NRL Report 6522

A System for Measuring the First and Second Order Probability Distributions of Signals and Noise

September 11, 1967

NAVAL RESEARCH LABORATORY
Washington, D.C.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

NRL Report 6522

A System for Measuring the First and Second Order Probability Distributions of Signals and Noise

C. McCoy, Jr. and H. L. Peterson

Electronics Branch
Sound Division

September 11, 1967

NAVAL RESEARCH LABORATORY
Washington, D.C.

CONTENTS

Abstract	iv
Problem Status	iv
Authorization	iv
INTRODUCTION	1
MEASUREMENT OF PROBABILITY DENSITY OF A CONTINUOUS WAVEFORM	2
Quantization	2
First-Order Probability Density	3
Second-Order Probability Density	6
GENERAL DESCRIPTION	9
DETAILED DESCRIPTION	10
Input Section	10
Core Memory	15
Binary Counter	16
Visual Section	16
Controller	17
RESULTS	21
FUTURE PLANS	26
CONCLUSIONS	27
REFERENCES	27
APPENDIX A — Legend	29

ABSTRACT

A system has been designed and developed for use, either in the laboratory or field, to measure the first- and second-order amplitude-probability densities of signals and noise in the area of underwater acoustics. The system consists of a wideband dc amplifier feeding a sample-and-hold network, followed by an analog-to-digital converter to encode the sampled instantaneous value to an amplitude resolution of ten-bits and sign. These digital codes become the address of a magnetic-core memory, and the value stored in a given address location is a tally for each time that address has been called up. Since the address was directed by the instantaneous amplitude value of a time function, the stored tallies are proportional to the amplitude density of that given time function when plotted out as a histogram. These measured values can be read into a general-purpose digital computer as desired to compute the normalized density function together with moments.

This method of code sorting can be applied to other problems of instrumentation, e.g., calibration checks of analog-to-digital converters, stability of analog magnetic tape recorders, and "drop-out" measurements in analog magnetic tape recording.

PROBLEM STATUS

This is an interim report on the problem; work is continuing.

AUTHORIZATION

NRL Problem S01-06
RF 101-03-44-4054

Manuscript submitted December 12, 1966.

A SYSTEM FOR MEASURING THE FIRST AND SECOND ORDER PROBABILITY DISTRIBUTIONS OF SIGNALS AND NOISE

INTRODUCTION

Various devices and methods (1-6) have been implemented at the Naval Research Laboratory to measure and estimate the first, second, and higher orders of the probability densities of amplitude of continuous voltage from experiments in underwater acoustics. The probability-density distribution of a time function, whether periodic or random, can be estimated by sampling the function and plotting the frequency of occurrence versus levels of amplitude. The larger the number of samples the better the estimate of the true statistical process, especially for the random case. The principal use for the analyzer has been to measure distributions of underwater acoustic signals and noise. Periodic functions are used only for calibration purposes.

Early density analyzers (7,8) employed various techniques for counting the time spent in a channel. Several methods employed a step analog transfer characteristic such that there was a fixed voltage output during the time the waveform was in a preset channel, and zero otherwise. The summation was done with an analog integrator. Another method was to modulate the pulse height of a pulse train, generated by a pulse generator. The summation was done by a pulse-height discriminator and a counter following standard nuclear pulse-height-counting techniques. The summing device was rezeroed in both of these methods, and a level was selected before the start or recording of a run. The procedure was to run for some fixed time and read out the contents of the integrator or the counter after completion of the run. The number read out was a tally for the frequency of occurrence in a given amplitude range or value. Runs had to be made for each level of amplitude to get the density over the entire range of amplitude levels. It becomes readily apparent that these methods are laborious and time consuming, since the analog signal must be tallied for a given level during a run. The tallies for all other levels must be neglected during that run, or at best only a few levels can be handled in parallel during a single run. To get the density distribution using these techniques, one must take the ensemble of all the runs to complete the analysis.

The most desirable method would utilize a system that could sample into the analog signal and keep a tally of all the incremental occurrences of the amplitude. These stored tally values would form the histogram of the amplitude-probability density of the given time function and would require only a single run by the system.

With the advent of accurate analog-to-digital converters, the quantized analog values, represented as digital numbers, can be handled with standard digital techniques by sampling the analog signal at some given rate. The speed, accuracy, and reliability of digitizing analog signals eliminates the difficulties of drift which plagued early methods. For example, it is easier to adjust current weights in an A/D converter than it is to adjust the transfer characteristics of an analog-to-pulse-height modulator over a wide range of voltage levels. Precision and uniformity of channel widths are basic requirements for A/D converters to obtain good density measurements.

This report describes a system to do real-time processing on an analog signal by converting it to digital form and obtaining the first- or second-order frequency of occurrences. From these numbers the corresponding amplitude-density distributions are computed. This system also can handle data while in digital numbers as stored on

magnetic tape. The system is, in effect, a low-cost, special-purpose computer that is mobile for field or laboratory use, where one can get a look at the amplitude-density distribution of some random or periodic time function, within a frequency range limited only by the A/D converter or other electronic considerations.

MEASUREMENT OF PROBABILITY DENSITY OF A CONTINUOUS WAVEFORM

Quantization

Any time-varying function, $x(t)$, whether random or periodic, can be plotted as amplitude versus time t. When one views such a graph, it becomes apparent that the amplitude appears in the neighborhood of some values more often than at others. It is also apparent that one cannot ascertain how much time is spent at a given amplitude, since that amplitude defines only a point in time. It is more meaningful to speak of the time spent in the vicinity of an amplitude. The projection of amplitude values on the time axis defines just a series of points, and a summation of these points provides no measure with respect to time.

It is advantageous when dealing with a continuous waveform that one quantize an analog signal before attempting to resolve the probability density. Figure 1 shows a continuous waveform, $x(t)$. On the right is a plot of the continuous probability-density function that describes the density $p(x)$ of all values of voltage amplitude x.

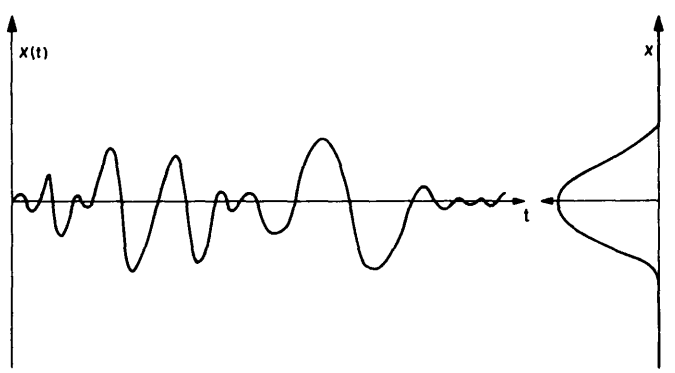


Fig. 1 - Continuous waveform $x(t)$ with corresponding probability-density function

If on the other hand a large time interval t is chosen, along with an amplitude interval Δx, then for each amplitude x, this region of amplitude is limited by $x + \Delta x$. The projection on the time axis of the segment of the waveform whose amplitude falls within this region $x + \Delta x$ defines our time interval.

For this reason, and also from a practical standpoint, the waveform characteristics are measured in a channel or window of amplitude Δx. Figure 2 represents the same function of time $x(t)$ plotted on a grid of channels. The amount of time spent in each channel can be shown at the right as before. This is a quantized density. There is the same area under the density function as before, but here the channel width is a finite value Δx, whereas in the continuous case the channel width is infinitesimally narrow. The quantization device, as previously mentioned, is an analog-to-digital voltage converter.

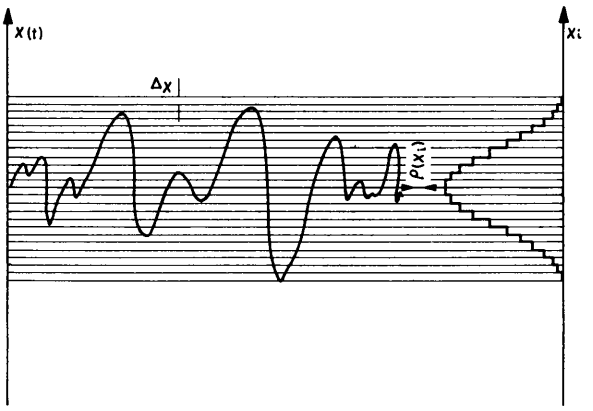


Fig. 2 - Function $x(t)$ with histogram showing the quantized density function

Assume that a random waveform (Fig. 3a) is to be analyzed to estimate the probability-density function of instantaneous amplitudes. The waveform is sampled periodically. This sampling gives a series of instantaneous amplitude values at the sample points (Fig. 3b). Each sampled value as it occurs is held and quantized into one of the coded levels by the action of the converter. An ith amplitude level indicates that the voltage sample had a value somewhere in the range from x_i to $x_i + \Delta x$, where i corresponds to the encoded number associated with that channel or level and Δx represents the width of the channel, which is essentially the resolution afforded by the least significant bit. Our measurements have been with a converter using ten bits and sign, which quantizes the voltage excursion into 1024 levels positive and 1024 negative, representing a resolution of about 0.1 percent of full-scale value. In order to study the effect of quantization, it is only necessary to think of the input-voltage waveform being applied directly to the step-transfer characteristic. A channel Δx is centered at each step of the transfer characteristic (Fig. 4). The coded value is the value of displacement.

First-Order Probability Density

As previously described, the probability that the fluctuating voltage falls into the incremental levels between x_i and $x_i + \Delta x$ is estimated by counting the number of occurrences N_i in the ith level or channel, during a run. The sum of all the successes is equal to the total number

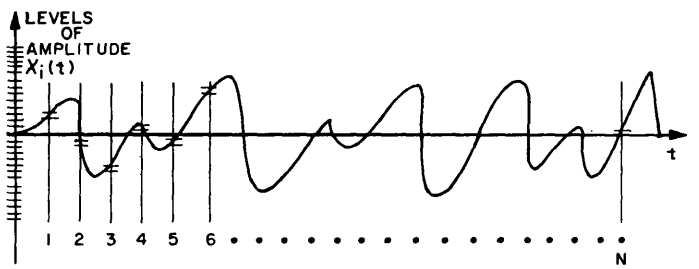
$$N_1 + N_2 + \cdots + N_i + \cdots + N_n = N \text{ for } n \text{ levels.}$$

Dividing N_i by N gives an estimate of the probability for that particular channel

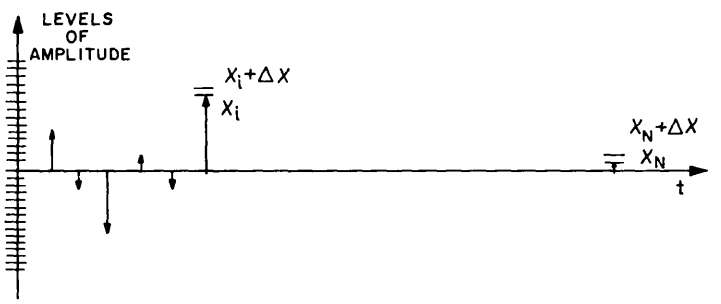
$$\frac{N_i}{N} \approx \text{Prob}(x_i \leq x < x_i + \Delta x) = P(x_i),$$

and assuming that the waveform will be contained within the full range of available levels, then

$$\frac{N_1}{N} + \cdots + \frac{N_i}{N} + \cdots + \frac{N_n}{N} = \sum_{i=1}^{N} P(x_i) = 1.$$



a. Sampling into a continuous random function of time



b. Sampled values from Fig. 3a

Fig. 3 - Periodic sampling of a random waveform

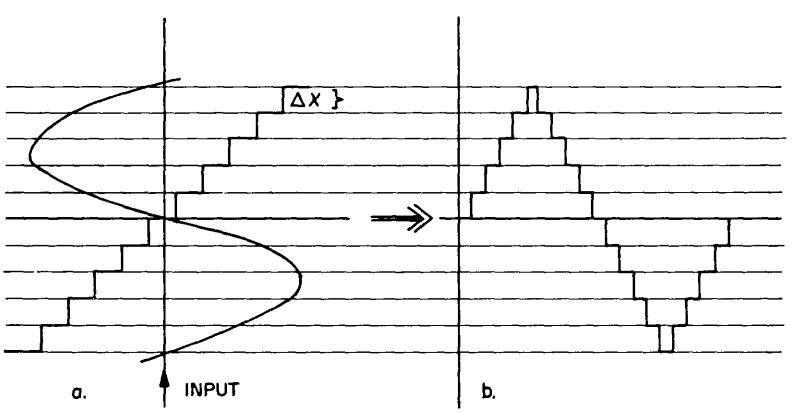


Fig. 4 - Effective transfer characteristic of an analog-to-digital converter. (a) Quantized transfer characteristic, and (b) Quantized output waveform

Each estimate is an approximation of the theoretical probability function describing the underlying statistical process such that

$$P(x_i) = \int_{x_i}^{x_i + \Delta x} p(x)\, dx$$

where $P(x_i)$ is the probability density and $p(x)\,dx$ is the probability that the voltage lies between x and $x + dx$. Thus the normalized measurement N_i/N is an estimate of the area under the probability-density curve bounded by a channel Δx wide at $x = x_i$ and hence is an estimate of the probability that the waveform falls in the ith channel, i.e., between x_i and $x_i + \Delta x$.

The closeness of the approximation depends upon the width of the channel Δx. If we divide by Δx and take the limit, as Δx goes to zero and N goes to infinity, the probability estimate $(N_i/N)/\Delta x$ becomes the slope at point x_i of the probability distribution function, or the value of the probability density $p(x_i)$ at x_i. Thus

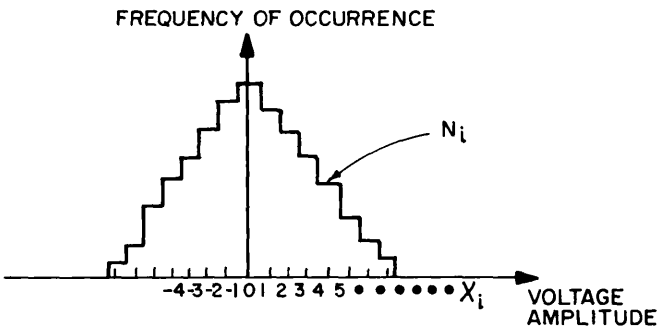
$$p(x_i) = \lim_{\substack{\Delta x \to 0 \\ N \to \infty}} \frac{(N_i/N)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\mathrm{Prob}(x_i \le x < x_i + \Delta x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_{x_i}^{x_i + \Delta x} p(x)\, dx .$$

The function $p(x)$ is the frequency function or the probability-density function, and the incremental value $p(x)\,dx$ the probability element. For a given sample size N there is a fluctuation associated with each channel estimate N_i/N. Increasing the sample size to infinity can be thought of as reducing the fluctuation to zero, in accordance with the law of large numbers. The number of levels to use for the probability estimate is determined somewhat by practical measurement considerations and is usually a compromise between the quantizing device and the amount of memory available.

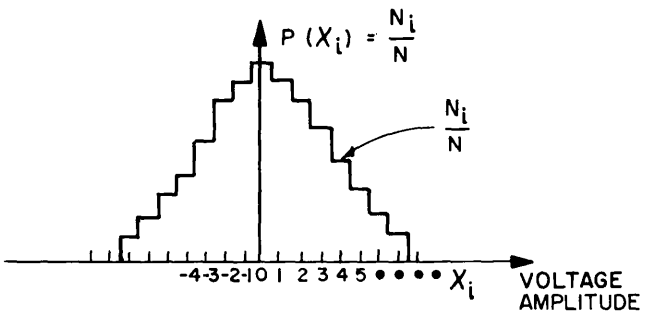
A plot of the frequency of occurrence for each level of amplitude is shown in Fig. 5a. The same data normalized is shown in Fig. 5b. When a sufficient number of total samples are quantized during a run, the measurement is stopped. The measurement consists of the numbers $N_1, N_2, \ldots N_n$ stored in memory. These numbers are read out of memory either for hand calculations or computer calculations of the normalized values N_i/N, and from the density various statistical averages may be computed from the data, e.g., the mean and the variance, standard deviation and comparison to a normal distribution density.

A calculation has been made using a theoretical probability-density function to show the error due to quantization. Assume that Gaussian noise has been sampled for an infinite interval, and assume that the quantizer has contiguous theoretically sharp channels over the full range, with each channel width equal to σ, where σ is the standard deviation. Figure 6 shows the normal curve plotted from a table with zero mean and unit standard deviation to represent the Gaussian noise-probability density of instantaneous amplitudes. This curve is divided into contiguous channels each σ in width, with the middle channel centered at $x = 0$. The area of each channel under the density curve is computed from a table of the cumulative normal distribution. This value is divided by the width of the channel, unity in this example, to get the height of the rectangle to plot on the density curve. Figure 7 shows the error is quantizing for a channel σ in width.

The error in variance due to quantization calculated for seven levels, using Shepherd's correction and disregarding all occurrences of X greater than $|3-1/2|\sigma$, is 0.007007, making the variance of quantized values 0.992993, compared to a theoretical variance of unity. If all occurrences greater than $|3-1/2|\sigma$ are grouped in the highest channels $|2-1/2|\sigma$ to $|3-1/2|\sigma$, then the error in variance is reduced to 0.003315. Thus



a. Frequency of occurrence vs x_i



b. Normalized frequency of occurrence, or estimate of probability density

Fig. 5 - Frequency of occurrence for each level of amplitude

Δx can be quite large and still afford a fairly accurate estimate of the probability for each channel.

Second-Order Probability Density

In addition to measuring the first probability density described above, the density analyzer is arranged to "tally" the joint occurrence of a pair of samples $\tau = t_i - t_{i-1}$ seconds apart. This is sometimes referred to as the second probability density. A given waveform with paired samples is shown in Fig. 8a. The corresponding representation of the sampled values is shown in Fig. 8b. Each value is held and quantized in an analog-to-digital converter. The quantized digital numbers are shown in Fig. 8c. The second sample of the previous joint pair is held as the first sample of the succeeding pair, for example, sample pairs N_1, N_2; N_2, N_3; N_3, N_4, ...; N_{n-1}, N_n. Tallying or sorting for the joint occurrences now becomes more complex. Sorting for the first distribution requires only a single line of numbers, whereas sorting for the joint density results in a two-dimensional array of numbers. As in the usual designation of an array or matrix of numbers, the value of the first sample of a pair can be characterized as selecting the row in a memory plane and the value of the second of the pair selecting the column. The memory cell at the coordinate numbers then is the tally number N_{ij} for the frequency of occurrence. The details of this referencing to memory are purposely being delayed until

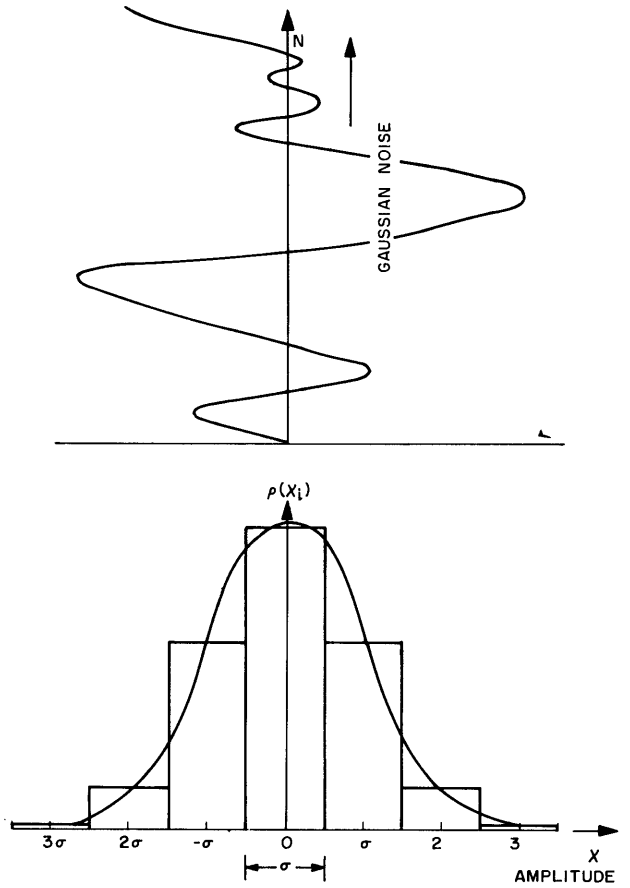


Fig. 6 - Example of theoretical Gaussian noise-probability density with quantizer channel width equal to standard deviation σ

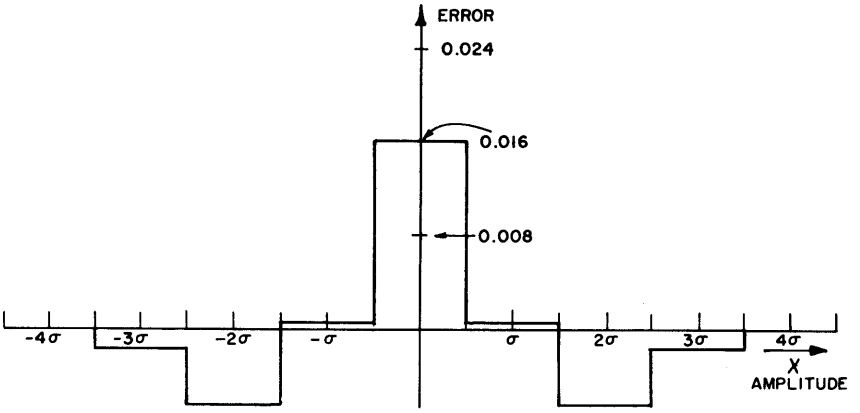
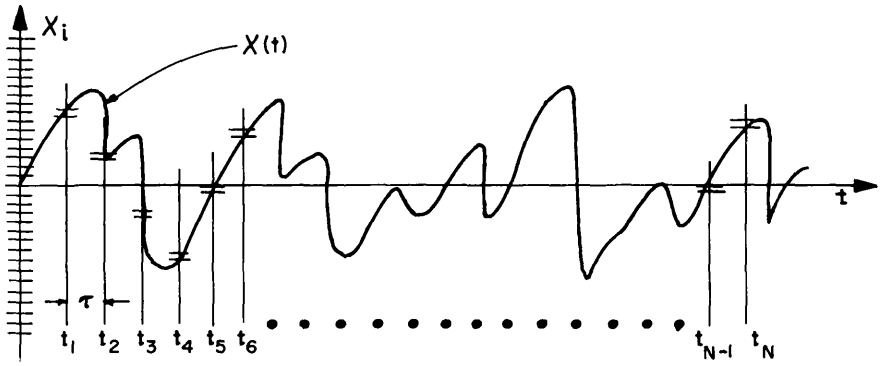
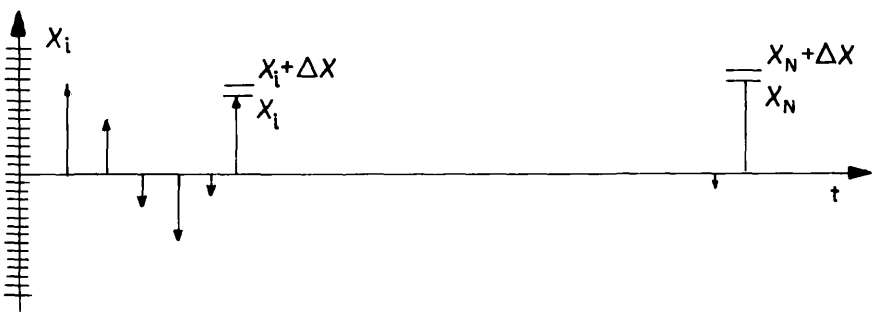


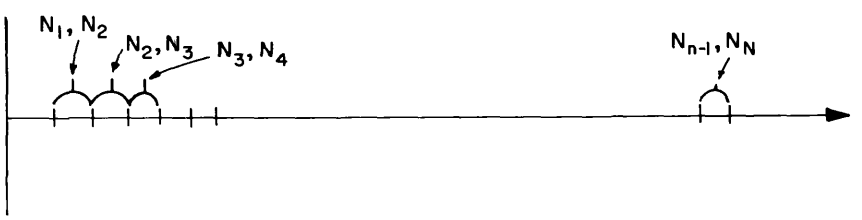
Fig. 7 - Error due to quantization



a. Waveform sampled at intervals τ



b. Sampled values from Fig. 8a



c. Sampled values of binary numbers paired from Fig. 8b

Fig. 8 - Joint occurrence of pairs of samples at intervals τ

later in the report. For the two-dimensional case, using the same rationale as in the first order, we get the following: $N_{ij}/N \approx p(x_i, x_j)\, dx_i\, dx_j$, where $p(x_i, x_j)$ is the second-order probability density of the amplitude.

GENERAL DESCRIPTION

As previously mentioned, the basic principle of this amplitude-density analyzer is to keep a tally of the occurrences of various values of amplitudes over a given period of time. This requires, as minimum basic items, a storage device or memory to hold the tally words, an address register to direct the values of amplitude to the storage device, and finally a counter to increment the number of times a given address has been called up.

The analyzer is comprised of five basic sections (Figs. 9, 10). The first section is the input; the second the core memory; the third the counter stage; the fourth the visual section; and the fifth section the controller.

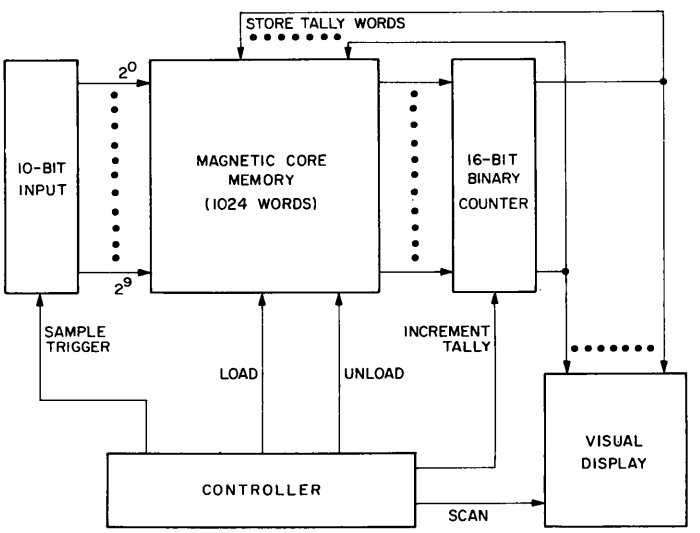


Fig. 9 - Amplitude-probability-density analyzer

The input section has a variety of stages and functions. There are three possible inputs that can be selected via a three-position rotary switch, the first of which is an analog-to-digital converter, whereby an analog signal can be fed in for real-time processing. The converter is a 10-kHz, ten-bits-plus-sign encoder which will quantize the instantaneous analog value to one part in 1024, representing about 0.1 percent resolution. The encoding is a ripple-down (serial) process, and new values are continuously generated upon receipt of a trigger pulse, until a stop pulse is initiated. Counting can be stopped at any time. For convenience a stop pulse is initiated whenever the contents of a particular address word has reached a value of $(2^{16} - 1)$.

The second input is a computer input that enables information stored in some computer to be tallied by the analyzer at a faster rate than could be achieved by software programming the computer. The advantage of using the analyzer as an auxiliary device to the computer is that the analyzer is a special fixed-wired device for tallying, whereas sorting and tallying in a large general-purpose computer is uneconomical.

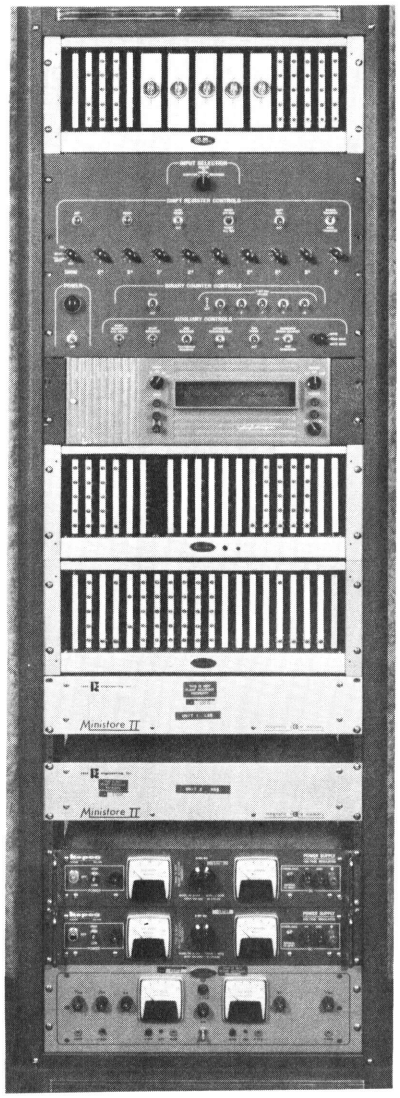


Fig. 10 - Front-panel view of amplitude-probability-density analyzer

The third input is from a 16-channel digital magnetic-tape recorder, enabling direct processing of data taken out in the field. Associated with the recorder is a "word selector" that demultiplexes the data words from the recorder in addition to selecting the "tau" steps. The recorder has a 16-bit word, but the actual data word is only 11 bits; the additional bits are used for record keeping when using the tape unit directly with a computer. The advantage of reading directly from a digital recorder is readily apparent when one has an abundance of raw field data for processing.

The three inputs when individually selected via input gates will feed the address portion of a random-access core memory. The core memory is 1024 words, with a word length of 16 bits. Since the address is directed by a data word, the data word must be limited to the ten most significant bits (MSB) so not to exceed the number of possible address locations. The sorting operation proceeds as follows. Initially all memory is cleared to zero, then a data word sets up the address via one of three possible inputs. Instead of storing the data word in memory in the conventional manner, the data word is used as an address word to call up a number located at that address in memory. The number in memory is read out into an external counter, and incremented by one digit; then the incremented number is read back into memory at the same location. It is to be noted that by this method the actual data word is lost after the classifying operation, and the number in memory is a record of the actual number of times that particular word has occurred during a run. A "run" consists of continuously sorting for some fixed length of time or, alternatively, sorting is set to stop when a "tally" number in memory reaches a maximum value, i.e., $(2^{16} - 1)$. The consecutive numbers in memory are the relative frequency of occurrence of the data words. The counter is the fourth stage of this system and is 16 bits in length, to match the 16-bit word length of the core memory. The counter increments the number of times an address is called from memory. The additional features of the system, such as the visual display, timing, and control, along with some manual and automatic features, will be discussed more thoroughly in the following section.

The fifth and final stage of the analyzer is the controller. The controller is all solid state and consists of a series of one-shot multivibrators forming a closed loop. Each stage performs a specific function and regulates the total timing of the analyzer.

DETAILED DESCRIPTION

Input Section

The first input stage is an analog-to-digital converter, commercially built by Adage, Inc. Use of a converter allows real-time processing of an analog signal, providing the frequency of the signal is within the range of the converter.

The analog-to-digital (A/D) converter that is used for the analyzer is a bipolar device requiring about 88 microseconds to convert ten bits plus sign and a parity bit, and it is used with a sample-and-hold circuit. It also functions as a bipolar digital-to-analog (D/A) converter, with the analog output being amplified by an operational amplifier. The operation of the A/D converter is a serial approximation of adding resistors to the circuit until the sum of the currents balances at the summation point feeding the comparator.

The time between decisions is approximately eight microseconds, which corresponds to the "interval clock" period. The total digital equivalent value of the balancing current is stored in a series of flip-flops giving the quantized value, in straight binary, of absolute value and sign. Each conversion is completely independent of the preceding one, and the minimum time between conversions is 20 microseconds. This interval gives the converter a conversion rate of almost 10,000 complete conversions per second. The output from the converter is gated in parallel into the random-access core memory via input gates and a ten-bit shift register.

The second input stage is for a Bendix G-15 computer. The computer has a word capacity of approximately 2176 as a result of having 20 lines with a capacity of 108 words per line and four short lines of four words each, plus a few special operational registers.

The density analyzer appears to have a word capacity far in excess of most computers, because the words are used as tallies. Although its core storage capacity is 1024 words, each word is a tally of 1024 words of 16 bits each. Since the actual data word is not stored in the memory, the analyzer has the ability to hold a maximum tally of $(2^{16} - 1)$ (1024), or 67,108,864 total events. This maximum tally could be realized only if the input signal was a perfect ramp wave over the full range of the converter. There are no known computers with this type of internal storage space exclusive of magnetic-tape units, where actual words are stored. Modern computers with fast input capability can take away numbers in a "tally" operation, but this arrangement ties up a much more expensive machine.

Using a computer with the analyzer requires a "fill-and-dump" operation when the computer is filled with data for other mathematical calculations, then dumped into the analyzer or stored on magnetic tape. This "fill-and-dump" operation could be done for an entire set of data for a density analysis, and the analyzer would then give the ensemble of the amplitude-probability distribution for the entire set.

The third input is for a digital magnetic-tape recorder, and the recorder used is a Potter M-906-II variable-speed, 16-channel, dual-head recorder with manual and remote control. This recorder is run with a packing density of 267 bits per inch, or 8 kHz, at a speed of 30 inches per second (ips).

The tape recorder has the raw field data that has been multiplexed onto the tape, in time-division multiplexing, and contains information from eight different analog sources (9). The recorder input on the analyzer allows direct processing from the recorder after it has been demultiplexed. This arrangement will then show what the distribution of the recorded data looks like before selection is made for computer processing.

The analyzer has the tallying capacity to read an entire reel of tape from the recorder, sorting each data-word tally into a maximum of 1024 amplitude "slots." A 2400-ft reel of tape runs for 16 minutes at a speed of 30 ips, and the data have been packed on at a maximum rate of 267 words per inch, giving a total of 7,689,600 words to be processed.

When using the magnetic-tape recorder as an input there is a five-bit tau selector (Figs. 11 and 12) for adjusting the tau steps.* This is a five-bit binary counter feeding a

*For a complete list of symbols appearing in the text and figures, see Appendix A.

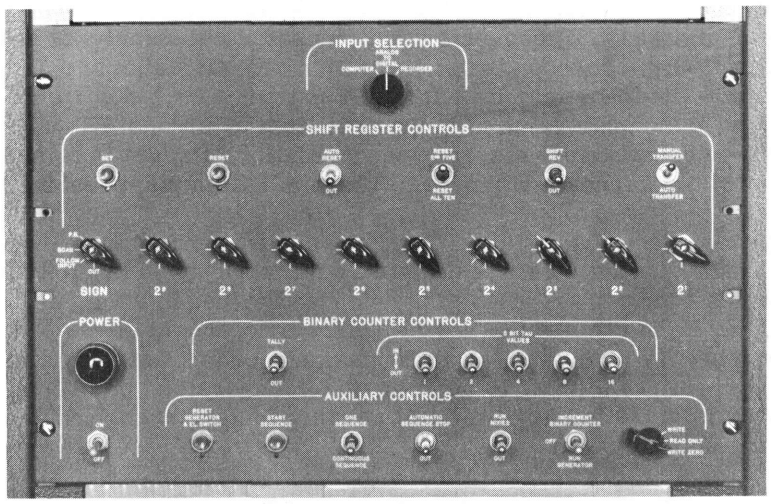


Fig. 11 - Control panel for density analyzer

five-input "and" gate. It reads the clock pulses from each recorded data word and, depending on the selection of the five toggle switches, will read only the appropriate data, such as every word, every second word, every third word, etc., up to a maximum selection of ever 31st word.

The input section also consists of a three-position rotary switch (Figs. 11 and 12) for selection of a given type input (i.e., A/D converter, computer, or tape unit) and with the aid of the input gates will feed the data in parallel from one of these sources into a ten-bit shift register.

The input lines to the shift register are directed to the shift register through ten switches. This system allows one to vary the number of address selections, manually set or reset an address, or cycle the address register for scanning out data from the core memory. The shift register takes the ten most significant bits of each data word, and these bits become the address for the ten-bit address register of the core memory. The shift register can set up second-order probability distribution by utilizing the five most significant bits from the inputs and shifting them down with the aid of the pulse generator (Fig. 12) into its least most significant five-bit section. The five bits of the next data word will complete the ten-bit address configuration, giving the second-order frequency of occurrence or bivariant distribution.

When the core memory has been sufficiently filled, readout of the stored information is accomplished by an address generator to scan the core properly (Fig. 13). The address generator is a recirculating reversible binary counter used with appropriate delay units. This readout arrangement is necessitated by the coding configuration of the A/D converter, and since the data are always derived by the same converter for the computer and magnetic-tape unit, the above readout technique is always the same.

This readout technique permits a display of decoded data on an oscilloscope. The core memory must be continuously scanned over all its addresses to get the correct display. If it were scanned from address zero to 1024 in sequence, the amplitude density would not appear symmetrical about the vertical axis (Fig. 14). This result is due to the coding on the negative half of the analog signal. The A/D converter is a bipolar device, and on the positive half of its coding the numbers increase in a straight binary fashion up to ten bits. When on the negative side the sign bit comes on, giving a binary indication of

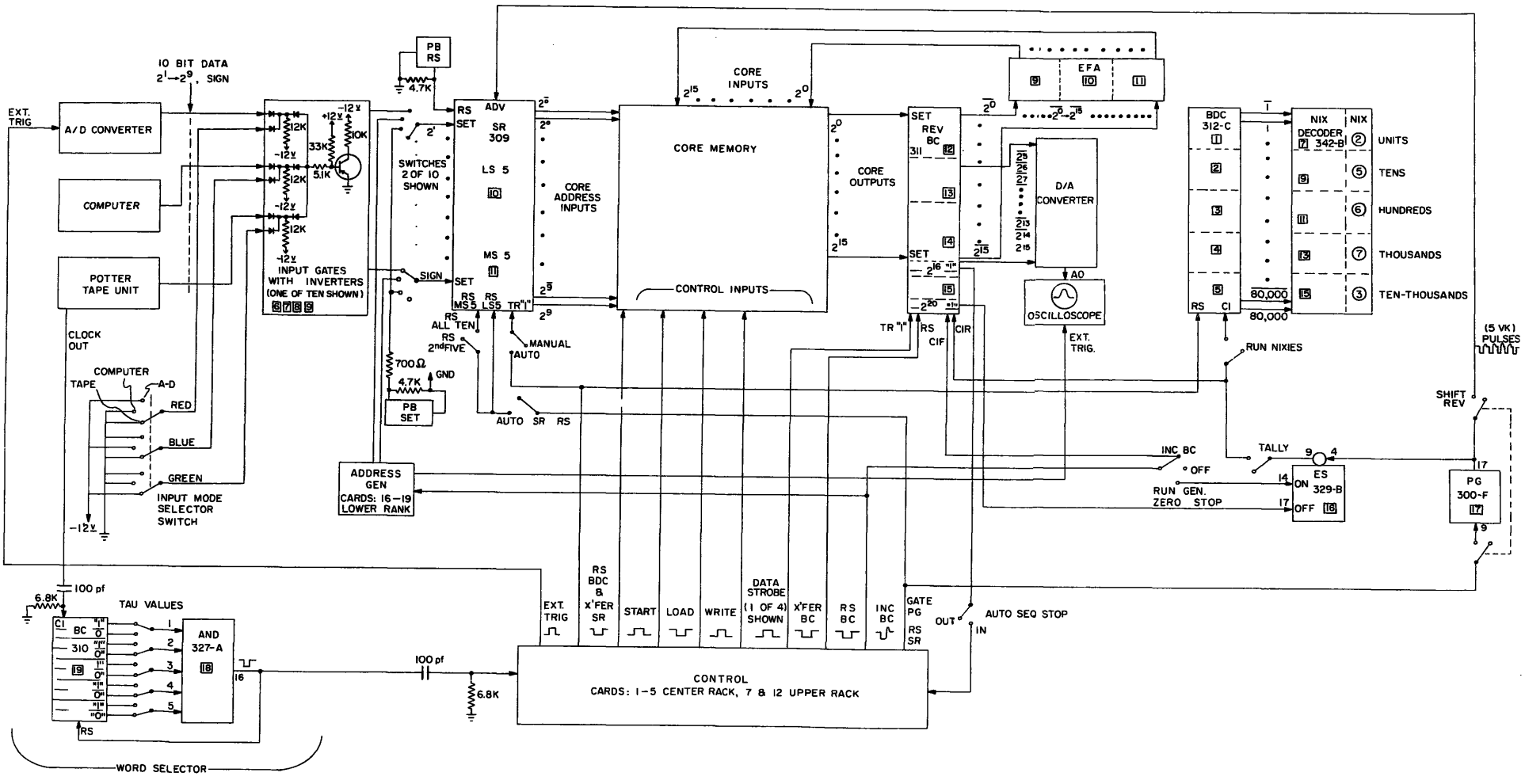


Fig. 12 - Block diagram of amplitude-probability-density analyzer

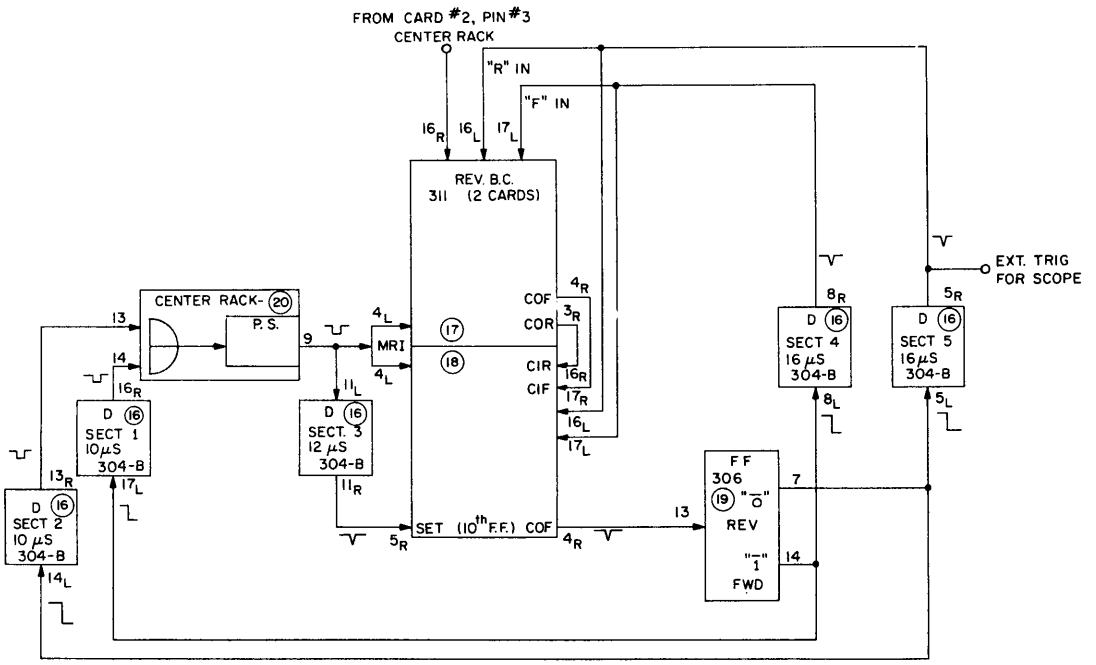
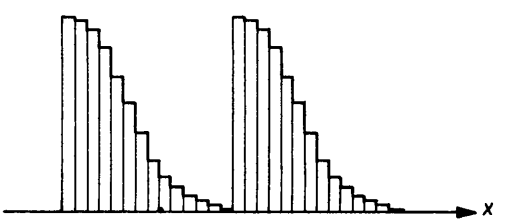


Fig. 13 - Block diagram of address generator for scanning core memory for readout

Fig. 14 - Splitting of display due to improper readout sequence



512 for negative zero, the sign bit stays on for all negative values, making the binary count start at 512 and increase to 1023 for full-scale negative. With this arrangement of coding it is more readily apparent that one cannot scan the addresses of the core memory and get the proper configuration. It is necessary to count from 511 in reverse to 0 and count forward from 512 to 1023, because the peak of the curve will be at address zero and 512. This counting technique is the reason for the special address generator for visual readout purposes (Fig. 13).

Core Memory

The second basic stage of the analyzer is the core memory (Fig. 12). The memory used in this system is a Rese Ministore II, Model 3708, random-access core memory, having a capacity of 1024 words eight bits in length. Since we wanted our tally to reach a maximum count of $(2^{16} - 1)$ it was necessary to employ two of these units in parallel. We had found empirically that a maximum tally of $(2^{16} - 1)$ gave good coverage over many periods of data and generated a smooth density curve.

The Model 3708 can retrieve information and subsequently regenerate the retrieved data, or it can clear an address location and subsequently write new data. The units include a magnetic-core array, sense amplifiers, digit drivers, input/output data registers, read and write current generators, access line current switches, power regulators, and an address register.

The address-selection scheme utilizes a coincident-current mode of word selection, and the bits of each word read out or written into the memory are in parallel form. The address format is pure binary, making it compatible with the A/D converter. A unique address location for all binary numbers from 0 to 1023 is provided. The cycle is such that a write operation will always follow a read operation. When an address is given to the memory, and a command is given to interrogate that address, the memory will read the specified address, leaving the location in the cleared state. The readout is immediately followed by a write current, which restores the information into the address from which it was taken. The total cycle time for the read-write operation is approximately 30 microseconds, making the memory speed slightly better than 33 kHz.

The function of the core memory, as was stated previously, is to tally the number of times a given address configuration appears, rather than storing the actual data words. The numbers stored in memory are a record of the actual number of times the data words appear during a given run and when read out will give a histogram of the frequency of occurrence of the data words. The data words are quantized values of the analog amplitude in a ten-bit binary configuration, and the frequency of such data words when read out and normalized by the total count N would then be an amplitude-probability distribution.

Binary Counter

The incrementing stage of the analyzer is the binary counter which makes up the third major section of the system. The counter is 16 stages long, to match the maximum word length of 16 bits.

The counter in this system is an external device and has the function of incrementing the number of times a given address is called up. It must be reiterated that the actual data words are lost after the tallying operation, and that the numbers in memory are a record of the number of times those particular data words have occurred during a run.

The counter used in this system is a reversible counter made from NAVCOR 300 series logic. The value of each address is incremented by one, and the new value is returned to that address until some particular address location reaches a value of $(2^{16} - 1)$ and causes the analyzer to stop its code sorting.

The outputs of the binary counter feed in parallel its ten most significant bits to a D/A converter (Fig. 12) which feeds the quantized steps to an oscilloscope for a visual representation. In addition to feeding the D/A converter, the counter also feeds all 16 of its output stages to the core memory via emitter followers, which serve to buffer the output loading of the counter.

Visual Section

The visual section of the density analyzer consists of two parts. One gives a visual representation of the quantized analog values as seen by a cathode-ray scope (Fig. 12). The second visual representation (Fig. 12) is the nixie readout lights, which give the individual quantized values in a five-digit decimal configuration.

The first visual display uses a ladder network for the D/A conversion. The ten most significant bits (MSB) are transferred in parallel to the 16-bit binary counter and to the ladder network, and the output becomes serially quantized steps of the binary code. For display purposes nothing is lost, as far as the visual configuration is concerned, by dropping the six least significant bits (LSB) from the ladder network. The total display is accomplished by repeatedly cycling the addresses of the core memory, as explained in a previous section, and reading out the address values to the ladder network. Although the address values are read to the ladder network via the binary counter, the counter does not serve any function during this process other than as a temporary storage device. The final display is shown on an oscilloscope, where oscillograms may be made of the distribution curve.

The nixie readout lights were employed so that one may get the respective values of the quantized steps in decimal. This process is basically manual, where an address is selected and its value is read out into the nixie decoder via the binary counter.

The sequence goes as follows. By putting the analyzer in the manual mode, an address is selected with the use of the ten input toggle switches and shift register. The value of that address is pulled from the core memory and dumped into the 16-bit binary counter. The counter is then run down in reverse, by a pulse generator, and runs up the binary decimal counter (BCD), which feeds the nixie decoders. The resulting binary number is now presented on the nixie light display in decimal. All phases of this operation are automatically initiated after one has manually selected the address to be read. This same procedure is employed for all addresses where the actual decimal number is desired.

Controller

The controller is the fifth and final major section of the analyzer and was made from commercially available solid-state logic circuits by NAVCOR, Inc., with a few modifications (Fig. 15). The controller is a series of one-shot multivibrators forming a continuous ring. As a pulse travels through each stage, a particular function is performed, as shown by the various control lines.

The lengths and voltage levels of the control pulses have been set up to satisfy the requirements of the buffer unit as shown in Fig. 12, where the voltage levels are 0 v and -6 v. The only other timing consideration was that of the A/D converter, which required 110 microseconds between conversions.

The pulse circulates continuously around the closed loop in the control unit, performing the various operations repeatedly until a stop pulse initiates a sequence stop. The stop pulse is derived from the binary counter by sensing on the 17th stage and generating a stop pulse when a tally has reached that value. The toggle switches for automatic read/write or manual selections are all located on the control panel (Fig. 11).

The control panel works with the digital controller. It is in effect the programmer for the system and allows one to select the type of input, whether to read or write, and the type of visual scanning one would desire.

Before a "read-in" operation, it is necessary to clear core memory. A clear operation goes as follows (Fig. 11). First, turn the power switch on, then select the type of input desired; assume in this case that the input comes from the A/D converter. Set the input rotary switch to the A/D converter.

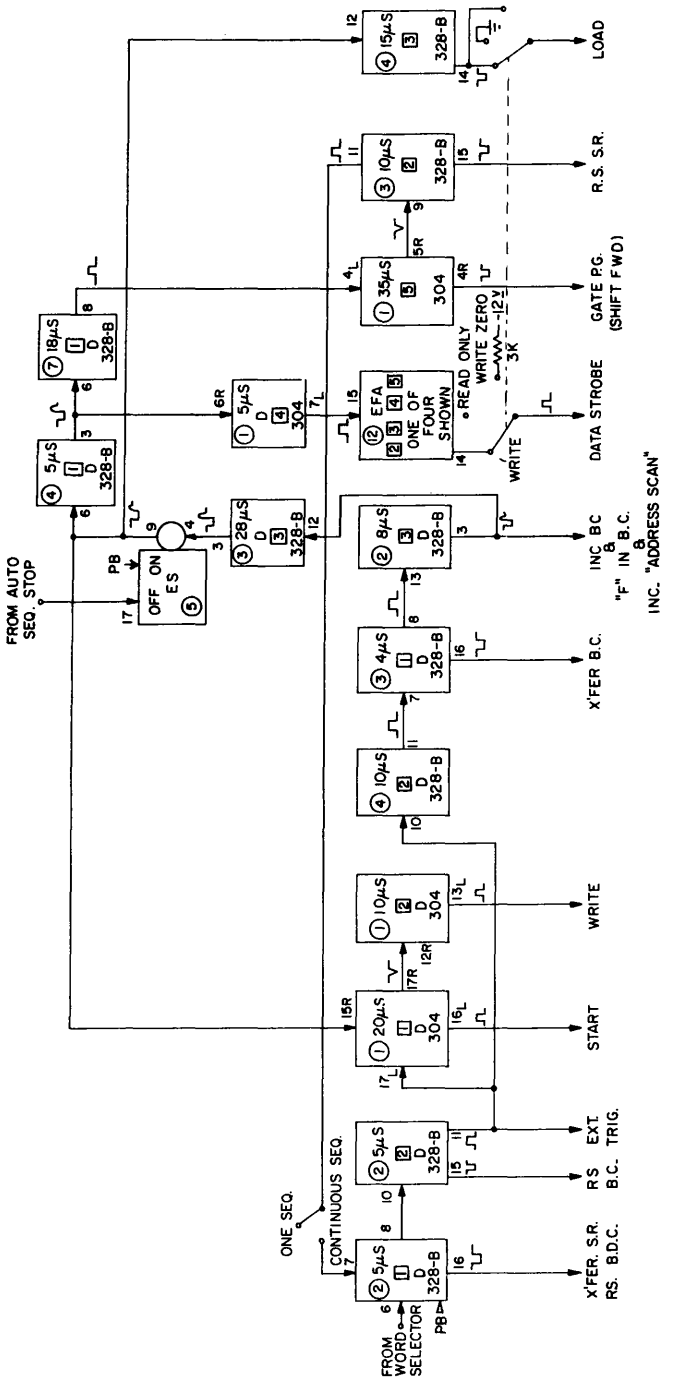


Fig. 15 - Block diagram of controller for amplitude-probability-density analyzer

Set the following switches on the control panel (Fig. 11):

A. Under shift register controls

1. Auto reset in
2. Reset all ten
3. Shift reverse out and manual transfer in automatic position
4. Place all ten rotary switches to the scan position

B. Binary counter control section is neglected in this operation

C. In the auxiliary control section

1. Push reset generator and electronic switch button
2. Place toggle switches in continuous sequence
3. Automatic sequence stop
4. Omit run nixies
5. Increment binary counter in the off position and place rotary switch in write zero mode
6. Push start sequence button, and this will clear core memory

After the system has cycled for a few seconds, place the sequence switch in the one sequence position; now the system is ready to accept data from the A/D input, with the following changes.

Under the shift-register controls, place all ten or any number thereof in the "follow input" position, and those inputs that are desired to be omitted should be placed in the "out" position. The number of inputs selected will determine the number of addresses that will be used for code sorting. Binary counter controls still remain unchanged, for they do not enter into this phase of operation. In the auxiliary controls section, set the rotary switch (extreme right-hand side) to the write position; set the toggle switch to increment binary counter position; place the sequence switch in continuous position, and push the start sequence button.

The analyzer will now sort analog data in real time with a sampling rate equal to or greater than the Nyquist rate, depending on the frequency range. This system will sort at a maximum rate of 8 kHz due to the limitations of the A/D converter, and the rate may be changed by adjusting the trim pots on the delay stages that feed the increment binary counter control (Fig. 11) or the one feeding the electronic switch.

The system will stop automatically when a particular address location has reached a value of $2^{16} - 1$, and the number of addresses that will contain data will be a function of the number of input switches that were used (i.e., number of addresses sorted equals 2^N, where N represents number of input lines used, up to a maximum of 10).

The numbers now contained in the core memory are the number of times those addresses have been selected by the data words, and are not the actual data words. To get an overall visual representation of data distribution, one can set up the proper controls on the control panel in the following manner:

A. Under auxiliary controls

1. Put toggle switch in "read only" position
2. Set increment binary counter to "off"
3. Set sequence switch to "continuous"

B. In the shift register control section

1. Set all switches that had been used in the "scan" position
2. Leave others in the "out" position*
3. Push "start sequence" button and see visual display on oscilloscope (Fig. 11)

When it is desired to stop the scanning process, place the sequence switch in the "one sequence" position. This action breaks the circulating pulse in the delay units and causes a program interrupt.

To get a visual numerical representation of the address values on the nixie display lights, the following procedure can be used providing all addresses have been utilized in the sorting operation.

Set the following switches on the control panel (Fig. 11):

A. Under Shift Register Controls

1. Place switch in "auto transfer"
2. Shift reverse switch "out"
3. Reset switch to "reset all ten"
4. Auto reset switch to "auto reset" position
5. Place all input switches to "scan" position

B. Under Binary Counter Controls, place tally switch to "tally" position

C. Under Auxiliary Controls

1. Place rotary switch to "read only"
2. Increment binary counter switch to "run generator"
3. Nixie switch to "run nixies"
4. Automatic sequence switch to "out" position
5. Sequence switch to "one sequence"
6. Push reset generator switch

*Push reset generator button. Due to the action of the automatic sequence stop, the electronic switch (Fig. 12) has been cut off, disabling any further continuous sequencing until it has been turned on again.

By then pushing the start-sequence switch, the contents of each address will be displayed in decimal form on the nixie lights. The address location will be displayed on the "address scan" binary counter neon lights (Fig. 13). It must be noted that the number displayed on the nixies will be the contents of the address, increased by two, as displayed on the "address scan" binary counter.

The advantage of this procedure is that all addresses are automatically cycled through with the aid of the address-scan generator, and all one need do, after setting up the proper switches, as previously explained, is to push the start-sequence button for a new address location. Numbers displayed on the nixies are the value of the address, decreased by one.

There is one other readout mode that is more manual than previously mentioned, and it enables one to select an arbitrary address and read out the contents of that address. To operate in this mode only, the following changes must be made.

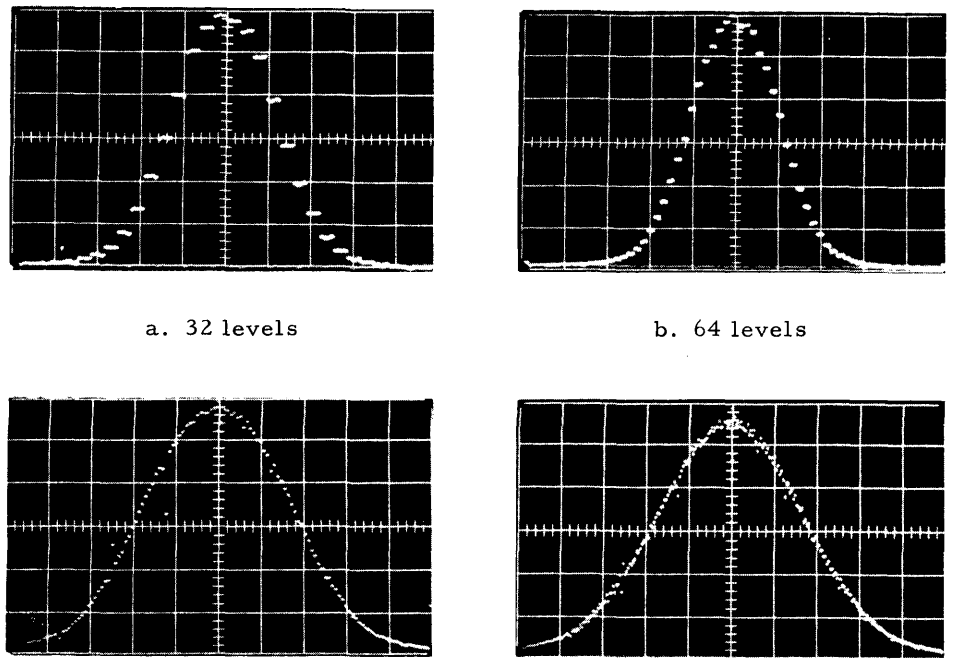
1. Place manual switch in "manual transfer" mode
2. Place both shift-reverse and auto-reset switches in the "out" mode
3. Place rotary switches in PB (pushbutton) or out position to set up the address configuration desired in binary
4. The PB position will give 1's, and out position will generate 0's
5. Push "reset" button to clear shift registers of previous information
6. Push "set" button for new address configuration
7. Push "reset generator"
8. Push "start sequence" buttons, and the number displayed on the nixies will be the value contained at the address as shown on the shift registers and rotary input switch configuration

This manual mode of operation allows one to go to any given address desired and read out its contents, and could serve as a random check on the analyzer.

RESULTS

The system allows us to do real-time amplitude-probability distributions with various analog signals. We have been able to get the distribution of thermal noise, pseudorandom noise, wideband or band-limited noise, and the distribution of sine waves.

Before the oscillograms of a few selected amplitude-probability distributions are considered, it may be instructive to explain the meaning of levels as used in this report. A ten-bit A/D converter has 1024 unique levels to resolve an analog signal into its binary representation. For each of these binary levels, representing a value of voltage amplitude, we assign a corresponding address in a 1024-word core memory. The memory will have a slot for each possible value of signal amplitude, or it can be said that there are 1024 levels of code sorting. If fewer levels should be used in code sorting than that afforded by the A/D converter, then resolution may be sacrificed in the sorting process. This results in a m-fold degeneracy of levels at each address, where $m = 2^{10-n}$, and n = number of bits used in sorting. Referring to Fig. 16b, where a six-bit address was used in the sorting process, the degeneracy $m = 2^{10-6} = 2^4$, or 16 levels per address. In Fig. 17b a ten-bit address was used, giving no degeneracy, but the curve is less smooth



a. 32 levels

b. 64 levels

c. 128 levels

d. 256 levels

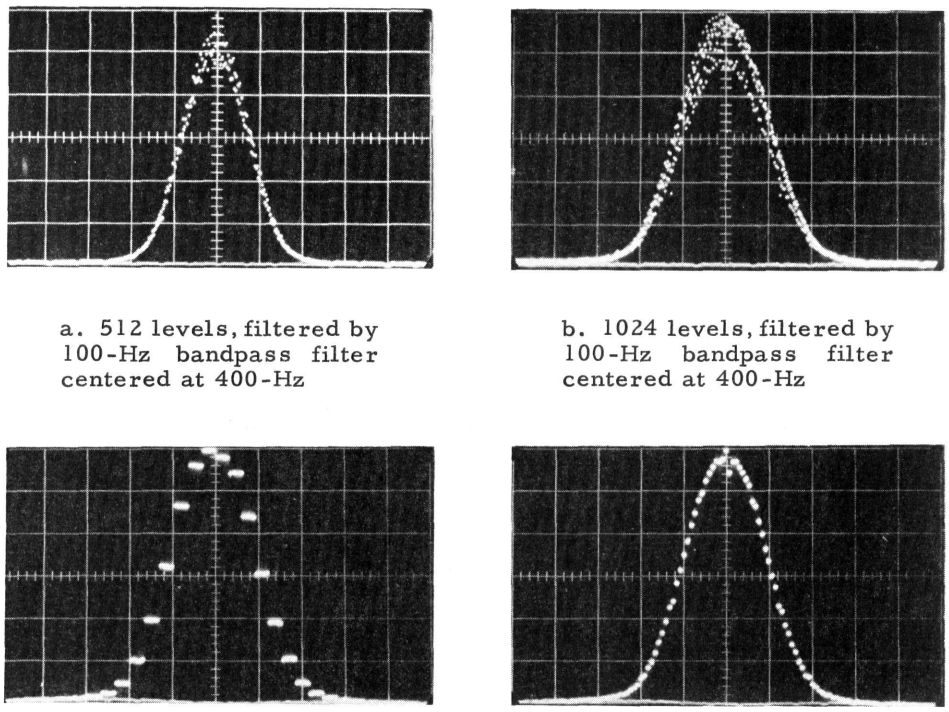
Fig. 16 - Levels of code sorting for thermal noise, filtered by 100-Hz bandpass filter centered at 400 Hz

than in the six-bit case. Note, also, that the curves are not symmetrical about the vertical axis. The A/D converter that was used in this analyzer had its negative zero code suppressed, creating the nonsymmetry about the vertical axis of the oscillograms.

Figures 16 and 17 show the effects of various code sortings in obtaining the first-order amplitude-probability distribution of the amplitude-probability distribution of thermal noise, filtered by a 100-Hz bandpass filter (BPF) centered at 400 Hz; also included in Fig. 17 are two oscillograms for wideband thermal noise, for code sortings of 32 and 128 levels respectively. Figure 18 is the first-order amplitude-probability distribution for a 400-Hz sine wave, with 256 and 512 levels of code sorting.

The first-order amplitude distribution was also made for pseudorandom noise of various sequence lengths, as indicated in Fig. 19. The pseudorandom noise (PN) was generated by a variable-length linear shift register (SR) with a period (P) equal to $2^n - 1$, where n = number of shift register bits used. Any sequence achieving a period equal to $2^n - 1$ is a maximum-length linear shift-register sequence. The output of the SR was fed through a 100-Hz BPF, centered at 400 Hz. The oscillograms show what type of first-order distribution that can be obtained for SR lengths of 9, 11, 17, and 19 bits respectively, using 64 levels of code sorting in all cases.

The density analyzer has facilities for measuring second-order amplitude-probability distributions, and Fig. 20 shows three oscillograms of second-order distributions for BPF thermal noise, wideband thermal noise, and PN of sequence length $(2^{17} - 1)$. To better understand what has taken place on the second-order amplitude-probability distributions, attention is called to Fig. 21, showing how the sampled pairs are formed. The time function $x(t)$ is sampled at some fixed sampled rate, and the 5 MSB of each sample



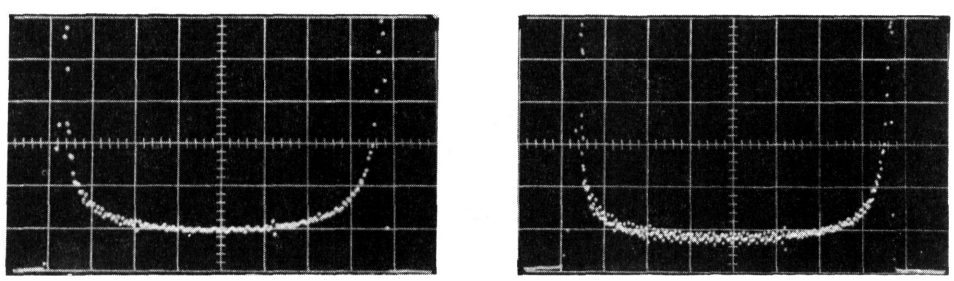
a. 512 levels, filtered by 100-Hz bandpass filter centered at 400-Hz

b. 1024 levels, filtered by 100-Hz bandpass filter centered at 400-Hz

c. 32 levels, for wideband 20-kHz thermal noise

d. 128 levels, for wideband 20-kHz thermal noise

Fig. 17 - Levels of code sorting for thermal noise



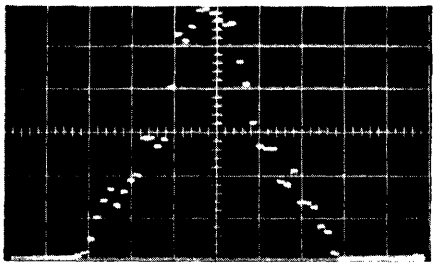
a. 256 levels

b. 512 levels

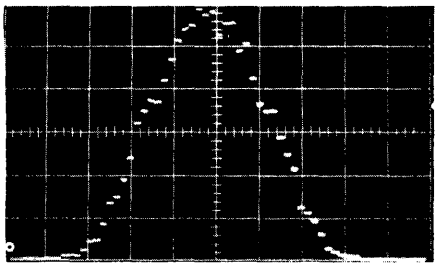
Fig. 18 - Levels of code sorting for a 400-Hz sine wave

pair x_1 and x_2 are combined to form a ten-bit word. This new ten-bit word becomes the address selection of the core memory.

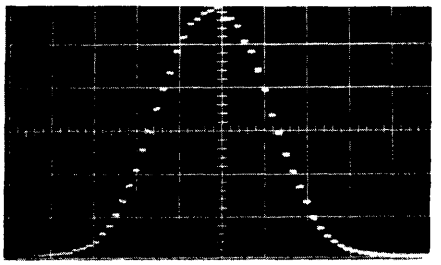
The core memory becomes a 32 by 32 matrix of the sampled pairs x_i and x_j. The tally operation is now the frequency of occurrence of the joint amplitude values of x_i and x_j and when the tallying has been completed these slots then approximate the joint probability, $P(x_i, x_j)$. Since a joint probability is a three-dimensional display, it cannot be depicted on a two-dimensional oscillogram. Therefore, the curves displayed in Fig. 20



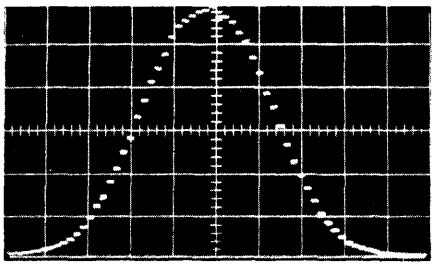
a. SR length 9 bits



b. SR length 11 bits

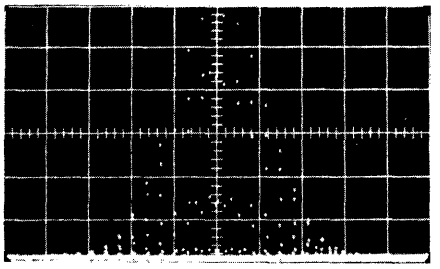


c. SR length 17 bits

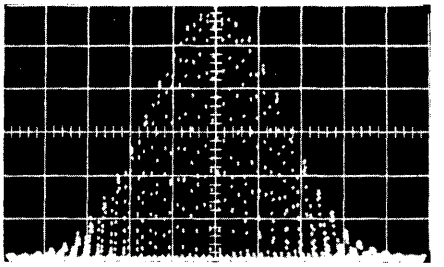


d. SR length 19 bits

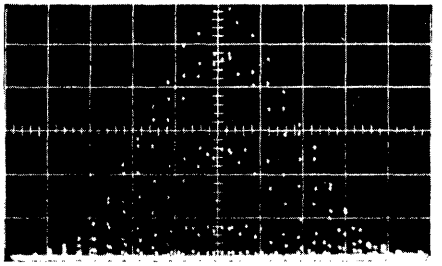
Fig. 19 - PN sequence with 64 levels of code sorting



a. Thermal noise, filtered by 100-Hz bandpass filter centered at 400 Hz



b. Wideband thermal noise



c. PN sequence, SR length 17 bits

Fig. 20 - Second-order amplitude densities

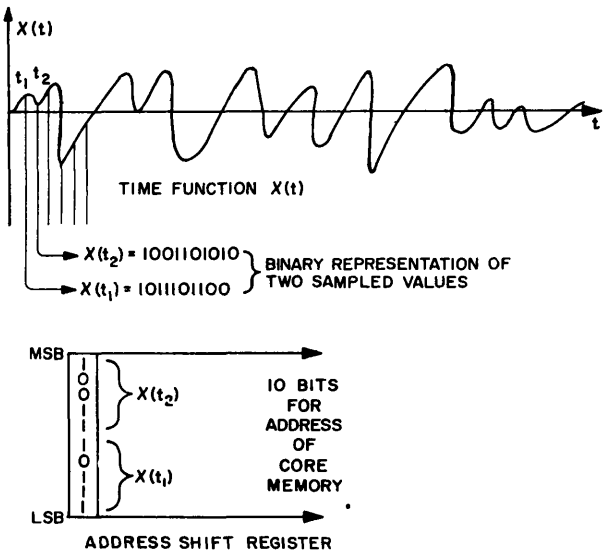


Fig. 21 - Sampled pairs being formed to get one address word in the second-order density scheme

are a series of nonoverlapping marginal probabilities of $p(x_i)$, and each $p(x_i)$ is obtained by summing over the j's for a fixed value of i, as given by equation:

$$p(x_i) = \sum_j p(x_i, x_j).$$

Our matrix contains 32 curves of $p(x_i)$, with 32 points associated with each curve.

The sampling rate for the analyzer is roughly 8 kHz, and data are band limited with a center frequency of 400 Hz and bandwidth of 100 Hz. This means that the sampling rate is almost 20 times the highest frequency component of the data. The result of this fast sampling rate means that successive sampled pairs will be highly correlated in value, since the time function can't make any extreme changes during the small time interval between samples. Joint occurrences of x_i and x_j are then rather close to the diagonal of $x_i = x_j$ in the matrix, as shown by the shaded region in Fig. 22. This fact explains in part why the curves for the band-limited thermal noise and PN (Fig. 20a and 20c) have fewer points in their center than the wideband thermal noise. In the band-limited case there will be certain unpermissible combinations of x_i and x_j; thus an extreme positive value of x_i can't be followed by an extreme negative value of x_j. This would imply that the band-limited signal could make an extreme excursion within a sampling interval.

In the case of the wideband thermal noise, the sampling rate is much less than the high-frequency components of signal. Joint pairing of x_i and x_j being highly uncorrelated as well as statistically independent, can take any range of values, within A/D converter limitations; therefore the matrix of x_i and x_j will have most or all of its slots used. This will cause a greater fill-in of points when the second-order density is displayed as series of marginal probabilities (Fig. 20b).

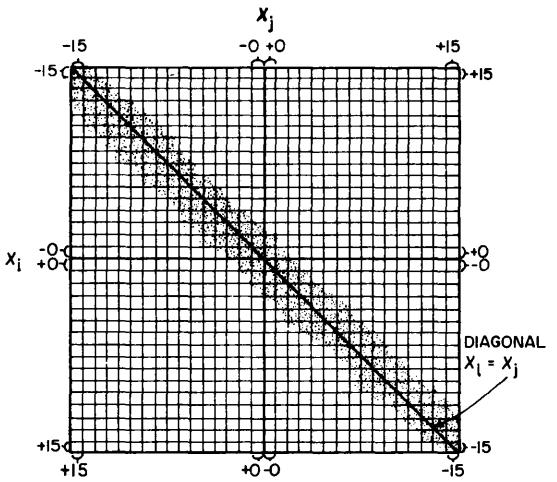


Fig. 22 - Core memory with 1024 bits, which becomes a 32 by 32 matrix for second-order density

FUTURE PLANS

With a direct hook-up, the contents of the analyzer can be transferred directly to the Computer Control Corp. (CCC) DDP-24 digital computer. Data transfer to a Bendix G-15 Computer has already been implemented. Since the G-15 has a drum memory rather than a core memory like the DDP-24, it required minimum interfacing, namely a shift register to get the transfer of data. With this arrangement, a statistical examination of the amplitude frequency of occurrences, namely, the mean, variance, and best possible fit to a theoretical Gaussian curve, may be obtained. This, of course, will allow a more thorough calculation of the cross-correlation or covariance of field data, since one of the major assumptions made on the data is that it is Gaussian and has a zero mean and a unit variance. The amplitude-distribution analyzer may give a better insight into the actual behavior of the data before small segments of it are pulled out for computer processing.

This technique of code sorting provides an excellent system for checking out digital tape recorders of magnetic tapes for dropouts. When used with a computer, the tally of the address occurrences enables one to run a reel of magnetic tape through the analyzer for computer dumping, and with a second pass of the same reel a comparison can be made of the tallies for various addresses. If there is a single dropout, a tally will change by one count and another tally would increase by one count. By noting where the increase and decrease have taken place, one could ascertain which bit actually dropped out. Since every data word selects an address, any variations in data words due to dropouts, either real or transient, would be noted in the final count, because <u>all</u> data words are counted for the number of times they occurred during a run.

A check on the frequency stability could also be made on an analog magnetic tape recorder using the analyzer to check on "wow" and "flutter" characteristics. This check could be implemented by feeding a very stable, crystal-controlled sine wave to the tape recorder, along with a trigger pulse also derived from the same crystal. The trigger pulse would be adjusted to sample at the peak and zero crossing of the sine wave. The setup would require an analog recorder with at least two channels, so one could contain the sine wave and the other the trigger pulse. When recording these waves on the analog recorder, one would feed the analyzer with the quantized values of the sine wave simultaneously to get the signature of the wave. Ideally the analyzer would have data at only

three addresses—one for the positive peak, one for the negative peaks, and another for zero crossings. These values would then be read out and noted, and after reading in from the analog recorder, a comparison could be made between the two sets of values. The variations of values obtained at the peaks would be an indication of amplitude variation, or "flutter," and the variations of the values obtained at the zero crossing would be indicative of frequency variations, or "wow." By making a series of runs one could see what the long-term or short-term stability over fixed time periods look like for a particular analog recorder.

This system can also serve as an excellent check on code reliability of an analog-to-digital converter. By feeding such a device a ramp wave that runs synchronously with the A/D trigger pulses, one can get a good look at the coding of the converter against another or against itself over a number of trials. This would even allow one to see if there are variations in code width for a given A/D converter, since a linear ramp wave covering full scale would cycle all the codes the same number of times.

CONCLUSIONS

This distribution analyzer serves as a low-cost, special-purpose computer capable of sorting digital data in real time at a speed comparable to that of a large-scale, high-cost computer. It is mobile and can be taken out in the field for on-line processing, or it can be used with other computer facilities to check on the first- and second-order amplitude-probability densities of raw data before the data are processed.

It also serves as a useful device for code checking A/D converters or making code comparisons between various converters. The analyzer may also check on dropouts of digital magnetic tape systems, or it can check the long-term frequency stability of analog tape recorders.

The same basic principles of this analyzer may be incorporated with a computer to get a broader range of the data distribution. An A/D converter with a greater dynamic range than that afforded by ten bits plus sign would directly address the computer, whereby each time an address is called up a tally goes to that address location. The number of addresses used would be a function of the A/D converter and would equal 2^N, where N is the number of bits used in the converter, including the sign bit. The magnitude of a given tally would be a function of the word length of the computer and would equal $2^K - 1$ where K is the number of bits in a computer word.

This technique would allow on-line processing with a digital computer (i.e., amplitude probability distribution). The readout could be programmed for type out or visual readout in some external device. The versatility of this analyzer is increased considerably with a computer acting as the storage and incrementing device. Although use with a computer may seem attractive, this does not lead to a portable system, and, in addition, it may unnecessarily tie up the use of an expensive computer.

REFERENCES

1. Peterson, H.L., and Finney, W.J., "Ultrasonic Noise Studies (Report of Instrumentation of Amplitude Analysis, MIT)," Report of NRL Progress, November 1950, page 86, Confidential

2. Peterson, H.L., and Finney, W.J., "Ultrasonic Noise Studies (Report of some distribution curves of data analysis—MIT)," Report of NRL Progress, December 1950, pages 64-65, Confidential

3. Peterson, H.L., and Finney, W.J., "Basic Sonar Electronic Research (Probability Amplitude Distribution Analyzer)," Report of NRL Progress, January 1958, pages 53-54

4. McCoy, C., Jr., Hammond, D.L., and Peterson, H.L., "Digital Amplitude Probability Distribution Analyzer," Report of NRL Progress, November 1960, pages 39-41

5. Peterson, H.L., "Probability Distribution Analysis of Sonar Echoes," paper given at the 6th Navy Symposium on Underwater Acoustics, held at the U.S. Navy Electronics Laboratory, San Diego, California, November 17, 18, 1952

6. Lieberman, G., Peterson, H.L., and Finney, W.J., "A Study of Noise Amplitude Distributions for Sonar Transmission," paper presented at the 11th U.S. Navy Symposium on Underwater Acoustics held at Pennsylvania State University, June 28-29, 1955

7. Davenport, W.B., Jr., "A Study of Speech Probability Distributions," Technical Report No. 148, Research Laboratory of Electronics, MIT, August 1950

8. Van Rennes, A.B., "Pulse-Amplitude Analysis in Nuclear Research," Nucleonics, Vol. 10, July 1952 through October 1952

9. McCoy, C., Jr., and Peterson, H.L., "A Magnetic Digital Recording System for Field Use," NRL Report 6016, November 15, 1963

* * *

APPENDIX A

LEGEND

PB	Push Button
RS	Reset
ADV	Advance (shift forward)
LS	Least Significant
MS	Most Significant
REV	Reversible
BC	Binary Counter
SR	Shift Register
CIF	Count In Forward
CIR	Count In Reverse
EFA	Emitter Follower Amplifier
BDC	Binary-coded Decimal Counter
NIX	Nixies
ES	Electronic Switch
PG	Pulse Generator
INC	Increment
XFER	Transfer
D	Delay
PS	Pulse Shaper
MRI	Master Reset In
FF	Flip-Flop
COF	Count Out Forward
COR	Count Out Reverse
FWD	Forward
N	Pin number
(N)	Section on card
[N]	Slot number in rack

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Naval Research Laboratory Washington, D. C. 20390	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP –

3. REPORT TITLE

A SYSTEM FOR MEASURING THE FIRST AND SECOND ORDER PROBABILITY DISTRIBUTIONS OF SIGNALS AND NOISE

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

An interim report on the problem; work is continuing.

5. AUTHOR(S) *(First name, middle initial, last name)*

Caldwell McCoy, Jr. and Herbert L. Peterson

6. REPORT DATE September 11, 1967	7a. TOTAL NO. OF PAGES 34	7b. NO. OF REFS 9
8a. CONTRACT OR GRANT NO. NRL Problem S01-06 b. PROJECT NO. RF 101-03-44-4054 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) NRL Report 6522 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*	

10. DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Department of the Navy (Office of Naval Research) Washington, D. C.

13. ABSTRACT

A system has been designed and developed for use, either in the laboratory or field, to measure the first- and second-order amplitude-probability densities of signals and noise in the area of underwater acoustics. The system consists of a wide-band dc amplifier feeding a sample-and-hold network, followed by an analog-to-digital converter to encode the sampled instantaneous value to an amplitude resolution of ten-bits and sign. These digital codes become the address of a magnetic-core memory, and the value stored in a given address location is a tally for each time that address has been called up. Since the address was directed by the instantaneous amplitude value of a time function, the stored tallies are proportional to the amplitude density of that given time function when plotted out as a histogram. These measured values can be read into a general-purpose digital computer as desired to compute the normalized density function together with moments.

This method of code sorting can be applied to other problems of instrumentation, e.g., calibration checks of analog-to-digital converters, stability of analog magnetic tape recorders, and "drop-out" measurements in analog magnetic tape recording.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6801

Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Probability						
Signals						
Noise						
Probability densities						
Distribution analyzer						

DD FORM 1 NOV 65 1473 (BACK)
(PAGE 2)