

March 26, 1963

M. V. MATHEWS ET AL

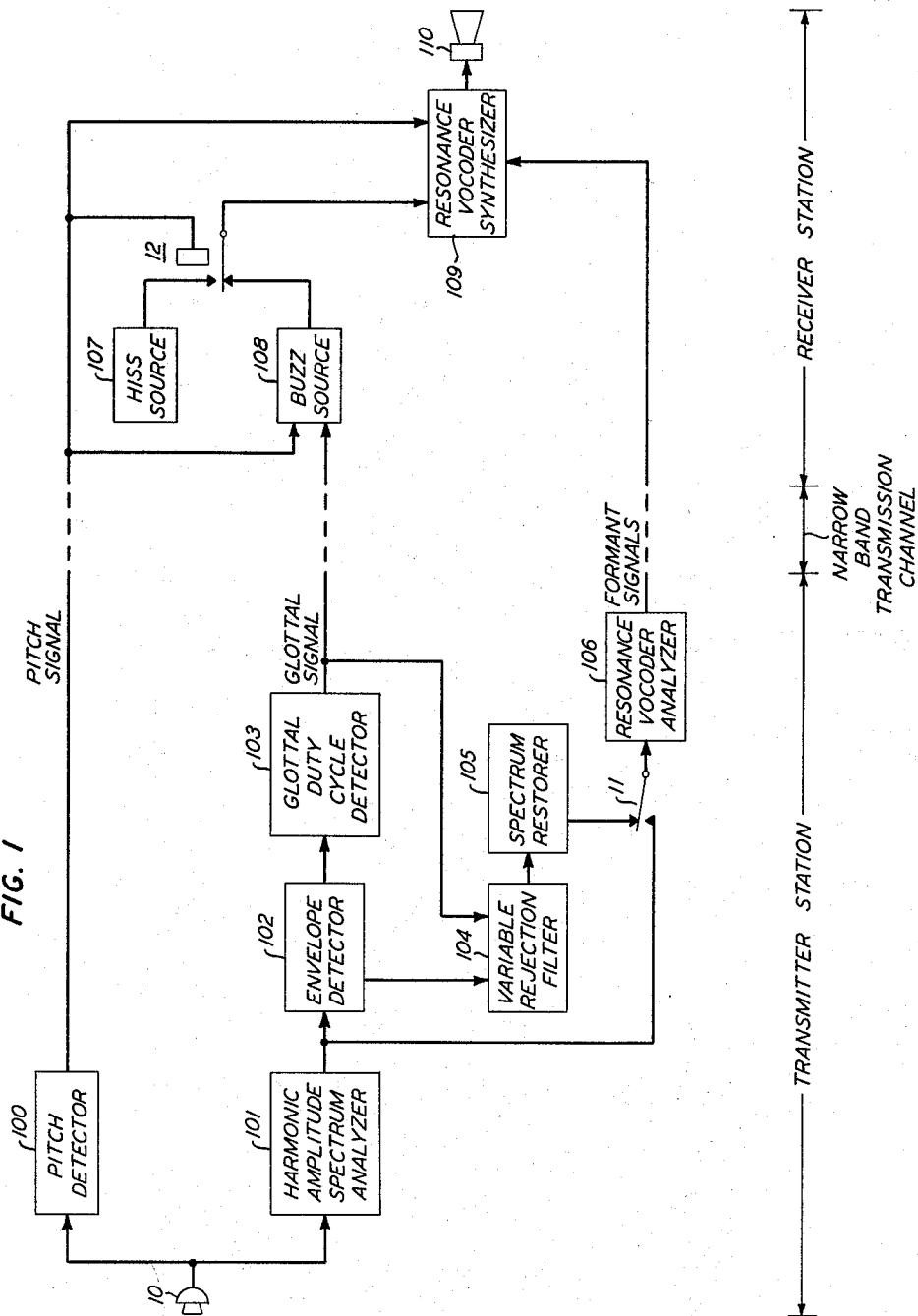
3,083,266

VOCODER APPARATUS

Filed Feb. 28, 1961

6 Sheets-Sheet 1

FIG. 1



INVENTORS
BY
M. V. MATHEWS
J. E. MILLER
G. E. Hirsch Jr.
ATTORNEY

March 26, 1963

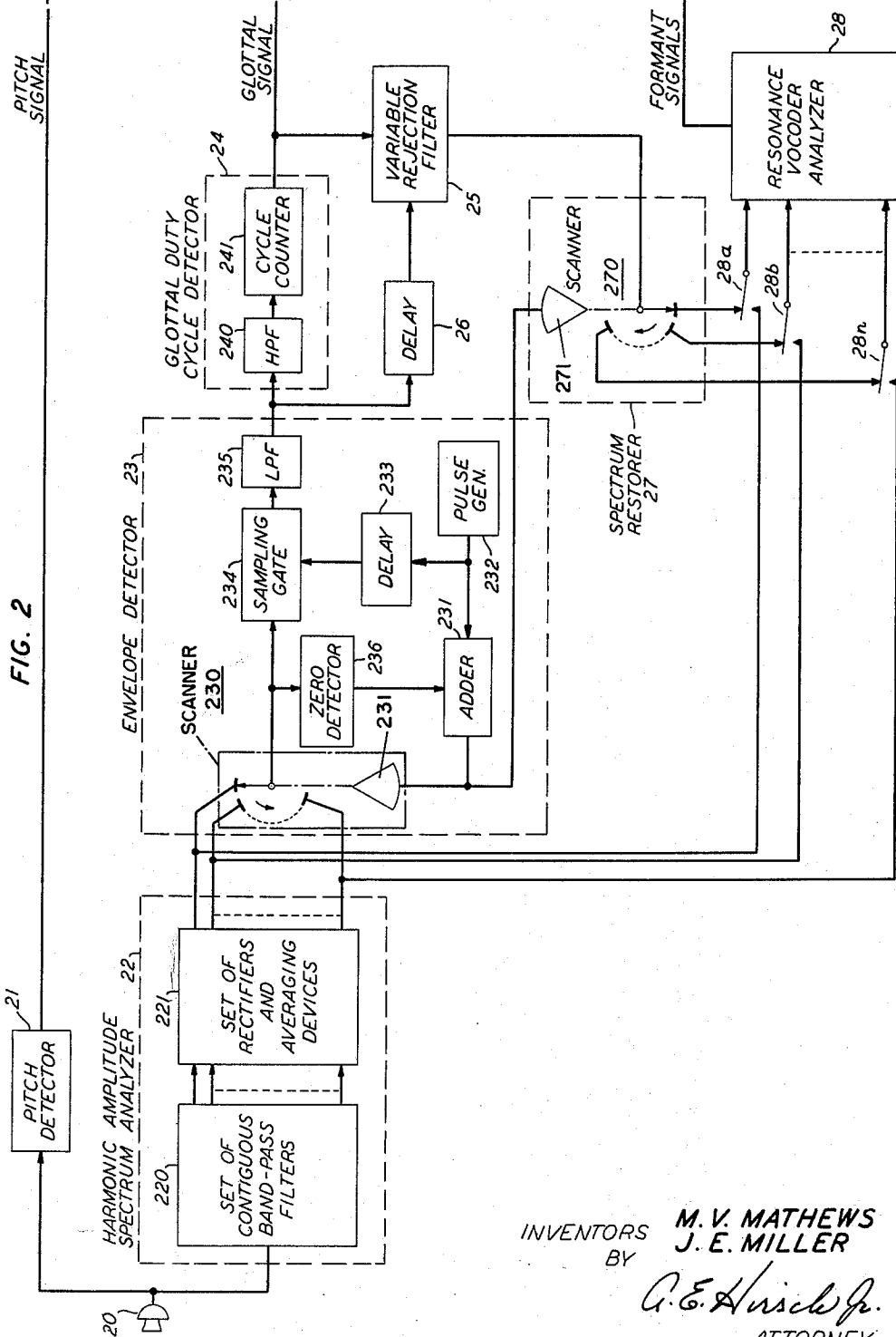
M. V. MATHEWS ET AL

3,083,266

VOCODER APPARATUS

Filed Feb. 28, 1961

6 Sheets-Sheet 2



INVENTORS
BY
M. V. MATHEWS
J. E. MILLER
G. E. Hirsch Jr.
ATTORNEY

March 26, 1963

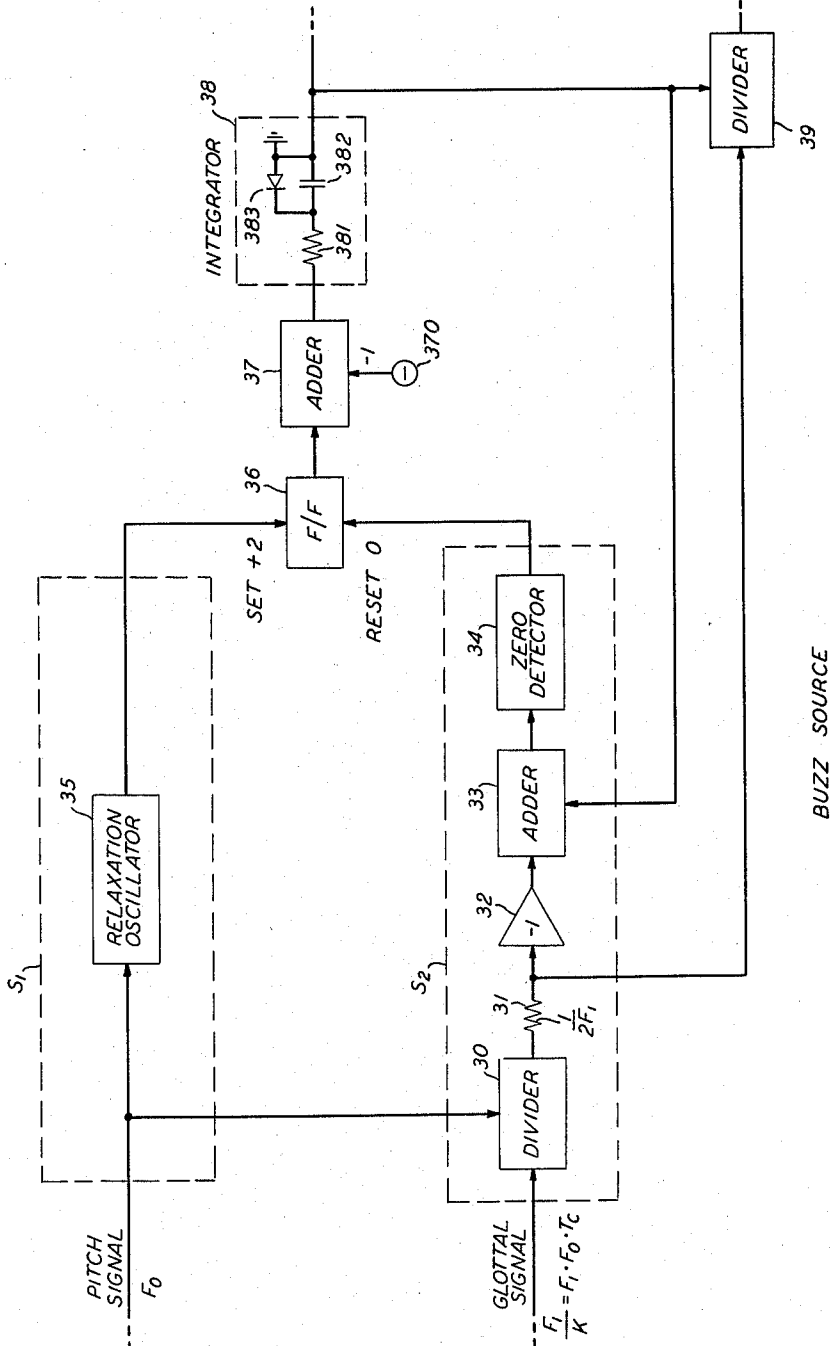
M. V. MATHEWS ET AL
VOCODER APPARATUS

3,083,266

Filed Feb. 28, 1961

6 Sheets-Sheet 3

FIG. 3



M. V. MATHEWS
J. E. MILLER
INVENTORS
BY

A. E. Hirsch Jr.
ATTORNEY

March 26, 1963

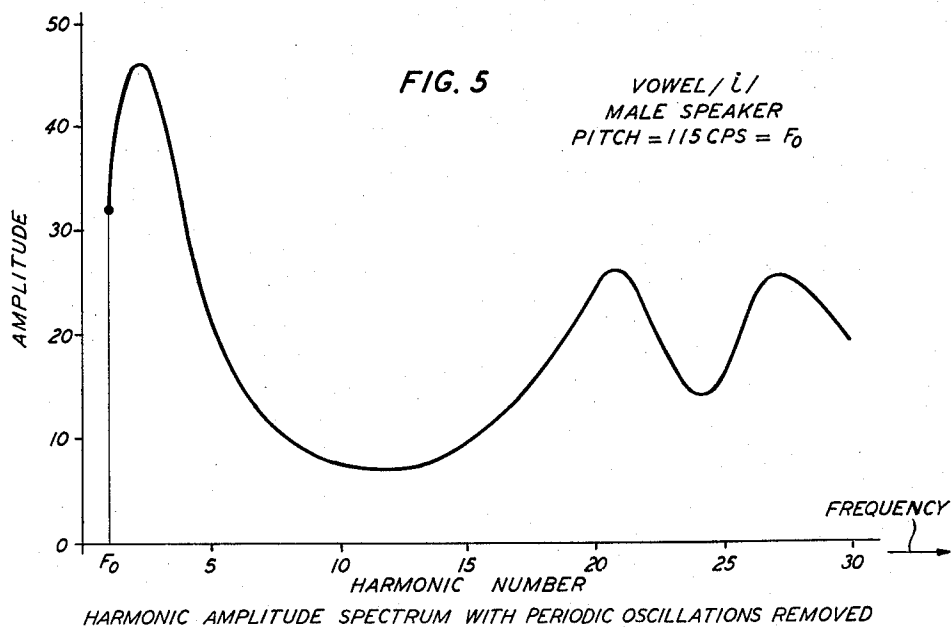
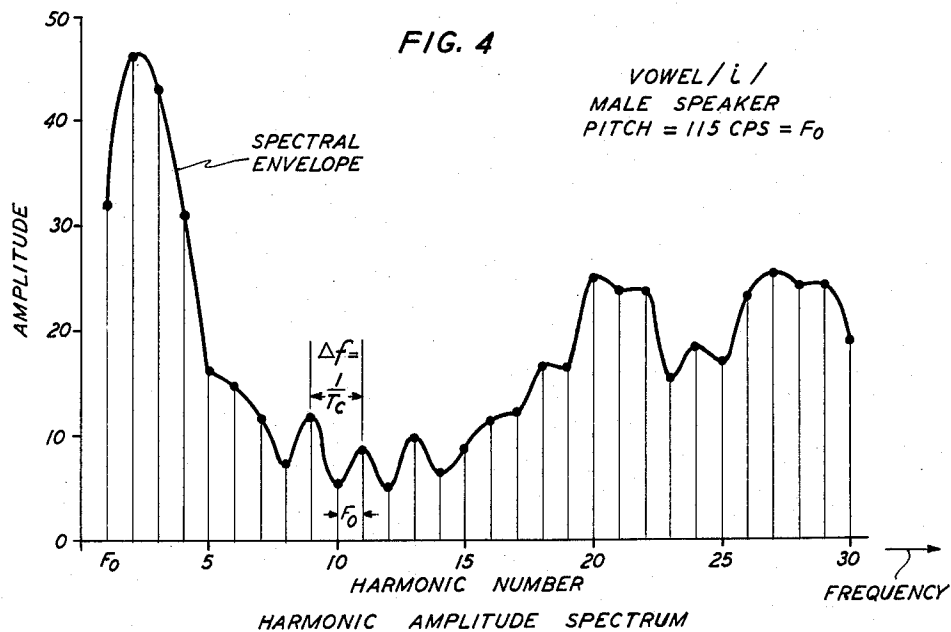
M. V. MATHEWS ET AL

3,083,266

VOCODER APPARATUS

Filed Feb. 28, 1961

6 Sheets-Sheet 4



INVENTORS **M. V. MATHEWS**
J. E. MILLER
BY
G. E. Hirsch Jr.
ATTORNEY

March 26, 1963

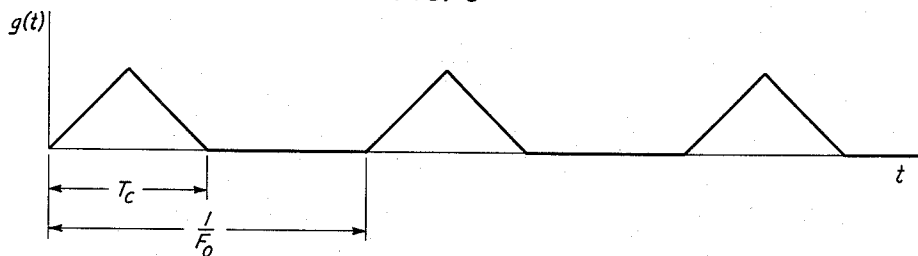
M. V. MATHEWS ET AL
VOCODER APPARATUS

3,083,266

Filed Feb. 28, 1961

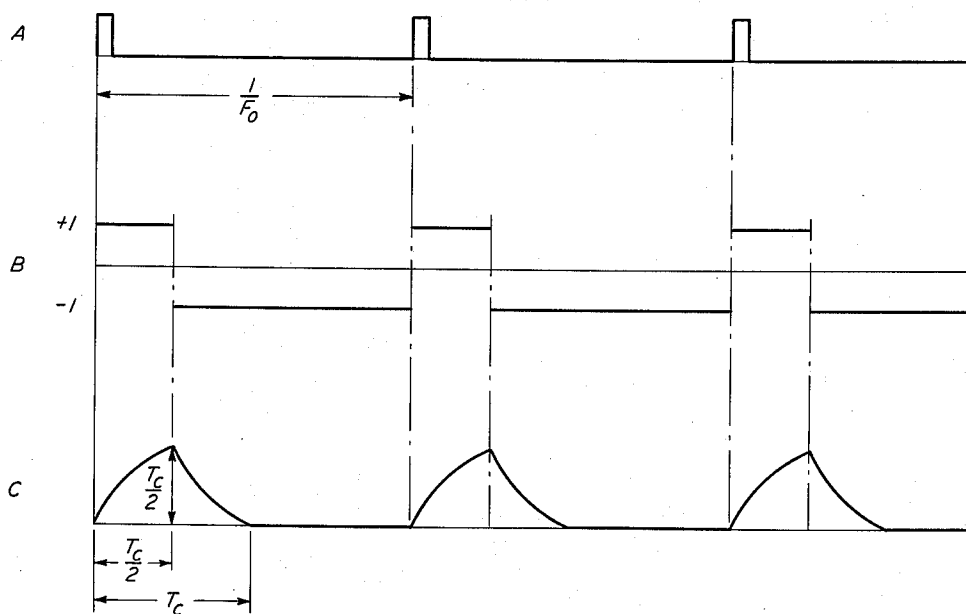
6 Sheets-Sheet 5

FIG. 6



IDEALIZED GLOTTAL EXCITATION FUNCTION

FIG. 7



INVENTORS **M. V. MATHEWS**
BY **J. E. MILLER**

G. E. Hirsch Jr.
ATTORNEY

March 26, 1963

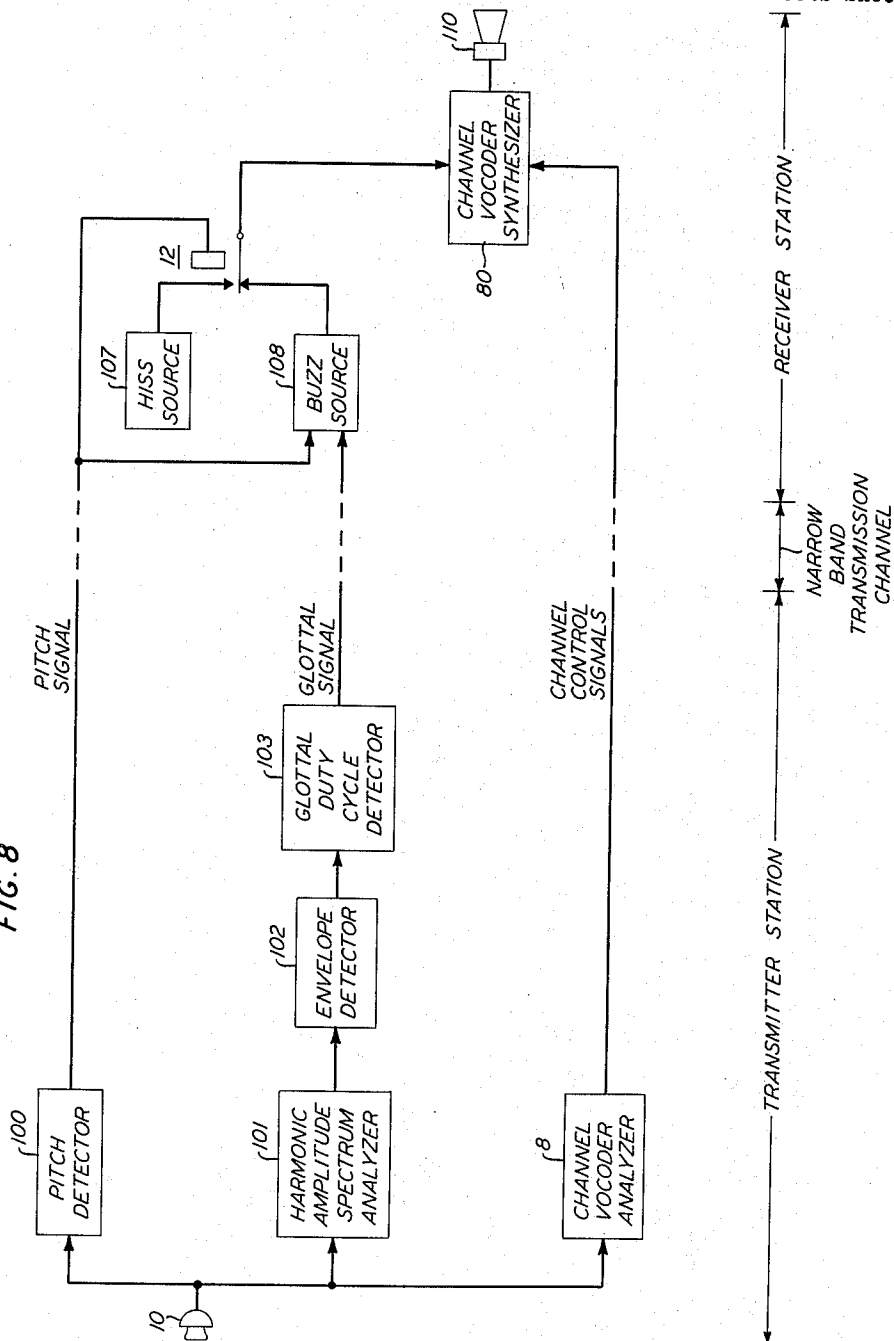
M. V. MATHEWS ET AL
VOCODER APPARATUS

3,083,266

Filed Feb. 28, 1961

6 Sheets-Sheet 6

FIG. 8



INVENTORS M. V. MATHEWS
J. E. MILLER
BY C. E. Hirsch Jr.
ATTORNEY

1

3,083,266

VOCODER APPARATUS

Max V. Mathews, New Providence, and Joan E. Miller, Plainfield, N.J., assignors to Bell Telephone Laboratories, Incorporated, New York, N.Y., a corporation of New York

Filed Feb. 28, 1961, Ser. No. 92,300

14 Claims. (Cl. 179-1)

This invention relates to the transmission of speech waves, and in particular to the transmission of speech waves by vocoder systems. One of its principal objects is to improve the quality of artificial speech produced by vocoder systems. Another object is to improve the accuracy with which the information-bearing characteristics of speech are detected by such systems.

Both the voiced sounds and the unvoiced sounds of which human speech is composed are produced by excitation of the vocal tract, but voiced speech sounds are produced by excitation of the vocal tract resonances or formants with puffs of air released by the glottis or vocal cord, whereas unvoiced speech sounds are produced by the passage of turbulent air through constrictions in the vocal tract. Resonance vocoders of the type described by J. C. Steinberg in Patent 2,635,146, issued April 14, 1953, and by H. L. Barney in Patent 2,819,341, issued January 7, 1958, transmit the information content of speech over a relatively narrow band of frequencies by analyzing a speech wave to obtain one narrow band signal representative of the speech characteristics attributable to the excitation, and a group of narrow band signals representative of the speech characteristics attributable to the vocal tract. In the Steinberg vocoder, for example, two of the excitation characteristics are represented by a voiced-unvoiced pitch signal: the instantaneous voiced or unvoiced nature of the speech sound, and, if the sound is voiced, the pitch or fundamental frequency of the excitation, or its reciprocal, the fundamental period of the excitation. Further, the characteristics of the vocal tract are represented by a group of formant signals that specify the instantaneous frequency and intensity of each vocal tract resonance or formant. Thus, in a resonance vocoder the pitch signal and the formant signals specify the information content of speech, with the important advantage that transmission of these information-bearing signals to a receiver station requires a substantially smaller band of frequencies than transmission of the speech wave itself. At the receiver station, an artificial speech wave is synthesized from the transmitted pitch signal and formant signals. The pitch signal controls the application of one or another of two artificial excitation sources to a network whose characteristics simulate those of the human vocal tract, where one of the artificial excitation sources is a buzz source that generates a periodic signal for the reconstruction of voiced sounds, and the other excitation source is a hiss source that generates an aperiodic signal for the reconstruction of unvoiced sounds. The operation of the vocal tract network is controlled by the formant signals to reconstruct from the artificial excitation signals a synthetic speech wave with the proper periodicity and in which the formants of the voiced portions are replicas of the formants of the original speech wave.

Recent investigations, however, reveal that although the typical vocoder pitch signal contains adequate information regarding the excitation characteristics of voiced sounds, the same pitch signal is an incomplete representation of the excitation characteristics of voiced sounds. In particular, the conventional pitch signal is a sufficiently accurate measure of the periodicity of the excitation function produced by the glottis during voiced sounds, but it contains no information regarding the internal

2

structure or waveform of individual periods of the glottal excitation function. As a result, the artificial excitation signal generated at a vocoder receiver station from a conventional pitch signal is an inaccurate replica of the original glottal excitation function, thereby impairing the quality of vocoder speech reproduced from such an artificial excitation signal.

It is a specific object of the present invention to improve the quality of vocoder speech by deriving from a speech wave an additional information-bearing signal representative of the internal structure or waveform of each period of the glottal excitation function.

Fourier analysis of individual periods of voiced sounds produces harmonic amplitude spectra whose envelopes contain periodic oscillations that are distinct from the major spectral peaks corresponding to vocal tract formants. Applicants have discovered that the frequency of these oscillations is related to the internal structure of each period of the glottal excitation function in the following specific manner: The glottal excitation function is a periodic function of time in which each period contains both a zero portion and a nonzero portion, and the duration of the nonzero portion of each period, which corresponds to the open or duty cycle of the glottis, is equal to the reciprocal of the frequency of periodic oscillations in the spectral envelope. By measuring the frequency of these oscillations, the apparatus of the present invention derives a so-called glottal signal that describes with accuracy the instantaneous duty cycle of the glottis, that is, the duration of the nonzero portion of each period of the glottal excitation function. By obtaining the glottal signal, in addition to a conventional pitch signal, this invention provides a more complete and a more accurate description of the excitation characteristics than is provided by the pitch signal alone. The present invention utilizes the glottal signal and a conventional pitch signal in a novel vocoder buzz source to generate an artificial excitation signal for the synthesis of voiced sounds. The periodicity of this artificial excitation signal is controlled by the pitch signal and the duration of the nonzero portion of each period of this artificial excitation signal is controlled by the glottal signal. By employing the novel artificial excitation signal generated by the buzz source of this invention in the synthesizer of a resonance vocoder, a highly intelligible, natural sounding replica of the original speech wave is reproduced.

Another important feature of this invention is the utilization of the glottal signal of this invention to improve the quality of vocoder speech by improving the accuracy of the formant signals obtained by a resonance vocoder analyzer. The periodic oscillations in the spectral envelope due to the glottal duty cycle tend to mask the major formant peaks in the envelope, and therefore the oscillations are a possible source of error in the detection of formant peaks by a vocoder analyzer. The present invention substantially reduces this source of error by utilizing the glottal signal to control the removal of the periodic oscillations from the spectral envelope before applying the speech wave to a resonance vocoder analyzer.

The invention will be fully understood from the following detailed description of preferred embodiments thereof taken in connection with the appended drawings, in which:

FIG. 1 is a schematic block diagram showing a complete resonance vocoder system embodying the apparatus of this invention;

FIG. 2 is a schematic block diagram showing the transmitter station of a resonance vocoder embodying apparatus for deriving a glottal signal and for removing periodic oscillations from the spectral envelope of a speech wave;

FIG. 3 is a schematic diagram showing buzz source

apparatus for generating an artificial excitation signal for the synthesis of voiced sounds from the glottal signal derived by the apparatus of FIG. 2;

FIG. 4 is a diagram of the harmonic amplitude spectrum of a typical voiced sound;

FIG. 5 is a waveform showing the envelope of the spectrum of FIG. 4 with the periodic oscillations removed;

FIG. 6 is a waveform showing an idealized version of several periods of the glottal excitation function;

FIG. 7 is a group of graphs of assistance in explaining the operation of the apparatus of FIG. 3; and

FIG. 8 is a schematic block diagram showing a complete channel vocoder system embodying the apparatus of this invention.

Theoretical Considerations

The voiced or vowel sounds of human speech are produced by quasi-periodic puffs of air released from the lungs into the vocal tract through opening and closing the glottis or vocal cord. The characteristics of voiced sounds reflect the properties of the vocal tract and the glottis, and some of the individual properties of the vocal tract and the glottis may be determined from a Fourier analysis of a single period of a voiced sound. A Fourier analysis of the sound [i], as in *mete*, for example, produces the harmonic amplitude spectrum shown in FIG. 4, in which the regularly spaced vertical lines represent the amplitudes of the harmonic frequency components of this sound, and in which the fundamental frequency, F_0 , of this sound is 115 cycles per second. It has been determined that the fundamental frequency component of this voiced sound, and of voiced sounds generally, is attributable to the quasi-periodic excitation function produced by the glottis. FIG. 4 also illustrates that the envelope of the spectrum contains several major peaks occurring at approximately harmonics number 2, 20, and 27, and it is well known that these major peaks, which are characteristic of all voiced sounds, are attributable to the vocal tract resonances or formants.

Although knowledge of the instantaneous fundamental frequency of the glottal excitation and the instantaneous locations and amplitudes of the vocal tract formants is sufficient to specify the information content of voiced speech sounds, it is now known that the glottal excitation function is not completely described by its fundamental frequency alone. Recent investigations by J. L. Flanagan, "Some Properties of the Glottal Sound Source," volume 1, *Journal of Speech and Hearing Research*, page 99 (1958), and by R. L. Miller, "Nature of the Vocal Cord Wave," volume 31, *Journal of the Acoustical Society of America*, page 667 (1959), reveal that the internal structure of each period of the waveform generated by the glottis is another important property of the glottal excitation function. Both Flanagan and Miller obtained glottal waveforms in which each period of the wave contained both a zero portion and a nonzero, triangular or pulse-shaped portion, corresponding to the closed and open positions of the glottis, respectively. An idealized version of the internal structure of several periods of a glottal excitation function, $g(t)$, is illustrated in FIG. 6, where the duration of the nonzero portion of each period, also referred to as the glottal duty cycle, is denoted by T_c .

Referring again to FIG. 4, applicants noted that the spectral envelope contains small periodic oscillations whose period is typically several harmonics in length. Analysis revealed that these oscillations are not only attributable to the glottal excitation function rather than to the vocal tract, but also that the frequency of these oscillations is equal to the reciprocal of the glottal duty cycle

This discovery provides the basis for the apparatus of the present invention, which by measuring the frequency of these oscillations in the spectral envelope of a speech wave, obtains a signal representative of the instantaneous duration of the glottal duty cycle. Further, this invention employs the glottal signal thus obtained to improve the quality of vocoder speech in two important respects: At a resonance vocoder transmitter station, the glottal signal is utilized to remove the periodic oscillations from the spectral envelope of the speech wave in order to improve the accuracy of the formant signals by eliminating a source of error in the location of formant peaks; and at a vocoder receiver station, the glottal signal is utilized to reconstruct an accurate replica of the original glottal excitation function for use as an artificial excitation signal in the synthesis of voiced sounds.

Complete System

With reference to FIG. 1, the apparatus of this invention is shown embodied in a complete resonance vocoder system. At the vocoder transmitter station, which is shown in detail in FIG. 2, an incoming speech wave from source 10, for example, an ordinary telephone transmitter, is applied in parallel to conventional pitch detector 100 and to harmonic amplitude spectrum analyzer 101. Pitch detector 100 derives a pitch signal of the type previously described, and harmonic amplitude spectrum analyzer 101 derives a group of signals representative of the amplitudes of the harmonic frequency components of the speech wave, that is, analyzer 101 performs a Fourier analysis upon the speech wave. Envelope detector 102 develops from the output signals of analyzer 101 a time-varying wave that closely follows the envelope of the speech amplitude spectrum, including the small periodic oscillations in the spectral envelope caused by the glottal duty cycle. The time-varying envelope wave from detector 102 is passed to glottal duty cycle detector 103, which measures the frequency of the periodic oscillations of the envelope wave to form at its output terminal a signal indicative of the glottal duty cycle. The glottal output signal of detector 103 is transmitted to the vocoder receiver station, together with the pitch signal from pitch detector 100 and the formant signals from resonance vocoder analyzer 106, as one of the group of narrow band vocoder signals which specifies the information content of the incoming speech wave.

The present invention, however, also utilizes the glottal signal at the vocoder transmitter station to improve the accuracy of the formant signals derived by analyzer 106. FIG. 5 shows the envelope of the spectrum of FIG. 4 with the periodic oscillations removed, and it is observed in a comparison of these figures that the periodic oscillations arising from the glottal duty cycle tend to mask the exact location of the formant peaks and thereby constitute a source of error in the detection of formants by vocoder analyzer 106. By removing these oscillations before formant detection, this invention enables vocoder analyzer 106 to locate formant peaks with greater precision and the quality of vocoder speech is improved. In the apparatus shown in FIG. 1, variable rejection filter 104, whose variable rejection band is controlled by the glottal signal, removes the periodic oscillations from the envelope wave, and spectrum restorer 105 reconstructs from the smoothed envelope wave a group of signals representative of the amplitudes of the harmonic components of the original speech wave for formant analysis by vocoder analyzer 106. Switch 11 is provided, however, if it is desired not to remove the periodic oscillations, in which case the amplitude signals from spectrum analyzer 101 may be applied directly to vocoder analyzer 106.

At the vocoder receiver station shown in FIG. 1, the pitch signal is passed to relay 12 to control the application of an artificial excitation signal from either conventional hiss source 107 or buzz source 108, shown in detail in FIG. 3, to resonance vocoder synthesizer 109.

5

When the pitch signal indicates the presence of unvoiced sounds, relay 12 is deenergized to connect hiss source 107 to synthesizer 109, whereas when the pitch signal indicates the presence of voiced sounds, relay 12 is energized, thereby connecting buzz source 108 of this invention to synthesizer 109. Synthesizer 109, which may be of any well-known construction, is controlled by the formant signals to reconstruct an artificial speech wave whose formants closely resemble the formants of the original speech wave. Reproducer 110, for example, a loudspeaker, converts the artificial speech wave output of synthesizer 109 into audible, high quality speech sounds.

Transmitter Apparatus

Referring now to FIG. 2, that part of the apparatus of this invention which is located at the transmitter station of a vocoder system is shown in detail. An incoming speech wave from source 20 is applied in parallel to pitch detector 21 and to harmonic amplitude spectrum analyzer 22, the latter element corresponding to analyzer 101 of FIG. 1. Within analyzer 22, a set of contiguous band-pass filters, denoted by element 220, separates the incoming speech wave into its individual frequency components. The pass bands of the individual filters should be sufficiently narrow to ensure that not more than one of the harmonic frequency components of voiced speech sounds is passed by each filter; for example, a set of bandpass filters spanning the frequency range from 100 to 3,000 cycles per second in contiguous bands of 75 cycles per second width satisfactorily meets this requirement. Each of the speech frequency components passed by filters 220 is applied to one of a set of rectifiers followed by averaging devices, represented by element 221, where the averaging devices may be low-pass filters, in order to obtain a direct-current signal proportional to the amplitude of each harmonic component. The output signals of analyzer 22 thus represent the harmonic amplitude spectrum of the incoming speech wave, and the envelope of the spectrum represented by these signals contains periodic oscillations due to the glottal duty cycle. By converting the parallel output signals of analyzer 22 into a serial train of samples, and by filtering the serial train of samples, envelope detector 23 obtains a time-varying wave that closely follows the envelope of the spectrum. Each of the output signals of analyzer 22 is applied to one of the input terminals of scanner 230, which may be a shaper 231 followed by a well-known ring circuit of multivibrators, and associated circuitry, stepped along from stage to stage by pulses supplied from pulse generator 232 at a rate of F_1 pulses per second; circuit details of a suitable scanner are illustrated in FIG. 1 of J. G. Kreer et al. Patent 2,527,638, issued October 31, 1960. Scanner 230 converts the parallel output signals of analyzer 22 into a serial train of signals spaced at uniform $1/F_1$ second intervals, and this train of signals is sampled at the end of each $1/F_1$ second interval by sampling gate 234. The sampling operation of gate 234 is also controlled by pulses from generator 232, which are delayed by passage through element 233 for a portion of each $1/F_1$ second interval, for example, $0.9/F_1$ second, in order to obtain at the output terminal of gate 234 a train of samples spaced uniformly at $1/F_1$ second intervals. Since the speech spectrum, as measured by analyzer 22, changes slowly with time at approximately a syllabic rate, scanner 230 must operate fast enough to make a single complete scan of the output signals of analyzer 22 before the spectrum has changed appreciably; it has been determined that a pulse rate on the order of 2,000 pulses per second is sufficient to achieve this result.

The amplitudes of the samples obtained at the output terminal of gate 234 are proportional to the magnitudes of the output signals of analyzer 22, but whereas the spacing between adjacent output signals of analyzer 22 is F_0 cycles per second on the frequency scale, where F_0 is the fundamental speech frequency as shown in FIG. 4, the

6

spacing between adjacent samples at the output terminal of gate 234 is $1/F_1$ second on the time scale. The train of samples from gate 234 is passed through low-pass filter 235, which has a cut-off frequency of $F_1/2$ cycles per second, to obtain a time-varying wave whose shape closely follows that of the spectral envelope of the speech wave.

The accurate determination of the glottal duty cycle by detector 24 requires that the sequence of samples delivered to filter 235 be composed solely of samples derived from the amplitudes of harmonic components. By making the pass bands of filters 220 sufficiently narrow to ensure that not more than one harmonic will fall within the pass band of each filter, it is possible that no harmonic component will fall within the pass band of a particular filter and that therefore the output signal of such a filter is of zero amplitude. Since a sample derived from a zero amplitude output signal is not derived from the amplitude of a harmonic component, such a sample must not appear in the sequence of samples if the glottal duty cycle is to be determined with accuracy. This invention prevents the appearance of such samples by employing zero detector 236, of conventional construction, which is connected to the output terminal of scanner 230 and serves to deliver a stepping pulse through adder 231 whenever the output signal of scanner 230 is zero, thereby advancing scanner 230 to the next stage containing a nonzero signal before the next regular stepping pulse is produced by generator 232. In the event that two or more adjacent filters have zero output signals, zero detector 236 will continue to deliver stepping pulses until a nonzero signal is reached; to provide for this contingency, zero detector 236 is designed to operate sufficiently fast to deliver more than one pulse in the $1/F_1$ second interval between successive pulses from generator 232.

A signal representative of the instantaneous glottal duty cycle is obtained by applying the envelope wave from envelope detector 23 to glottal duty cycle detector 24, which comprises high-pass filter 240 followed by cycle counter 241, both of which elements are of well-known design. Since the periodic oscillations in the spectral envelope which are caused by the glottal duty cycle are considerably more rapid than other fluctuations in the envelope, the periodic oscillations in the corresponding time-varying envelope wave are separated from other fluctuations by providing filter 240 with a cut-off frequency of approximately 500 cycles per second. Cycle counter 241 then derives from the periodic oscillations passed by filter 240 a signal that is proportional to the frequency of these oscillations.

The exact relationship of the output signal of detector 24 to the glottal duty cycle is obtained from the following considerations. From an examination of the spectral envelope shown in FIG. 4, it is observed that a single "period," Δf , of the envelope oscillations attributable to the glottal duty cycle is larger than the spacing, F_0 , between harmonics, that is,

$$\Delta f = KF_0, K > 1 \quad (1)$$

The "frequency" of the periodic oscillations in the spectral envelope is therefore the reciprocal of Equation 1,

$$\frac{1}{\Delta f} = \frac{1}{KF_0} \quad (2)$$

Since

$$\frac{1}{\Delta f} = T_0$$

the glottal duty cycle, and

$$\frac{1}{F_0} = T_0$$

the fundamental period, Equation 2 may be rewritten

$$T_0 = \frac{T_0}{K} \quad (3)$$

In physical terms, Equation 3 expresses the well-known observation that the glottis is closed for a fraction of each fundamental period. Typically, K varies from 2 to 3, that is, the glottis is closed for $\frac{1}{2}$ to $\frac{1}{3}$ of each fundamental period, and the spectrum shown in FIG. 4 illustrates a glottal closure equal to one-half of the fundamental period. When the apparatus of envelope detector 23 described above converts the parallel output signals of analyzer 22 into a serial train of samples uniformly spaced at intervals of $1/F_1$ second, the period, $\Delta f'$, of the oscillations in the corresponding time-varying envelope wave derived from the samples by detector 24 is

$$\Delta f' = \frac{K}{F_1} \quad (4)$$

and the frequency of these oscillations is

$$\frac{1}{\Delta f'} = \frac{F_1}{K} \quad (5)$$

Substituting for K in Equation 5 from Equation 1,

$$\frac{1}{\Delta f'} = \frac{F_1 \cdot F_0}{\Delta f} = F_1 \cdot F_0 \cdot T_c \quad (6)$$

hence by measuring the frequency of the periodic oscillations in the envelope wave, the output signal developed by detector 24 is directly proportional to the glottal duty cycle, T_c .

Although the glottal signal from detector 24 is transmitted to the vocoder receiver station as an information-bearing signal together with the pitch signal and the formant signals, as shown in FIG. 2, it is also employed at the transmitter station to improve the accuracy of the formant signals by removing from the spectral envelope of the speech wave the periodic oscillations caused by the glottal duty cycle which mask the formant peaks. This is accomplished by applying the glottal signal to the control terminal of variable rejection filter 25, and by applying the envelope wave from detector 23 to the input terminal of filter 25, after passage through delay element 26 to compensate for the delay in deriving the glottal signal. Filter 25, which may be of the type described by C. R. Howard in volume 28, Journal of the Acoustical Society of America, page 1096 (1956), has a transmission characteristic that is continuously variable over a preselected frequency range in response to an applied control signal. The variable transmission characteristic of filter 25 is controlled by the glottal signal, and the frequency range of this characteristic is selected to produce a substantial attenuation of the frequency components of the periodic oscillations in the envelope wave. By thus removing the periodic oscillations attributable to the glottal duty cycle, the major formant peaks of the envelope wave, as illustrated in FIG. 5, are located by analyzer 28 with a greatly reduced likelihood of error.

The output wave filter 25 may be applied directly to an analyzer of the type described by E. E. David, Jr., in volume CT-3, I.R.E. Transactions on Circuit Theory, page 239 (1956); or, as illustrated in FIG. 2 the smoothed envelope wave from filter 25 may be applied to spectrum restorer 27, which comprises a scanner 270 operated in synchrony with scanner 230, to develop a group of signals representative of the amplitudes of the harmonic components of the incoming speech wave, but without oscillations in amplitude due to the glottal duty cycle. Scanner 270 may comprise a shaper 271 followed by a ring circuit of multivibrators, and associated circuitry, of construction identical with that of scanner 230. The output signals of restorer 27 are then passed to resonance vocoder analyzer 28, for example, the analyzer shown in FIG. 1 of J. L. Flanagan Patent 2,891,111, issued June 16, 1959. The formant signals obtained by analyzer 28 are transmitted together with the pitch signal and the glottal signal to the vocoder receiver station over a suitable channel, indicated in the drawing by broken lines. Switches 23a, 23b, . . . , 23n are provided, however, if it is desired to derive the formant signals

directly from the original harmonic amplitude signals generated by analyzer 22 without removing the periodic oscillations in the spectral envelope.

Buzz Source

With reference to FIG. 3, the buzz source apparatus of this invention, corresponding to buzz source 103 of FIG. 1, comprises an upper subpath, S_1 , and a lower subpath, S_2 , where the output terminals of these subpaths are connected to the set and reset terminals of flip-flop 36, respectively. The transmitted glottal signal is applied to the dividend terminal of a well-known divider circuit 30 in subpath S_2 , and the transmitted pitch signal is applied both to the divisor terminal of divider 30 and to the input terminal of relaxation oscillator 35 in subpath S_1 . From Equation 6, the quotient signal formed at the output terminal of divider 30 is proportional to $F_1 \cdot T_c$, and by passing the quotient signal through resistor 31 whose resistance is proportional to the constant

$$\frac{1}{2F_1} \cdot \frac{T_c}{2}$$

a signal proportional to

is obtained at the output terminal of resistor 31. The output signal of resistor 31 is applied through a polarity inverter, for example, a minus one amplifier 32, to one of the input points of adder 33, and the output terminal of integrator 38 is connected to the other input point of adder 33. Whenever the output signals of polarity inverter 32 and integrator 38 are of equal amplitude but of opposite polarity, the output signal of adder 33 becomes zero, thereby causing zero detector 34 to generate a pulse that resets flip-flop 36 to its zero state.

The pitch signal applied to relaxation oscillator 35 in subpath S_1 produces a train of periodic pulses at the output terminal of oscillator 35 at the rate of F_0 pulses per second, as illustrated in graph A of FIG. 7. These pulses, which in prior art vocoders are utilized without further modification as an artificial excitation signal for the reconstruction of voiced sounds, are used in the present invention to set flip-flop 36 to a plus two (+2) state. By combining in adder 37 the output signal of flip-flop 36 with a constant minus one (-1) signal supplied by source 370, the output signal of adder 37 is either plus one or minus one, as shown in graph B of FIG. 7, depending upon the state of flip-flop 36. The plus one-minus one output signal of adder 37 is integrated by integrator 38, composed of resistor 381 connected in series with capacitor 382. When a plus one signal from adder 37 is integrated in response to a pulse from oscillator 35, a positive charge is developed on capacitor 382 which increases until the charge reaches a value equal to $T_c/2$, as shown in graph C of FIG. 7, at which point it cancels the $-T_c/2$ signal from polarity inverter 32, thereby causing zero detector 34 to reset flip-flop 36 to its zero state and changing the output signal of adder 37 from plus one to minus one. When a minus one signal from adder 37 is applied to integrator 38, capacitor 382 is discharged from a value of $T_c/2$ to zero, as illustrated in graph C of FIG. 7, at which point rectifier 383 holds or clamps capacitor 382 until the next pulse from oscillator 35 changes the output signal of adder 37 to plus one and causes a positive charge to develop again on capacitor 382. As shown in graph C of FIG. 7, the alternate charging and discharging of capacitor 382 produces at the output terminal of integrator 38 a periodic signal whose waveform approximates the idealized glottal excitation function shown in FIG. 6; that is, the length of each period of the output signal is equal to the fundamental period, $1/F_0$, and the duration of the nonzero portion of each period is equal to the glottal duty cycle T_c .

Since the amplitude of the output signal of inte-

grator 38 varies with the value of T_0 , as illustrated graphically in FIG. 7C, the output signal of integrator 38 is suitable for use as an artificial excitation signal in vocoder synthesizers where information regarding the amplitudes or intensities of the formants of such sounds is conveyed by the artificial excitation signal; an example of this type of system is described in the aforementioned H. L. Barney patent. For use in a vocoder system of the variety described in the aforementioned Steinberg patent, however, it is necessary to modify the output signal of integrator 38 to make its peaks uniform in amplitude, since the Steinberg vocoder transmits separate information-bearing signals specifying the formant amplitudes. The output signal of integrator 38 is modified to make its peaks uniform in amplitude by applying it to the dividend terminal of divider 39, and by connecting the output terminal of resistor 31 to the divisor terminal of divider 39. The quotient signal obtained at the output terminal of divider 39 has peaks of uniform amplitude, despite variations in T_0 , and this quotient signal may be utilized as an artificial excitation signal for the reproduction of voiced sounds.

Channel Vocoder

Although the apparatus of this invention has been described as a part of a resonance vocoder system, it is to be understood that this invention may also be used to improve the quality of speech produced by a channel vocoder analyzer of the type described by H. W. Dudley in Patent 2,151,091, issued March 21, 1939. Referring to FIG. 8, the pitch signal and the glottal signal are derived from an incoming speech wave as described above, and channel vocoder analyzer 8, the operation of which is set forth in detail in the Dudley patent, obtains a group of channel control signals representative of the amplitudes of the harmonic components of the speech wave. The pitch signal, glottal signal, and channel control signals are transmitted to a receiver station, where the pitch signal and glottal signal are utilized to generate a voiced artificial excitation signal of the type produced at the output terminal of divider 39 of the buzz source apparatus shown in FIG. 3. The artificial excitation signal is passed to channel vocoder synthesizer 80, the construction of which is also described in the Dudley patent, where the transmitted channel control signals operate the synthesizer to reconstruct a high quality artificial speech wave.

It is to be understood that the above-described arrangements are merely illustrative of applications of the principles of the invention. Numerous other arrangements may be devised by those skilled in the art without departing from the spirit and scope of the invention.

What is claimed is:

1. In a vocoder speech transmission system, the combination that comprises a source of a speech wave, means for deriving from said speech wave a plurality of control signals indicative of the information-bearing characteristics of said speech wave, including means for measuring periodic oscillations in the spectral envelope of said speech wave to obtain a control signal indicative of the glottal duty cycle characteristic of said speech wave, means for transmitting said plurality of control signals to a receiver station, and at said receiver station, means for synthesizing an artificial speech wave from said transmitted control signals.

2. In a resonance vocoder speech transmission system, the combination that comprises a source of a speech wave, means for deriving from said speech wave a group of information-bearing signals representative of the fundamental period and the formants of said speech wave, means for measuring the frequency of periodic oscillations in the spectral envelope of said speech wave to obtain an information-bearing signal representative of the glottal duty cycle of said speech wave, means for transmitting all of said information-bearing signals to a receiver station, and at said receiver station, means for synthesizing an

artificial speech wave from said information-bearing signals.

3. In a channel vocoder speech transmission system, the combination that comprises a source of a speech wave, means for deriving from said speech wave a first signal indicative of the fundamental period of the glottal excitation function of said speech wave, means for counting periodic oscillations in the spectral envelope of said speech wave to obtain a second signal indicative of the duty cycle of the glottal excitation function of said speech wave, means for obtaining from said speech wave a group of channel control signals representative of the amplitudes of the harmonic components of said speech wave, means for transmitting said first signal, said second signal, and said channel control signals to a receiver station, and at said receiver station, means for generating from said first and second signals an artificial excitation signal, and means under the control of said channel control signals for synthesizing a speech wave from said artificial excitation signal.

4. Apparatus for improving the quality of speech produced by a resonance vocoder system which comprises at the transmitter station of said system a source of a speech wave, means for deriving from said speech wave a pitch signal, analyzing means for deriving from said speech wave a group of formant signals, means for obtaining from periodic oscillations in the spectral envelope of said speech wave a signal representative of the glottal duty cycle of said speech wave, means under the influence of said glottal signal for removing periodic oscillations due to the glottal duty cycle from the spectral envelope of said speech wave before applying said speech wave to said analyzing means, means for transmitting said pitch signal, said formant signals, and said glottal signal to a receiver station, and at said receiver station, means for generating from said pitch signal and said glottal signal an artificial excitation signal, and means responsive to said formant signals for synthesizing a speech wave from said artificial excitation signal.

5. In a system for improving the quality of speech produced by a resonance vocoder, the combination that comprises a source of a speech wave, means for deriving a pitch signal from said speech wave, means for obtaining from said speech wave a wave that closely resembles the spectral envelope of said speech wave, means for counting the periodic oscillations in said envelope wave to derive a signal representative of the glottal duty cycle of said speech wave, means controlled by said glottal signal for removing said periodic oscillations from said envelope wave to produce a smoothed envelope wave, means for obtaining from said smoothed envelope wave a group of signals representative of the formants of said speech wave, means for transmitting said pitch signal, said glottal signal, and said formant signals to a receiver station, and at said receiver station, means for utilizing said pitch signal and said glottal signal to produce an artificial excitation signal, and means under the influence of said formant signals for reconstructing a synthetic speech wave from said artificial excitation signal.

6. Apparatus as defined in claim 5 wherein said means for counting the periodic oscillations in said envelope wave to derive a signal representative of the glottal duty cycle of said speech wave comprises a high-pass filter followed by a cycle counter.

7. In a system for improving the quality of speech produced by a resonance vocoder, the combination that comprises means for deriving a pitch signal from said speech wave, means for separating said speech wave into its individual harmonic frequency components, means for deriving from each of said harmonic components a signal proportional to the amplitude of each component, a first sampling means for obtaining from said harmonic amplitude signals a time sequence of uniformly spaced samples whose magnitudes are proportional to the amplitudes of said harmonic frequency components, means

for obtaining from said time sequence of uniformly spaced samples a continuous wave that closely resembles the spectral envelope of said harmonic frequency components, means for isolating periodic oscillations in said continuous wave, means for measuring the frequency of said periodic oscillations to develop a signal proportional to the glottal duty cycle of said speech wave, means controlled by said glottal signal for removing said periodic oscillations from said continuous wave to produce a smoothed continuous wave, a second sampling means synchronized with said first sampling means for reconstructing from said smoothed continuous wave a group of signals representative of the amplitudes of the harmonic frequency components of said speech wave, but without said periodic oscillations, means for analyzing said group of reconstructed signals to obtain a plurality of signals indicative of the formants of said speech wave, means for transmitting said pitch signal, said glottal signal, and said formant signals to a receiver station, and at said receiver station, means for generating from said pitch signal and said glottal signal an artificial excitation signal, and means under the influence of said formant signals for synthesizing a speech wave from said artificial excitation signal.

8. Apparatus for determining the glottal duty cycle of a speech wave which comprises a source of a speech wave, means for deriving from the harmonic amplitude spectrum of said speech wave a wave closely resembling the envelope of said spectrum, means for isolating periodic oscillations in said envelope wave, and means for counting the frequency of said periodic oscillations to develop a signal proportional to the glottal duty cycle of said speech wave.

9. Apparatus for determining the glottal duty cycle of a speech wave which comprises a source of a speech wave, means for separating said speech wave into its individual harmonic frequency components, means for deriving from each of said harmonic components a signal proportional to the amplitude of each component, means for sampling said harmonic amplitude signals to produce a time sequence of uniformly spaced samples whose magnitudes are proportional to the amplitudes of said harmonic frequency components, means for obtaining from said time sequence of uniformly spaced samples a continuous wave that closely resembles the spectral envelope of said harmonic frequency components, means for isolating periodic oscillations in said continuous wave, and means for measuring the frequency of said periodic oscillations to develop a signal proportional to the glottal duty cycle of said speech wave.

10. Apparatus as defined in claim 9 wherein said means for sampling said harmonic amplitude signals comprises a scanner provided with a stepping terminal, an output terminal, and a plurality of input stages in one-to-one correspondence with said harmonic amplitude signals, combining means provided with two input terminals and an output terminal, wherein the output terminal of said combining means is connected to the stepping terminal of said scanner, a first pulse generator connected to one of the input terminals of said combining means for advancing said scanner from stage to stage at a predetermined rate, a second pulse generator connected between the output terminal of said scanner and the other input terminal of said combining means for advancing said scanner to the next stage when an output signal fails to appear at the output terminal of said scanner, gating means provided with two input terminals and an output terminal, means for connecting the output terminal of said scanner to one of the input terminals of said gating means, and delay means connected between said first pulse generator and the second input terminal of said gating means.

11. In a system for the synthesis of speech, a source of a first control signal representative of the fundamental period of a speech wave, a source of a second control signal representative of the glottal duty cycle of said

speech wave, a source of a group of control signals representative of the formants of said speech wave, means for generating from said first and second control signals a first artificial excitation signal for the synthesis of voiced portions of an artificial speech wave, means for generating from said first control signal a second artificial excitation signal for the synthesis of unvoiced portions of an artificial speech wave, and means under the influence of said group of control signals for synthesizing an artificial speech wave from said first and second artificial excitation signals.

12. In a system for synthesizing speech, a source of a first control signal indicative of the fundamental period of the glottal excitation function of a speech wave, a source of a second control signal proportional to the duty cycle of the glottal excitation function of said speech wave, a source of a group of control signals representative of the formants of said speech wave, means for generating from said first and second control signals a first artificial excitation signal for the synthesis of voiced speech sounds including a first subpath containing a relaxation oscillator and provided with an input terminal and an output terminal, means for applying said first control signal to the input terminal of said first subpath, a second subpath provided with an input terminal and an output terminal between which is connected the series combination of a divider circuit provided with a divisor terminal and a dividend terminal, a resistor, a polarity inverter, an adder circuit provided with two input terminals, and a zero detector, means for applying said first control signal to the divisor terminal of the divider circuit in said second subpath, means for applying said second control signal to the dividend terminal of the divider circuit in said second subpath, a flip-flop whose set terminal is connected to the output terminal of said first subpath, and whose reset terminal is connected to the output terminal of said second subpath, a source of a negative polarity signal of unit amplitude, adding means provided with two input terminals and an output terminal, means for connecting the output terminal of said flip-flop to one of the input terminals of said adding means, means for connecting the output terminal of said negative signal source to the other input terminal of said adding means, an integrating means provided with an input terminal and an output terminal, means for connecting the output terminal of said adding means to the input terminal of said integrating means, means for connecting the output terminal of said integrating means to one of the input terminals of the adder in said second subpath, a dividing means provided with a dividend terminal, a divisor terminal, and an output terminal, means for connecting the output terminal of said integrating means to the dividend terminal of said dividing means, and means for connecting the output terminal of the resistor in said second subpath to the divisor terminal of said dividing means, means for generating from said first control signal a second artificial control signal for the synthesis of unvoiced speech sounds, and means under the control of said group of formant signals for synthesizing speech from said first and second artificial excitation signals.

13. Apparatus for generating an artificial excitation signal for the reconstruction of voiced portions of synthetic speech which comprises a source of a first signal representative of the fundamental frequency of a speech wave, a source of a second signal proportional to the instantaneous glottal duty cycle of said speech wave, means supplied with said first signal for generating a first train of pulses with the same period as said fundamental period, and means under the influence of said second signal for modifying said first train of pulses to obtain a second train of triangular-shaped pulses with the same period as said fundamental period and in which the duration of each triangular shaped pulse is proportional to the instantaneous glottal duty cycle of said speech wave.

14. Apparatus for generating an artificial excitation

13

signal for the reconstruction of voiced portions of synthetic speech which comprises a source of a first incoming signal indicative of the fundamental period, F_0 , of the glottal excitation function of a speech wave, a source of a second incoming signal proportional to the duty cycle, T_c , of said glottal excitation function, a first subpath containing a relaxation oscillator and provided with an input terminal and an output terminal, a second subpath comprising the series combination of a divider circuit provided with a divisor terminal and a dividend terminal, a resistance element, a polarity inverter, an adder circuit provided with two input terminals, and a zero detector, means for simultaneously applying said first signal to the input terminal of said first subpath and to the divisor terminal of said divider circuit in said second subpath, means for applying said second signal to the dividend terminal of the divider circuit in said second subpath, a flip-flop whose set terminal is connected to the output terminal of said first subpath and whose reset terminal is connected to the output terminal of said second subpath, a source of a negative polarity signal of constant amplitude, adding means provided with two input terminals and an output terminal, means for connecting the output terminal of said flip-flop to one of the input terminals

14

of said adding means, means for connecting said negative signal source to the other input terminal of said adding means, integrating means whose input terminal is connected to the output terminal of said adding means and whose output terminal is connected to one of the input terminals of the adder circuit in said second subpath, whereby the signal developed at the output terminal of said integrating means has a period of $1/F_0$, the duration of the nonzero portion of each period is proportional to T_c , and the peak amplitude of the nonzero portion of each period is proportional to $T_c/2$, a dividing means provided with a dividend terminal, a divisor terminal, and a quotient terminal, means for connecting the output terminal of said integrating means to the dividend terminal of said dividing means, and means for connecting said resistor in said second subpath to the divisor terminal of said dividing means, whereby the signal developed at the quotient terminal of said dividing means has a period of $1/F_0$, the duration of the nonzero portion of each period is proportional to T_c , and the peak amplitude of the nonzero portion of each period is constant.

No references cited.