

[72] Inventor **Donald Anthony Acott Roworth**  
London, England

[21] Appl. No. **649,518**

[22] Filed **June 28, 1967**

[45] Patented **Dec. 29, 1970**

[73] Assignee **International Standard Electric Corporation**  
New York, N.Y.  
a corporation of Delaware

[32] Priority **July 1, 1966**

[33] **Great Britain**

[31] **Nos. 29684/66, 29685/66 and 29686/66**

[56] **References Cited**

**UNITED STATES PATENTS**

|           |        |          |           |
|-----------|--------|----------|-----------|
| 2,819,341 | 1/1958 | Barney   | 179/1(AS) |
| 3,330,910 | 7/1967 | Flanagan | 179(AS)X  |
| 3,423,530 | 1/1969 | Coulter  | 179/1(AS) |

*Primary Examiner*—Kathleen H. Claffy  
*Assistant Examiner*—Charles W. Jirauch  
*Attorneys*—C. Cornell Remsen, Jr., Rayson P. Morris, Percy P. Lantzy, J. Warren Whitesel, Phillip A. Weiss and Delbert P. Warner

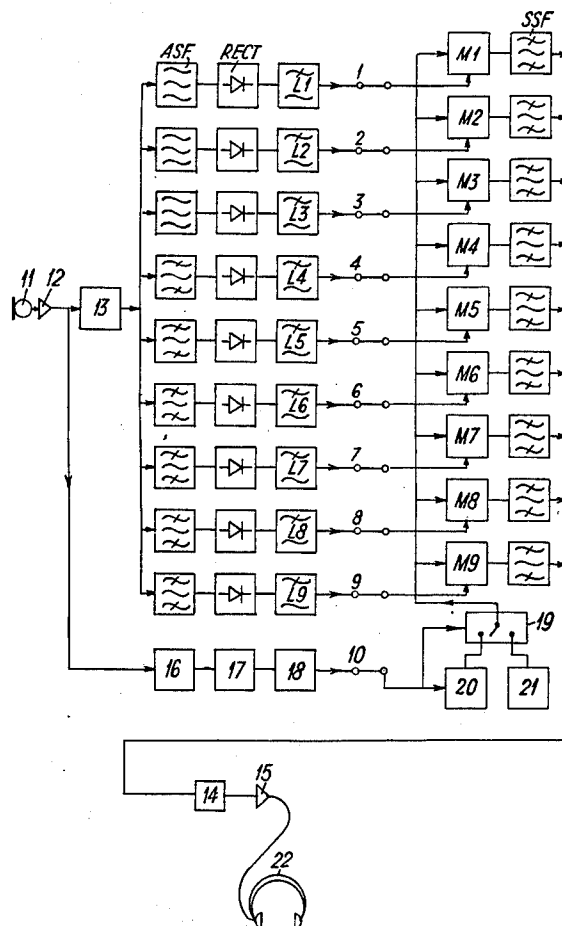
[54] **VOCODER FILTER SYSTEM**  
13 Claims, 11 Drawing Figs.

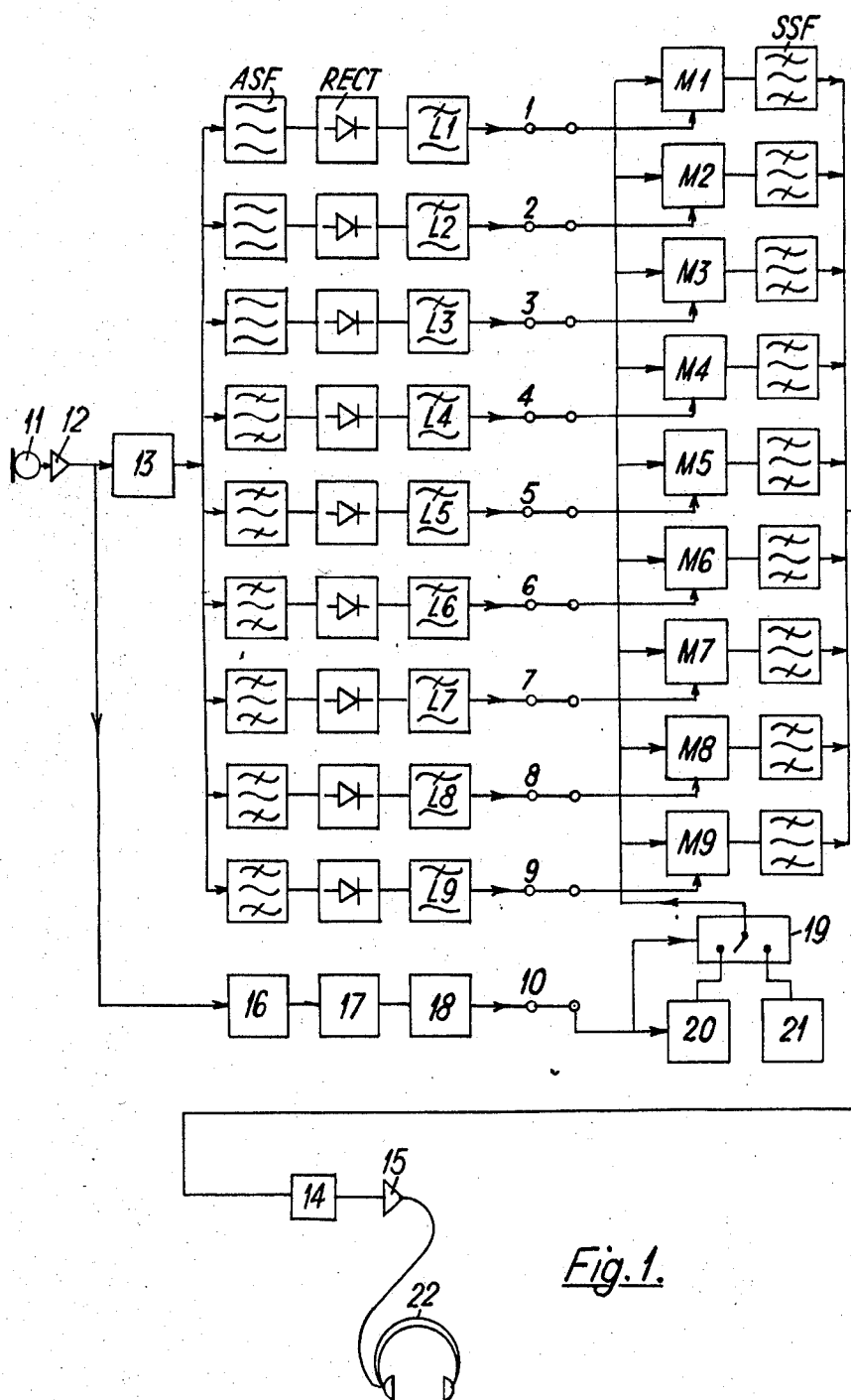
[52] U.S. Cl. .... **179/1,**  
179/15.55

[51] Int. Cl. .... **G101/00**

[50] Field of Search. .... **179/1AS;**  
333/70R, 70T; 323/125; 179/15.55

**ABSTRACT:** A vocoder or voice synthesizer includes a plurality of filters which are used to analyze human speech signals so that they may be synthesized. The original voice signal is separated by a plurality of high Q-filters. A variable resistor is coupled across these filters to selectively increase or decrease the quality factor or Q of the filters. Time controlled clamping means are provided for alternately sampling speech via high Q-filters and then discharging the filter elements by lowering the Q. This way, the energy stored in the filters during one speech sample does not degrade the response to the next speech sample.





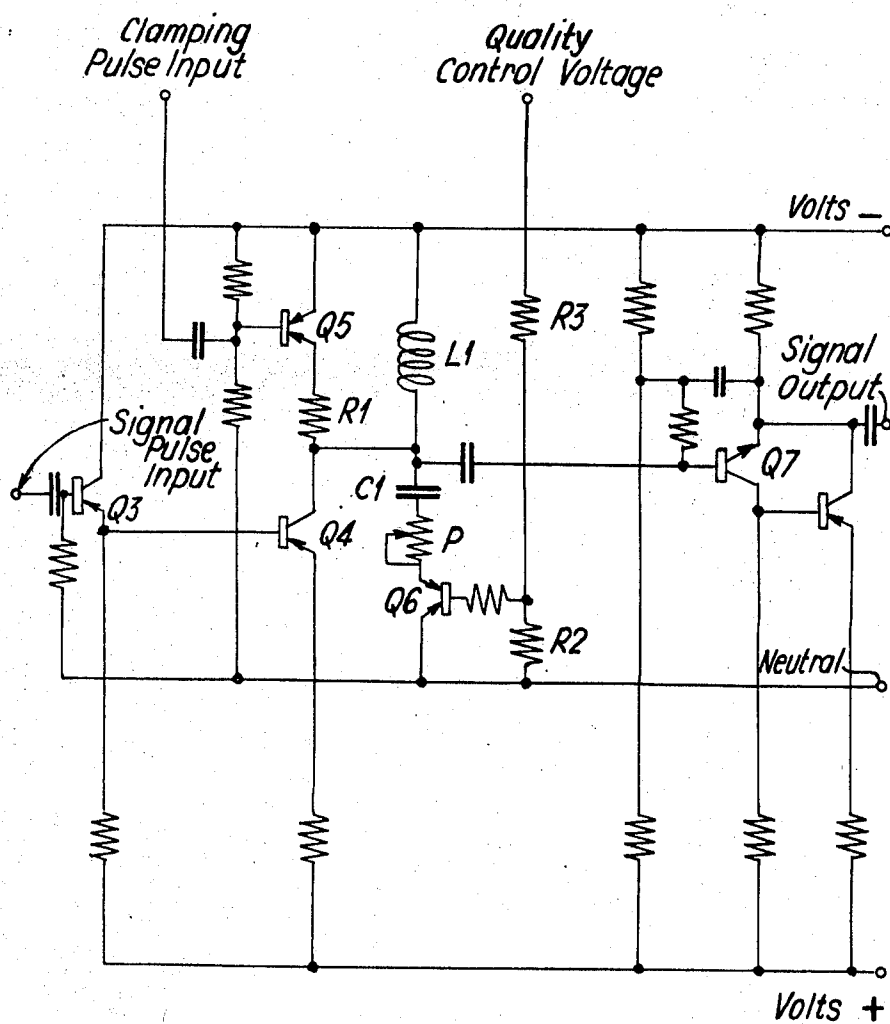


Fig. 2.

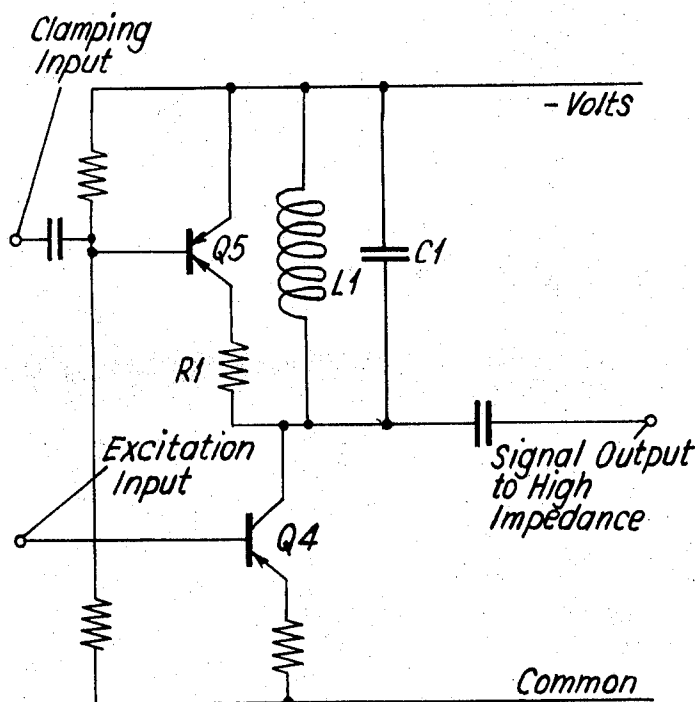


Fig. 3.

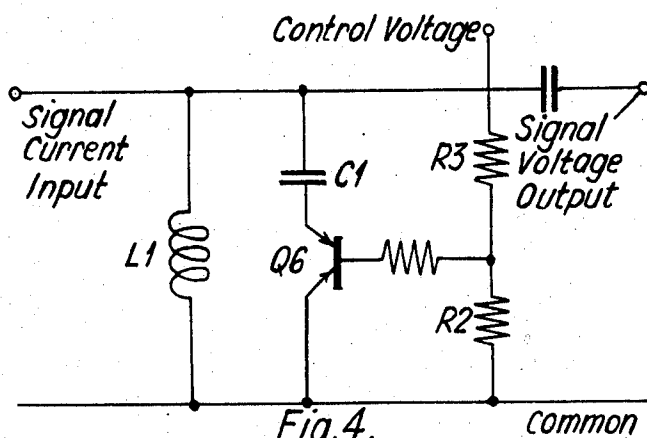


Fig. 4.

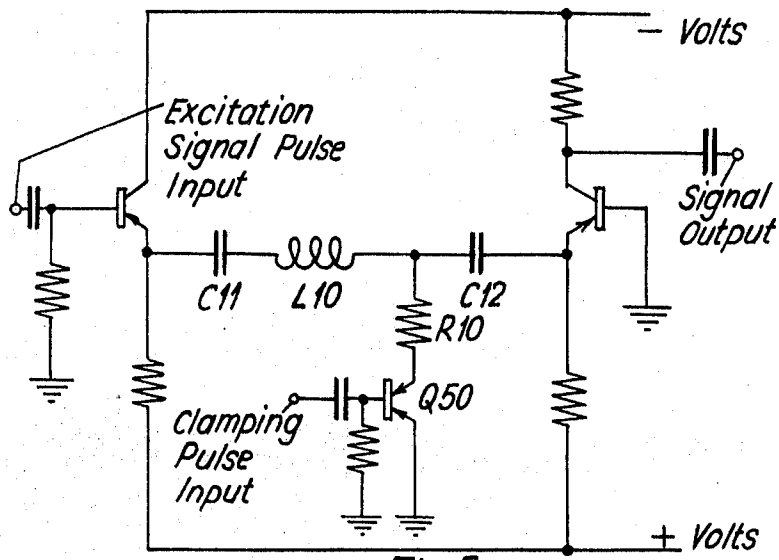


Fig. 5.

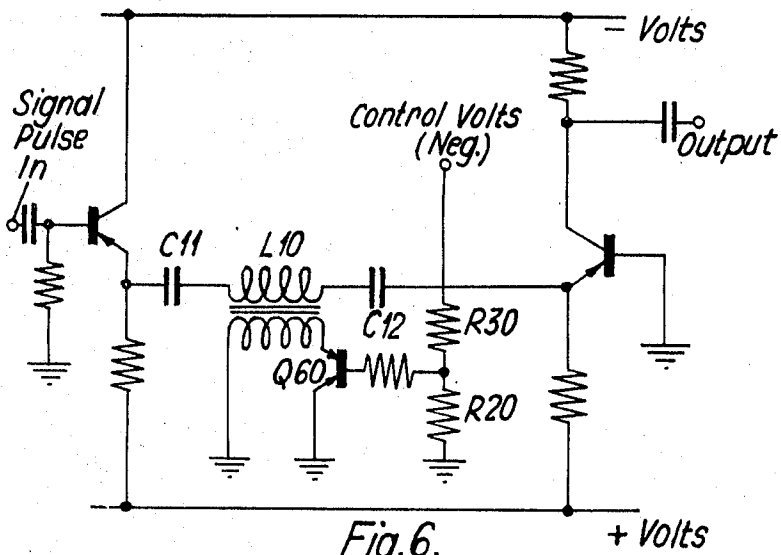


Fig. 6.

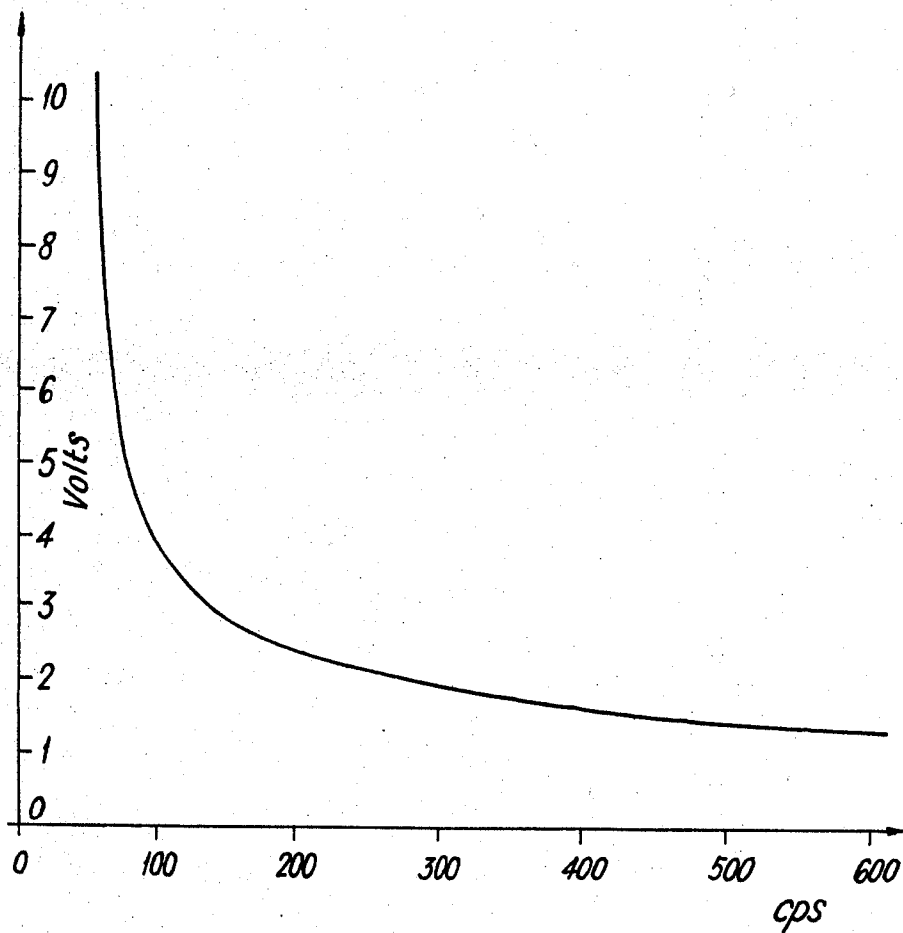
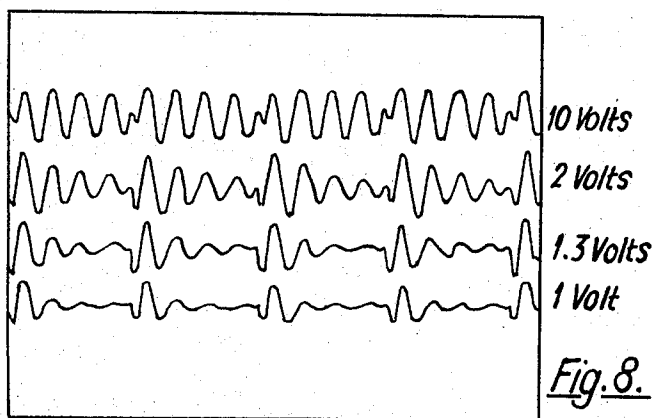
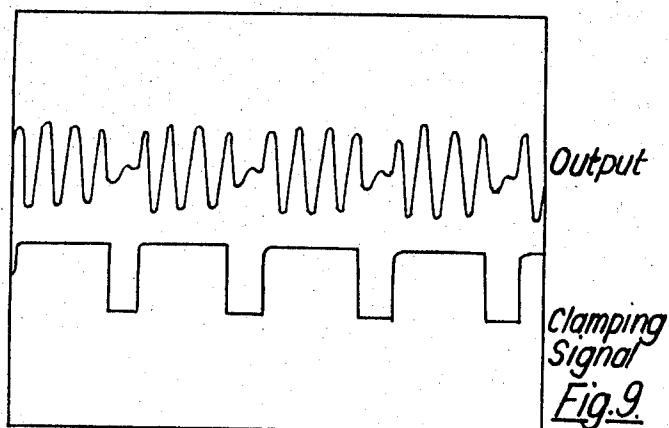
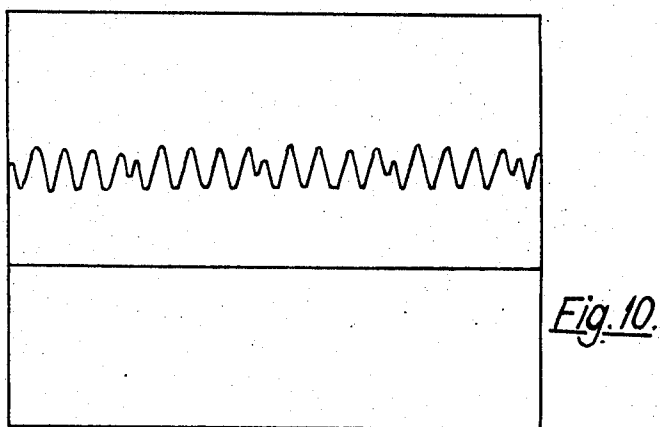


Fig. 7.



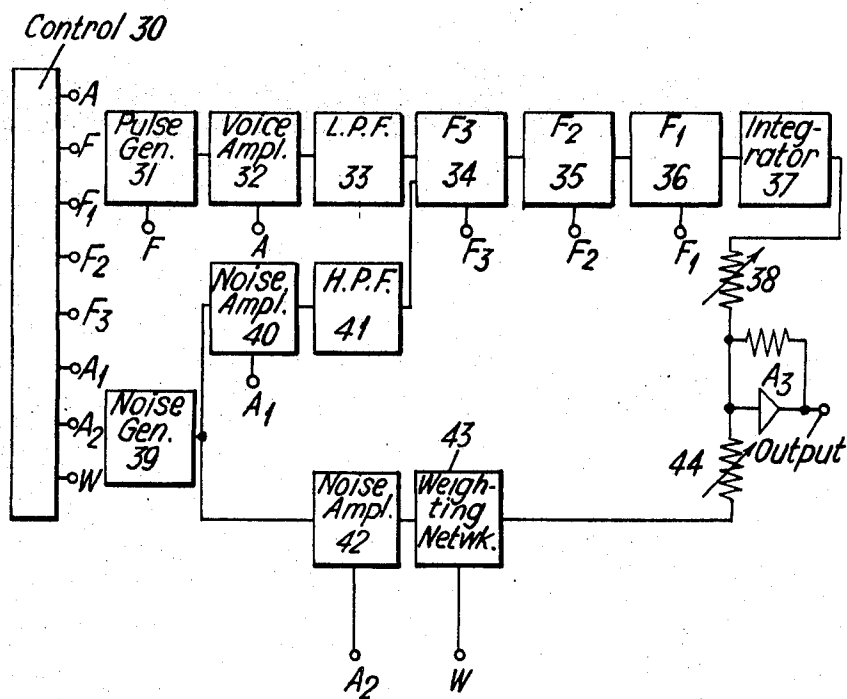


Fig. 11.

## VOCODER FILTER SYSTEM

The present invention relates to circuitry and methods for periodically sharply varying the quality factor of resonant circuits so as to allow the stored energy of the circuit to be dissipated and has particular though not exclusive application in a speech synthesizer or in the synthesizer stage of a vocoder.

A vocoder is a speech transmission system comprising an analyzer which produces control signals derived from a speech input, a transmission line for the said control signals, and a synthesizer.

The bandwidth required to transmit the control signals is less than that required to transmit the speech in analogue form. The control signals can be multiplexed and are not subject to the same kinds of distortion as are analogue speech signals. In a voice excited vocoder, for example, a base band covering a frequency range such as 80-600 c.p.s. is transmitted in analogue form and is added at the synthesizer to the output of the parallel frequency channels which cover the remaining part of the speech bandwidth. The base band is also subjected at the synthesizer to nonlinear distortion to provide the excitation signal for the stimulation of the frequency channels.

Other types of vocoder include, for example, the correlation vocoder and the formant tracking vocoder.

The synthesizer of many types of vocoder consists of a number of frequency channels in parallel, each channel containing a modulator and a band-pass filter. A signal whose pitch is that of the original speech and which contains sufficient harmonic content to cover the frequency range of interest is used to excite the channels. Commonly the above signal consists of a train of narrow pulses which in the case of a full channel vocoder are spaced at pitch rate and in the case of a voice-excited vocoder are spaced at a rate which is related to pitch rate but not necessarily equal to it. If the channel filters are sharply tuned (as is generally the case), they will ring after the application of each pulse for a time greater than the time between pulses. There will thus be energy remaining in the filters when the next pulse is applied, and depending on the relative phase of the damped oscillation it will interfere additively or otherwise with the response of the resonant circuit to the succeeding excitation signal.

The present invention is suitable for use in the synthesizer of a channel vocoder which is excited by a series of pulses and in the channel portion of the synthesizer of a voice-excited vocoder.

In vocoders which employ single tuned filters for the frequency channels in the synthesizer stage and wherein sample excitation pulses are applied to the filters once in every pitch period, the quality of the reproduced speech is degraded by the energy which is carried over in the filters from one pitch period to the next. If the quality factor  $Q$  of the filter as hereinafter defined is lowered sufficiently to avoid this energy storage, i.e. if the filters of the synthesizer are broadened sufficiently, then the speech quality is further degraded. We propose therefore to employ a technique which we refer to as "clamping" which consists of allowing the filter to be of as high a  $Q$  during the major portion of each pitch period as is required and considerably reducing the  $Q$  at the end of each pitch period for a short time immediately before the application of the next excitation pulse, thus ensuring that there is little or no energy remaining in the filter when it is again excited.

The invention provides a method of dissipating the energy stored in a resonant circuit in a vocoder synthesizer without giving rise to undesirable transients in the output.

The duration of the clamping signal is as short as possible while being sufficiently long to allow substantially all of the stored energy to be dissipated and the signal is timed so as to end at the application of the succeeding excitation pulse.

Various preferred embodiments of the invention include a number of variable bandwidth filters whose bandwidths are variable independently of one another or a number of filters the bandwidths of some or all of which are variable in concert.

The resonant frequency  $f$  in cycles per second of a parallel LC circuit is given approximately by

$$f = \frac{1}{2\pi\sqrt{LC}} \quad (1)$$

where  $L$  and  $C$  are the values in M.K.S. units of the inductance and the capacitance respectively.

The range of frequencies around the resonant frequency over which the response of the resonant circuit does not fall more than 3 decibels below the value of the output at resonance is called the bandwidth and is denoted by  $\delta f$ .

The quality factor  $Q$  of the circuit is defined by the formula

$$Q = \frac{f}{\delta f} \quad (2)$$

where  $f$  is the resonant frequency.  $Q$  is thus a measure of resolution.

$Q$  may be shown to satisfy the relation

$$Q = \frac{1}{R} \sqrt{\frac{L}{C}} \quad (3)$$

where  $R$  is the effective resistance present in the circuit.

From (1), (2) and (3) we have

$$\delta f = \frac{R}{2\pi L} \quad (4)$$

if  $C$  has been chosen to give the required  $f$ .

Thus the bandwidth is proportional to the overall resistance of the circuit. If the resistance is arranged to be variable in response to a control signal, we obtain a filter whose bandwidth is externally controllable.

For a mathematical analysis of resonant circuits, reference may be made, for example, to J. Millman and H. Taub's Pulse, Digital and Switching Waveforms; devices and currents for their generation and processing (McGraw-Hill 1965). Obviously the complete mathematical analysis of a series resonant circuit will differ from that of a parallel resonant circuit. However the techniques of clamping and of smoothly varying the bandwidth of a filter which are provided by the present invention have application both in series and in parallel tuned circuits.

In one version of the preferred embodiment of the invention, the bandwidth of a single tuned LC filter is altered by changing the effective resistance in the circuit.

A resonant circuit with positive damping is one in which the output decays with time, e.g. the output might be of the form  $\exp(-kt) \cdot \sin(t+c)$  where  $\exp$  denotes the exponential function,  $c$  is a constant,  $k$  is a positive constant called the decay factor, and  $t$  is time measured from the moment of application of the excitation pulse.

The circuit is said to be critically damped when the rate of energy decay is a maximum. In this case the energy is dissipated in the minimum possible time.

Any real resonant circuit has inherent damping due to resistive losses in the circuit elements, particularly in the inductance. Generally it is required to minimize this damping in order to realize the maximum possible selectivity, but if an additional resistance is switched in parallel with a branch of a resonant circuit, for example, the result will be critical damping of the circuit if the additional resistance has been selected appropriately.

According to the invention there is provided a circuit arrangement which includes a bank of resonant circuits resonating each at a different frequency and having quality factor control means whereby the bandwidth of one of the resonant circuits may be varied.

The invention resides in the method employed and in circuitry or apparatus which includes a synthesizer of the type described, and in a resonant circuit wherein critical damping is introduced by the method described.

In one embodiment of the invention, the critical damping resistance of the resonant circuit is introduced by being switched in electronically.

In a particular embodiment of the above paragraph a symmetrical transistor is in series with a resistor, a pulse of the appropriate duration being applied at the appropriate moment to the base of the transistor so as to switch into the resonant circuit the resistance of the transistor together with the resistance of the resistor.

The polarities of the power supply to the transistor, of the resonant circuit, of the excitation pulses and of the pulse applied to the base of the transistor are not necessarily independent of one another.

In a preferred embodiment of the invention, the bandwidths of the resonant circuits of the speech synthesizer are variable independently of one another, so that, for example, subjective tests can be made as to what settings give rise to most acceptable speech, enabling research into acceptableness and other characteristics of speech to be carried out.

The term "speech synthesizer" includes, for example, the synthesizer stage of a vocoder, a reading machine for the blind, and a parametric artificial talker.

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 diagram of a channel vocoder;

FIG. 2 is a circuit diagram of one form of filter for a vocoder synthesizer channel;

FIG. 3 is a circuit diagram which is equivalent to a part of FIG. 2;

FIG. 4 is a circuit diagram which is equivalent to a different part of FIG. 2;

FIG. 5 is a circuit diagram of a second form of filter for a vocoder synthesizer channel with clamping facilities;

FIG. 6 is a circuit diagram of a second form of filter for a vocoder synthesizer channel with bandwidth control;

FIG. 7 is a graph illustrating the variation in bandwidth in the circuits of FIGS. 2 and 4 for a range of applied quality factor control voltages;

FIG. 8 shows oscilloscope traces of output waveforms obtained for various control voltages in the circuit of FIGS. 2 and 3;

FIG. 9 is an oscilloscope trace showing a clamping signal and the output obtained when the clamping signal is applied;

FIG. 10 is an oscilloscope trace of the output obtained with zero clamping signal; and

FIG. 11 is a block diagram of a parametric speech synthesizer.

The definitions of some terms used in the following description will now be introduced.

A symmetrical transistor is one which is constructed in such a way that its characteristics remain substantially unchanged when the emitter and collector connections are interchanged.

A transistor is said to be saturated when the difference in potential between collector and emitter is less than that between base and emitter. When this occurs, the transistor undergoes bottoming, that is to say the output is determined by the characteristics of the transistor and not by the applied signal.

A symmetrical transistor operating in the saturated mode has an effective emitter-collector resistance  $R_T$  given by

$$R_T = \frac{2}{I_b(\beta + 1)\lambda} \quad (5)$$

where  $I_b$  is the base current,

$\beta$  the common emitter current gain, and

$\lambda$  is typically 40 per volt at 16°C.

For a typical transistor of the kind utilized as Q5 or Q6 in the circuits of FIGS. 2, 3, and 4, or as Q50 in FIG. 5 or as Q60 in FIG. 6,  $R_T$  is about 10 ohms when  $I_b$  is 100  $\mu$  A.

FIG. 1 illustrates a so-called channel vocoder having an analyzer stage from which signals are derived and carried by a number of transmission paths, one of which, 10, carries the fundamental frequency information of the speech, while the remainder convey the spectral information. These signals control the synthesizer stage which reconstructs the original speech. In more detail, a speech signal is picked up by a microphone 11 and amplified by an amplifier 12. The output of the amplifier 12 is fed to a pitch detector 16 which detects whether the fundamental voice frequency is random or

periodic. From 16 the pitch information goes to a frequency-to-voltage converter 17. The output of 17 is fed through a low-pass filter 18 and carried by the pitch channel 10 (which thus conveys the fundamental frequency or pitch information of the speech, if any) to a switch 19 and a pitch generator 20. If the pitch information is zero, a hiss generator 21 is switched on. If the pitch information exceeds a threshold value, "voicing" in the 80—300 c.p.s. range is provided.

The output of the amplifier 12 goes also to a pre-emphasis equalizer 13 and thence to a bank of analyzer spectrum filters ASF whose outputs are rectified by the bank of rectifiers RECT and then applied to the bank of low-pass filters L1—L9. The signals carried by channels 1—9 thus convey the spectral information of the speech. The outputs of channels 1—9 are applied to the modulators M1—M9 where they modulate the pitch or hiss signal from 20 or 21 respectively. The modulated outputs of M1—M9 pass through a bank of synthesizer spectrum filters SSF to be combined at the deemphasis equalizer 14, whence the reconstituted speech signal goes to an amplifier 15, being output in audible form at the headphones 22.

The technique of switching in of critical damping can be utilized in some or all of the synthesizer spectrum filters SSF.

The bandwidths of the analyzer spectrum filters ASF can be varied independently of one another or in concert.

If the synthesizer spectrum filters are of the type shown in FIG. 2, then variation of bandwidth can be achieved in filters which can also be critically damped between successive excitation pulses.

FIG. 2 (and FIGS. 3 and 4, which are different portions of FIG. 2, corresponding respectively to the clamping facility and to the control for smoothly varying the bandwidth) illustrates one spectral filter and some associated circuitry from the filter bank SSF of the synthesizer stage of the vocoder of FIG. 1, the filter consisting of a resonant circuit. A series of excitation pulses spaced at voice fundamental rate are generated (by 20 of FIG. 1, from information provided by 16, 17 and 18) and applied to the synthesizer where they excite the tuned circuits whose outputs are combined to give an analogue response which is converted to audio form by an electromechanical transducer.

The parallel resonant circuit consists of an inductor L1 and a capacitor C1 in the collector circuit of Q4. Excitation pulses are applied to the base of Q4. The signal output is taken to a high impedance (Q7 of FIG. 2, not shown in FIG. 3) to avoid loading the resonant circuit. Clamping is accomplished by placing in parallel with L1 and C1 a resistor R1 which is in series with a symmetrical transistor Q5. Q5 is normally biased to be nonconducting, but the application of a negative pulse of at least 10 volts amplitude turns it hard on and places the low resistance R1 across the resonant circuit thus dissipating the energy stored therein. The clamping pulse is timed to be just long enough to allow all the energy to be dissipated and to end just before the application of the next excitation pulse.

The capacitor C1 is taken to earth through a variable resistance P and a symmetrical transistor Q6.

The variable resistance enables the minimum bandwidth to be preset. The transistor is used as a variable resistance, and so must always be kept saturated. This means that only small signals can be handled and so the gain of the stage is kept low by making the emitter resistance of Q4 large, say 120 kilohms.

A voltage applied for quality control delivers a current through the potential divider formed by R2 and R3 to the base of the transistor Q6. For an applied negative voltage in the range 10 volts to 1.5 volts the bandwidth is in the range 60 c.p.s. to 500 c.p.s.

The symmetrical PNP transistor Q6 is used as a variable resistance by keeping it saturated, in which case the collector current is independent of the base current.

Then the bandwidth  $\delta f$  for a given frequency  $f$  is given by equation (4) where R is the total resistance of the circuit, and so

$$\delta f = R_L + \frac{2}{I_B(\beta \neq 1)\lambda} \quad (6)$$

(since  $R = R_L + R_T$  where  $R_L$  is the resistance of the inductor).

$$\delta f = A + \frac{B}{I_B} \quad (8)$$

where A and B are constants. Thus a parabolic relationship holds between  $\delta f$  and  $I_B$  over a certain range.

(The ratio  $R_3/R_2$  is approximately 10/1 in order to obtain in a convenient way the range of variation in transistor base current from Q6.)

FIGS. 5 and 6 illustrate as an alternative a series resonant form of resonant circuit for the synthesizer stage of a vocoder and having respectively clamping facilities and a control for smoothly varying the filter bandwidth. R10, R20, R30, L10, Q50 and Q60 correspond to R1, R2, R3, L1, Q5 and Q6 of FIGS. 2-4. The effective capacitance C10 of the circuit of FIGS. 5 and 6 is given by

$$\frac{1}{C10} = \frac{1}{C11} + \frac{1}{C12} \quad (9)$$

Application of a clamping pulse input to the base of Q50 in FIG. 5 results in critical damping being introduced across the resonant circuit for the duration of the pulse.

The duration of the pulse is chosen to be such that all the energy stored in the circuit is dissipated in as short a time as possible in the interval between the application of successive excitation pulses and the timing is such that the clamping pulse ends before the application of the next excitation pulse.

In FIG. 6 the inductance L10 is seen to be provided by the primary winding of a transformer. Variation in bandwidth is achieved by applying a negative control voltage through the potential divider formed by the resistance R20 and R30 to the base of a saturated symmetrical transistor Q60.

In FIGS. 2-6 the symmetrical transistors have been PNP transistors. NPN transistors can be used throughout, (though symmetrical NPN transistors are not readily commercially available at present), appropriate changes e.g. in the polarity of the power supplies being made in the circuits. The excitation pulse of FIGS. 2-6 and the clamping pulse of FIGS. 2, 3 and 5, are negative going signals. The control voltage of FIGS. 2, 4 and 6 is a negative voltage. These require to be positive if the polarities are reversed. Field effect transistors can be used in place of symmetrical PNP transistors, with appropriate modifications in the circuits.

FIG. 7 illustrates the control characteristic obtained by applying a quality control voltage as described with reference to FIGS. 2 and 4. The curve is seen to vary smoothly over the range 100 c.p.s. to 300 c.p.s. which is commonly of most utility in vocoder application.

FIG. 8 illustrates output waveforms obtained by repetitive stimulation of the tuned circuit for various values of the control voltage (equivalently, for various bandwidths).

Referring again to FIGS. 2, 3 and 5, "clamping" is achieved by placing the critical damping resistance in series with a symmetrical transistor Q5 across the inductor. The base of Q5 is biased so that normally it is turned off, but the application at the input labeled "clamping pulse" of a negative pulse (of 10 volts amplitude for the filter of FIGS. 1 and 2) turns it hard on and the oscillations of the tuned circuit die away rapidly. The time required for all the energy to be dissipated depends on the resonant frequency, being greater for the lower frequencies. For the lowest frequency channels, about 400  $\mu$ sec. is required, but as this may represent about 20 percent of the time between excitation pulses some compromise will be needed. Normally the "clamping" pulse should end at or just after the commencement of the excitation pulse.

The upper trace of FIG. 9 illustrates the response of a single filter of resonant frequency 1,000 c.p.s. which is excited by a train of pulses whose repetition rate is constant at 250 p.p.s. The lower trace, which is to be interpreted as a relatively short negative pulse of amplitude 10 volts and not as a long positive pulse illustrates the form of the "clamping" signal which is ap-

plied to the base of Q5 so as to prevent energy from being carried over between one excitation period and the next.

FIG. 10 shows the response of the same filter to the same pulse train but with the clamping signal removed, and shows the reduced response due to destructive interference between the oscillations of successive pitch periods.

FIG. 11 illustrates a synthesizer which produces artificial speech. A control unit 30 produces control signals, the control unit itself being governed by, for example, a digital computer with paper tape output, or by a voltage reader which reads frequency and spectral information presented as manually prepared patterns of electrically conducting ink on rolls of paper. For the production of vowels, a fundamental frequency signal F is applied to a pulse generator 31, the output of which is applied to a voice amplitude circuit 32 whose operation is controlled by an amplitude signal A and whose output is fed through a low-pass filter 33 and thence through resonant circuits 34 (F3), 35 (F2) and 36 (F1) to an integrator 37, and thence through a network which includes a variable resistance 38 to an amplifier A3, the output being an audible signal. For the production of "voiced" consonants and front fricatives, a noise generator 39 feeds a first noise amplitude circuit 40 which is controlled by the parameter A1. The output of 40 passes through a high pass filter 41 and thence goes to the resonant circuits 34 (F3), 35 (F2) and 36 (F1), passing next to the integrator 37 and the variable resistance 38 and finally to the output amplifier A3. For the production of other consonants, including, for example, the back fricatives, the output of the noise generator 39 is fed to a second noise amplitude circuit 42 which is controlled by the parameter A2 and thence to a weighting network 43 whose performance is controlled by the W parameter. The output of the weighting network is fed through a variable resistor 44 and is then output as an audible signal from the amplifier A3.

The above description of a parametric artificial talker is somewhat simplified.

Essentially 36, 35 and 34 mirror the first three formants of natural speech i.e. they resonate at around the natural frequencies of the vocal tract. In the synthesis of an utterance, they are modified in accordance with the alteration in shape which the human vocal tract would undergo during speech production.

36-34 can be connected in parallel rather than in series if appropriate modifications are made to the rest of the circuitry of FIG. 11.

The term "voice" as used in the preceding paragraph means the periodicity superimposed on certain speech sounds by the vibration of the vocal chords. The term "noise" means the turbulent waveform produced by the flow of air in the constricted vocal tract. "Front fricatives" include, for example, the English labiodental fricatives /f/ and /v/ and the alveolar fricatives /s/ and /z/. Back fricatives include the English palatal

fricatives  $\left| \frac{V}{S} \right|$  and  $\left| \frac{V}{Z} \right|$  and the glottal spirant /h/.

The facility of smoothly varying the bandwidth of a tuned circuit has application in the filters 34 (F3), 35 (F2) and 36 (F1).

Variation of the bandwidth of a filter during the production of a single phoneme may be desirable in order to reproduce adequately, for example, the English bilabial nasal /m/, the alveolar nasal /n/ and the velar nasal /ng/ of English "sing."

The facility of dissipating energy in the interval between successive excitation pulses also has application in the filters 34 (F3), 35 (F2) and 36 (F1).

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. A speech synthesizer which includes a plurality of resonant circuits, each having inherent damping, each resonating at different frequencies, excitation means associated with each of said resonant circuits for repeatedly applying momen-

tary excitation pulses to the associated circuit thereby causing damped oscillations therein, and means for selectively controlling said excitation means to cause the amplitude of the oscillations in any one of the circuits to decay to an arbitrarily low value before the application of the next successive excitation pulse to that same circuit.

2. A speech synthesizer as claimed in claim 1, wherein each resonant circuit has a quality factor, and said excitation control means comprises means for lowering the quality factor of each of the resonant circuits for a portion of the time between the application of successive excitation pulses.

3. A speech synthesizer as claimed in claim 2, including resistance means associated with each resonant circuit for damping the associated resonant circuit, and means for switching a resistance means into the associated resonant circuit to give critical damping in the resonant circuit for some discrete time after the application of one excitation pulse, and means for removing said resistance means from said resonant circuit before the application of the next succeeding excitation pulse.

4. A speech synthesizer as claimed in claim 1, wherein one of the resonant circuits includes an inductance in parallel with a capacitance.

5. A speech synthesizer as claimed in claim 1, wherein one of the resonant circuits includes an inductance in series with a capacitance.

6. A speech synthesizer as claimed in claim 3, further in-

cluding an inductor and transistor connected in parallel, and means for connecting the critical damping resistance between said inductance and the emitter of said transistor.

7. A speech synthesizer as claimed in claim 6, wherein the transistor is symmetrical, and wherein there are means for applying a pulse of predetermined duration and timing to the base of the transistor to switch it on.

8. A speech synthesizer as claimed in claim 1, and further including means for varying the bandwidths of the resonant circuits independently of one another.

9. A speech synthesizer as claimed in claim 1, and further including an inductance in one of the resonant circuits adjacent one winding of a transformer.

10. A speech synthesizer as claimed in claim 9, further including a variable resistance in series with another winding of the transformer.

11. A speech synthesizer as claimed in claim 10, wherein the said variable resistance is a transistor.

12. A speech synthesizer as claimed in claim 11, wherein the transistor is a symmetrical PNP transistor which is kept saturated and an external control signal is applied as a potential to the base of the transistor.

13. A speech synthesizer as claimed in claim 1, wherein said resonant circuit has associated therewith a variable resistance for connection into said resonant circuit to preset the bandwidth of the filter.

30

35

40

45

50

55

60

65

70

75