result of the conversion, and a statement should be made in each such case as to the seismological implications of the loss of resolution.

c. Write a report covering the study and the results. In addition, the report shall be prefaced with a section designed to educate the reader on the subject of gain ranging (e.g., its purpose, how it works, the basis for choosing a particular gain range format, etc.). The report should be submitted NLT 31 Mar 80.

4. Any questions concerning this matter should be directed to Capt Paul Terry of this office, AV 221-7572.

FOR THE COMMANDER

LARRY J. COE, Lt Colonel, USAF
Chief, VELA Seismological Center