324-77.

SR

OR 3,573,612

[11] 3,573,612

[72] Inventor **Robert Walter Alister Scarr**
Stansted, England

[21] Appl. No. **770,690**
[22] Filed **Oct. 25, 1968**
[45] Patented **Apr. 6, 1971**
[73] Assignee **Standard Telephones and Cables Limited**
London, England
[32] Priority **Nov. 16, 1967**
[33] **Great Britain**
[31] **52,146/67**

[54] APPARATUS FOR ANALYZING COMPLEX WAVEFORMS CONTAINING PITCH SYNCHRONOUS INFORMATION

5 Claims, 2 Drawing Figs.

[52] U.S. Cl. 324/77, 179/1
[51] Int. Cl. G01r 23/16
[50] Field of Search ..... 324/77 (D), (H), 77 (G), 78 (E); 179/1 (AS); 328/151, 130; 307/235

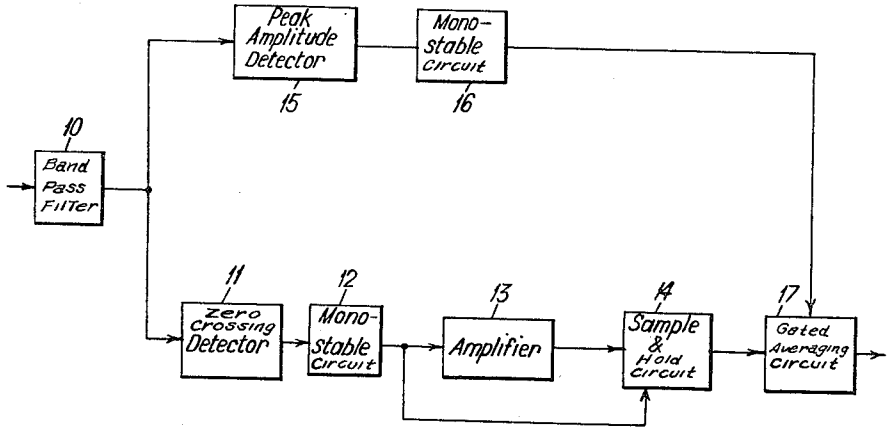
[56] **References Cited**

UNITED STATES PATENTS

3,335,225 8/1967 Campanella et al. 179/1
3,416,080 12/1968 Wright 324/77

Primary Examiner—Edward E. Kubasiewicz
Attorneys—C. Cornell Remsen, Jr., Walter J. Baum, Percy P. Lantzy, Philip M. Bolton, Isidore Togut and Charles L. Johnson, Jr.

ABSTRACT: A complex waveform containing pitch synchronous information is applied in parallel to a peak detector and a zero crossing detector. A first monostable is activated by the output signal of the zero crossing detector to provide a first pulse of fixed duration. A sawtooth waveform is activated by the trailing edge of the first pulse. The sawtooth waveform will have an amplitude linearly proportional to time. The leading edge of the next sawtooth waveform will activate a sample and hold circuit to sample and store the sawtooth waveform. A second monostable is activated by the output signal of the peak detector to provide a second pulse of a duration less than the fixed duration. The second pulse switches on (gates) an averaging circuit which is coupled to average the output signal of the sample and hold circuit for the duration of the second pulse. This results in an output signal which is an average of the times between successive zero crossings of the complex waveform within the duration of the second pulse.



Patented April 6, 1971 3,573,612

2 Sheets-Sheet 1

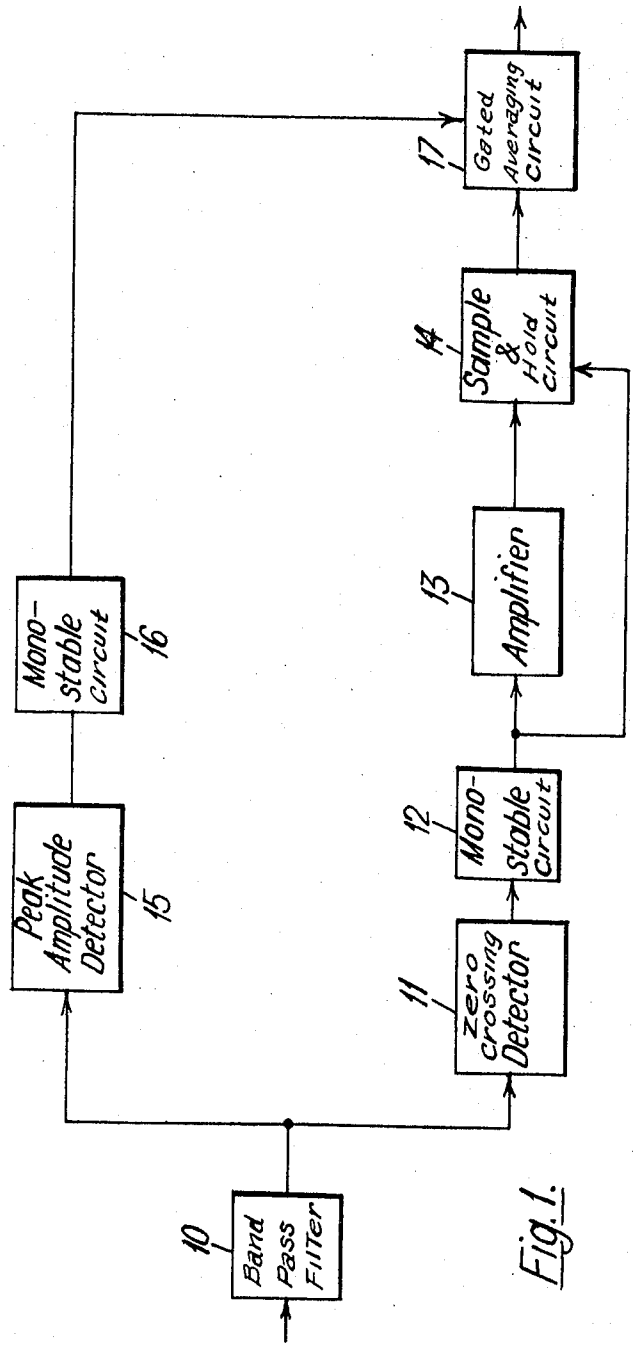


Fig. 1.

Inventor
ROBERT W. A. SCARR
By Philip M. Bolton
Attorney

Patented April 6, 1971 3,573,612

2 Sheets-Sheet 2

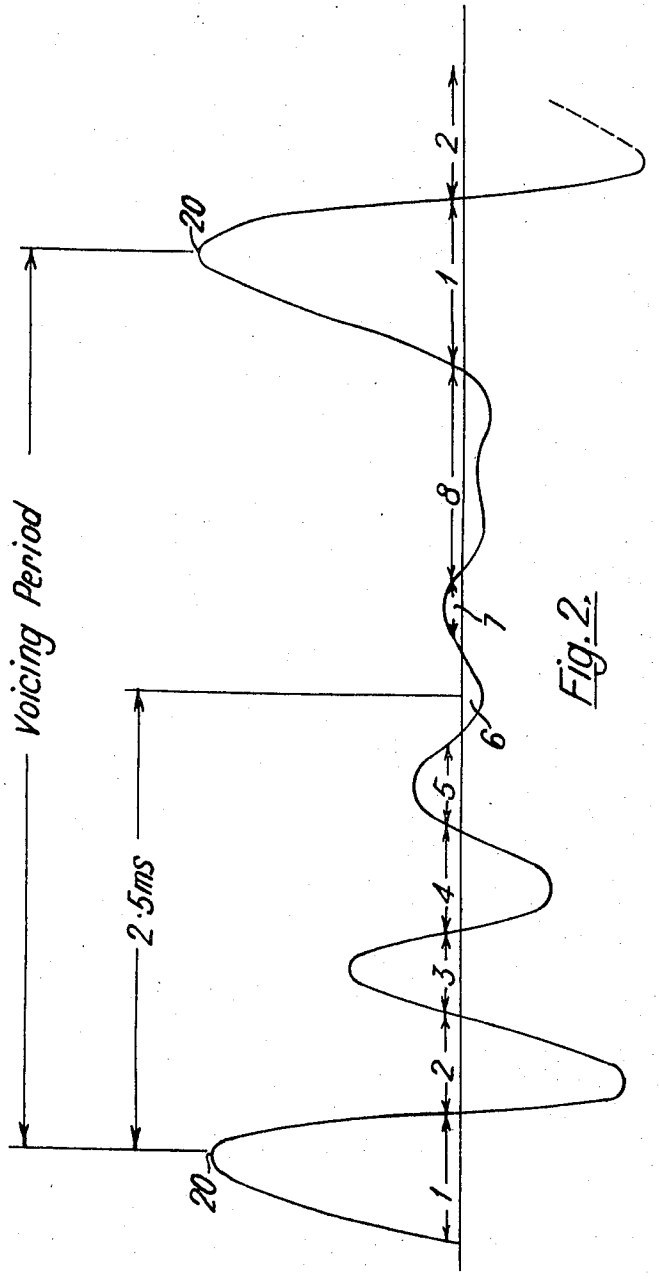


Fig. 2.

Inventor
ROBERT W. A. SCARR
By Philip M Bolton
Attorney

APPARATUS FOR ANALYZING COMPLEX WAVEFORMS CONTAINING PITCH SYNCHRONOUS INFORMATION

This invention relates to apparatus for analyzing complex waveforms containing pitch synchronous information, such as speech waveforms.

A steady state waveform, even a complex one, lends itself to mathematical treatment. For example, the performance of the vocal tract for a steady state vowel sound can be expressed in precise mathematical terms. In particular, the behavior of a single resonance, or formant, can be expressed in a relatively simple manner. The exciting source, the larynx, vibrates at a fairly constant frequency and produces harmonics. These harmonics are arbitrarily related to the formant frequency and, depending on the difference between the formant frequency and the nearest harmonic, a different wave shape and zero crossing pattern will result for which is presumably heard as having the same vowel quality. If the harmonic series representing the larynx source can also be expressed in mathematical terms, then it is a relatively simple, if somewhat laborious, matter to calculate the zero crossing pattern.

Conventional instruments for special analysis depend largely on having a repetitive waveform giving a frequency-line structure. The mathematical basis for this is the Fourier series. For single events a Fourier integral transform is a useful tool (though the Laplace transform is often preferred) and an event can be transformed from the time domain to the frequency domain and vice versa. The Fourier integral transform and the Laplace transform are useful in going from the frequency domain to the time domain in order to obtain, for example, the transient response of a specified circuit. Frequency analysis of single events in the time domain on the other hand is not always easy to instrument or useful when done.

Speech falls between the single event and the repetitive waveform. Voiced speech may be regarded as quasi-repetitive. In continuous speech where as much, or more, information is conveyed by formant transitions as by the steady state value of the formants, it is obviously important to have a good dynamic measure of vowel quality.

Assuming a Sona-graph type of presentation is accurate, formant transitions last for periods of the order of 40—100 msec. and can involve frequency changes of the second formant in excess of 1,000 c/sec. (P.S.GREEN "Consonant-Vowel Transitions, a Spectrographic Study" Acta Linguistica 12, (1958) pp. 57—105). dF/dt can therefore exceed 10,000 c/sec. per sec. Or to put it another way, the resonant frequency of the second formant can change by 100 c/sec. in one (male) voicing cycle.

The basic problem is to take a source (the larynx) assumed to be constant in frequency and with a known harmonic structure and to compute either the frequency spectrum or time output as a function of the moving resonant frequencies of the vocal tract. Each harmonic can be treated separately and the results summed.

If the discrete harmonic structure of the larynx vibration can be neglected, then a formant transition will appear to the recipient as a frequency modulated wave with an "instantaneous frequency" related to the resonant frequency of the formant. "Instantaneous freuqncy" is a contradiction in terms because frequency cannot be measured in an instant. Nevertheless, in the theory of frequency modulation (S. Goldman "Frequency Analysis, Modulation and Noise" McGraw Hill (1948), pp. 146—154) a useful definition of instantaneous frequency can be given when the carrier frequency is high compared with the modulation frequency. In the present case this is not necessarily true because the ratio of the fundamental of the carrier to the modulation frequency can be as low as 10 to 1. But the ratio of the formant frequency to the modulation frequency is higher and it will be assumed that the usual definition of instantaneous frequency is still useful, i.e.,

$$\text{Instantaneous frequency} = \frac{1}{2\pi}\frac{d\theta}{dt}$$

where the frequency modulated signal is expressed as

$$a = A \sin \theta$$

To calculate the spectrum of a frequency modulated wave is straight forward when the modulating frequency is repetitive. The result might be called the 'long term spectrum' of the frequency modulated wave. However, in the case of speech, the modulating wave is not repetitive and it is the 'short term' situation that is of interest. The short term situation is the best expressed by the instantaneous frequency as this is presumably related closely to the actual format resonances which are in turn related to the position of the articulators.

Thus some measure of instantaneous frequency should be less ambiguous than a presentation of the Sona-graph type which requires a justification that the peaks in the energy/frequency plot correspond to the instantaneous position of the formant and requires relatively complex instrumentation to 'locate' the peaks.

In copending application, Ser. No. 684,903, filed Nov. 17, 1967 there is described apparatus for analyzing a complex waveform containing pitch synchronous information in which a peak amplitude point in the voicing cycle is used to establish a phase reference to which zero crossing may be referred, and extracting one zero crossing distance at the pitch synchronous rate from the voicing cycle. Such apparatus requires a certain complexity in the provision of counter and gating circuits to extract the value of the one relevant distance, particularly at high frequencies, for example at frequencies above 800 c/sec.

According to the present invention there is provided apparatus for analyzing a complex waveform containing pitch synchronous information including means for detecting the first occurrence of a selected pitch synchronous characteristic of the waveform, means for gating the waveform for a predetermined period of time after the detection of the selected characteristic, means for integrating the periods of time between successive occurrences of the characteristic within the gating period and means for averaging the times between the successive occurrences of the characteristics within the gating period.

The above-mentioned and other features of the invention and the manner of attaining them will become more apparent and the invention itself will be best understood by reference to the following description of an embodiment of the invention, taken in conjunction with the accompanying drawings, in which:

FIG. **1** is a block schematic of an analyzing apparatus; and

FIG. **2** illustrates a typical waveform containing pitch synchronous information.

In the arrangement of FIG. **1** the input waveform is s separated into frequency ranges by a number of band-pass filters, one such filter **10** being shown. The filter output is applied to a zero crossing detector **11**, which is a limiting amplifier. Following the zero crossing detector **11** there is a monostable **12** which provides a pulse of fixed duration, said 30 microseconds. Every time the detector **11** gives an output, i.e., every time the input to the detector goes through zero, the trailing edge of the monostable pulse starts an integrating amplifier **13** which provides a sawtooth waveform having an amplitude linearly proportional to time. The leading edge of the next monostable pulse operates a sample and hold circuit **14** which causes the output of the amplifier **13** to be sampled and held in circuit **14**. The trailing edge of the monostable pulse resets the integrating amplifier and starts the next cycle of operation.

Also connected to the output of the band-pass filter **10** is a peak amplitude detector **15** which provides a pulse once per voicing cycle at the peak in that cycle. The peak amplitude detector output goes to a second monostable **16** which gives a pulse of fixed duration, say 2.5 milliseconds. For the duration of this pulse an averaging circuit **17** is switched on and averages the output of the sample and hold circuit **14** for that period.

Considering the waveform shown in FIG. **2**, one complete voicing period is deemed to start at point **20**, where a peak in the voicing cycle occurs. The sample and hold circuit will pro-

vide outputs proportional to the zero crossing distances marked 1—5 in the FIG. In the sixth zero crossing distance the 2.5 millisecond period terminates and so only distances 1—5 are averaged by the averaging circuit, the remaining zero crossing distances in the voicing cycle being ignored. The output of the averaging circuit is therefore the average of distances 1—5. The remaining distances are associated with the low-amplitude part of the waveform and are not wanted in the averaging process because they contain little relevant information and very often include relatively high-noise components, thus presenting completely misleading information.

It is clear from FIG. 2 that the circuit works best when the number of zero crossing distances involved in the averaging process is relatively large and therefore the range of frequencies concerned is large compared with the fundamental voicing period. This process is therefore preferred for use on the higher frequency ranges. In practice it is most suitable for use on frequencies above 800 c/sec. At frequencies below this, i.e., 200—800 c/sec. the number of zero crossings in, say 2.5 milliseconds would be too few to arrive at a reasonable average.

It is to be understood that the foregoing description of specific examples of this invention is made by way of example only and is not to be considered as a limitation on its scope.

I claim:

1. Apparatus for analyzing a complex waveform containing pitch synchronous information including:

an input means for said waveform;

first means coupled to said input for detecting the first occurrence of a selected pitch synchronous characteristic of said waveform;

second means coupled to said input for integrating the periods of time between adjacent zero crossings of said waveform between successive occurrences of said selected characteristic; and

third means coupled to said second means and said first means for averaging said periods of time during a given duration less than the duration between successive occurrences of said selected characteristic.

2. Apparatus for analyzing a complex waveform containing pitch synchronous information including:

a band-pass filter through which said waveform is passed; and

a zero crossing detector coupled to said filter to which said filter waveform is applied;

a peak amplitude detector coupled to said filter to which said filtered waveform is applied;

timing means coupled to said peak amplitude detector responsive to an output of the peak amplitude detector to produce a timing pulse having a given duration less than the duration between successive outputs of said peak amplitude detector; and

averaging means coupled to said zero crossing detector and said timing means to average the periods of time between successive outputs of said zero crossing detector for a duration equal to said given duration.

3. Apparatus according to claim 2 in which said averaging means includes;

a first monostable circuit coupled to said zero crossing detector responsive to an output from said zero crossing detector;

an integrating amplifier coupled to said first monostable circuit responsive to an output of one significance from said first monostable circuit whereby said amplifier produces a sawtooth waveform;

a sample and hold circuit coupled to said amplifier to sample said sawtooth waveform and hold the sampled value, said sample and hold circuit being responsive to an output of another significance from said first monostable circuit; and

an averaging circuit coupled to said sample and hold circuit to average the output of said sample and hold circuit for said given duration of said timing pulse from said timing means.

4. Apparatus according to claim 3 in which said timing means is a second monostable circuit responsive to an output of said peak amplitude detector.

5. Apparatus according to claim 4 in which said zero crossing detector is a limiting amplifier.