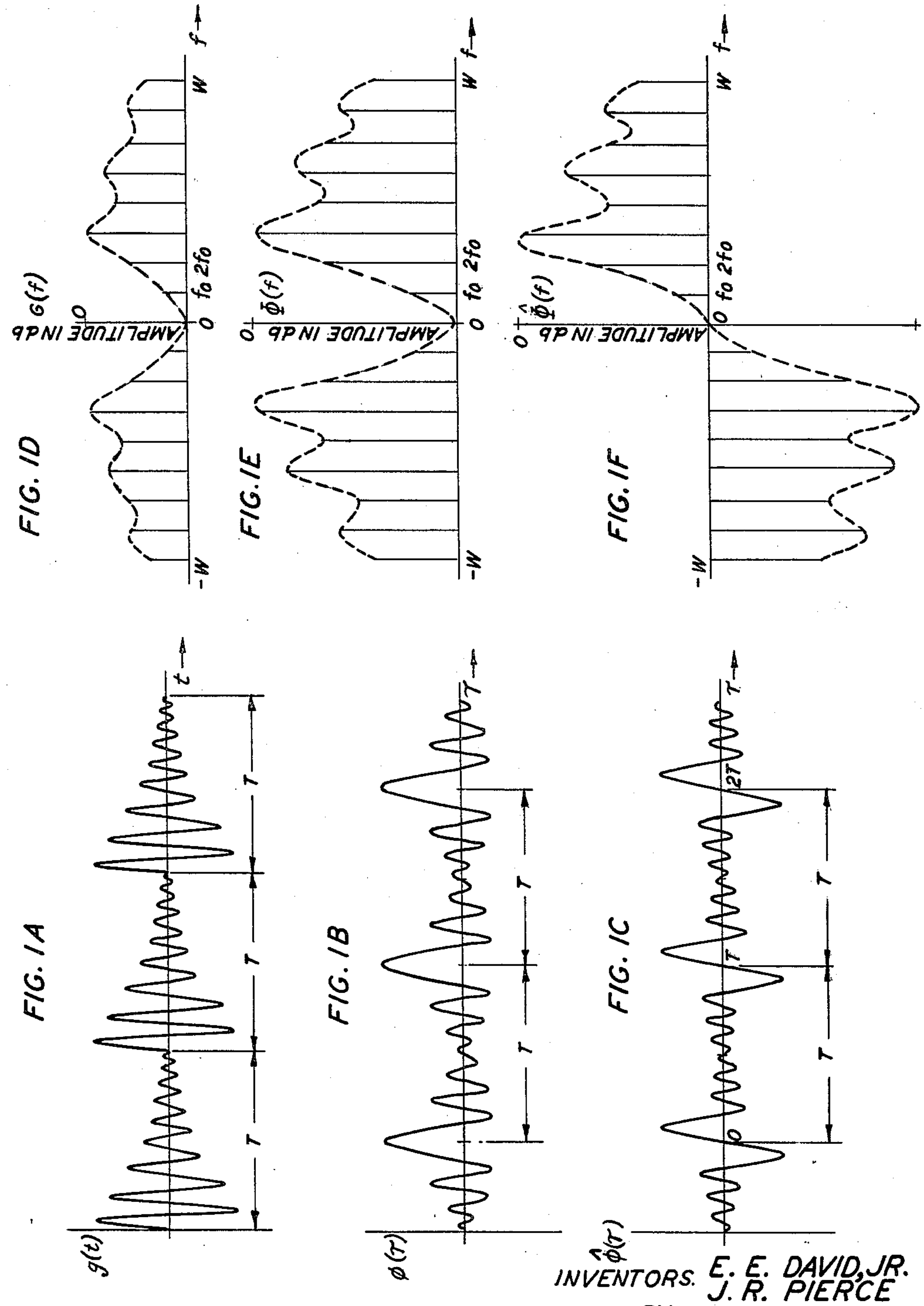
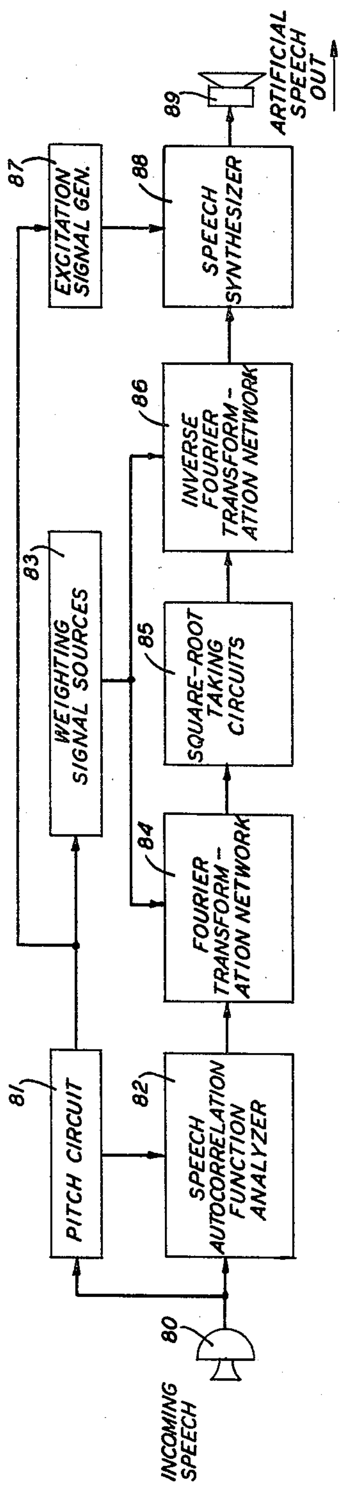
PITCH SYNCHRONOUS AUTOCORRELATION VOCODER

Filed Aug. 9, 1960 11 Sheets-Sheet 1



INVENTORS. E. E. DAVID, JR.
J. R. PIERCE
BY
C. E. Hirsch Jr.
ATTORNEY

FIG. 1G



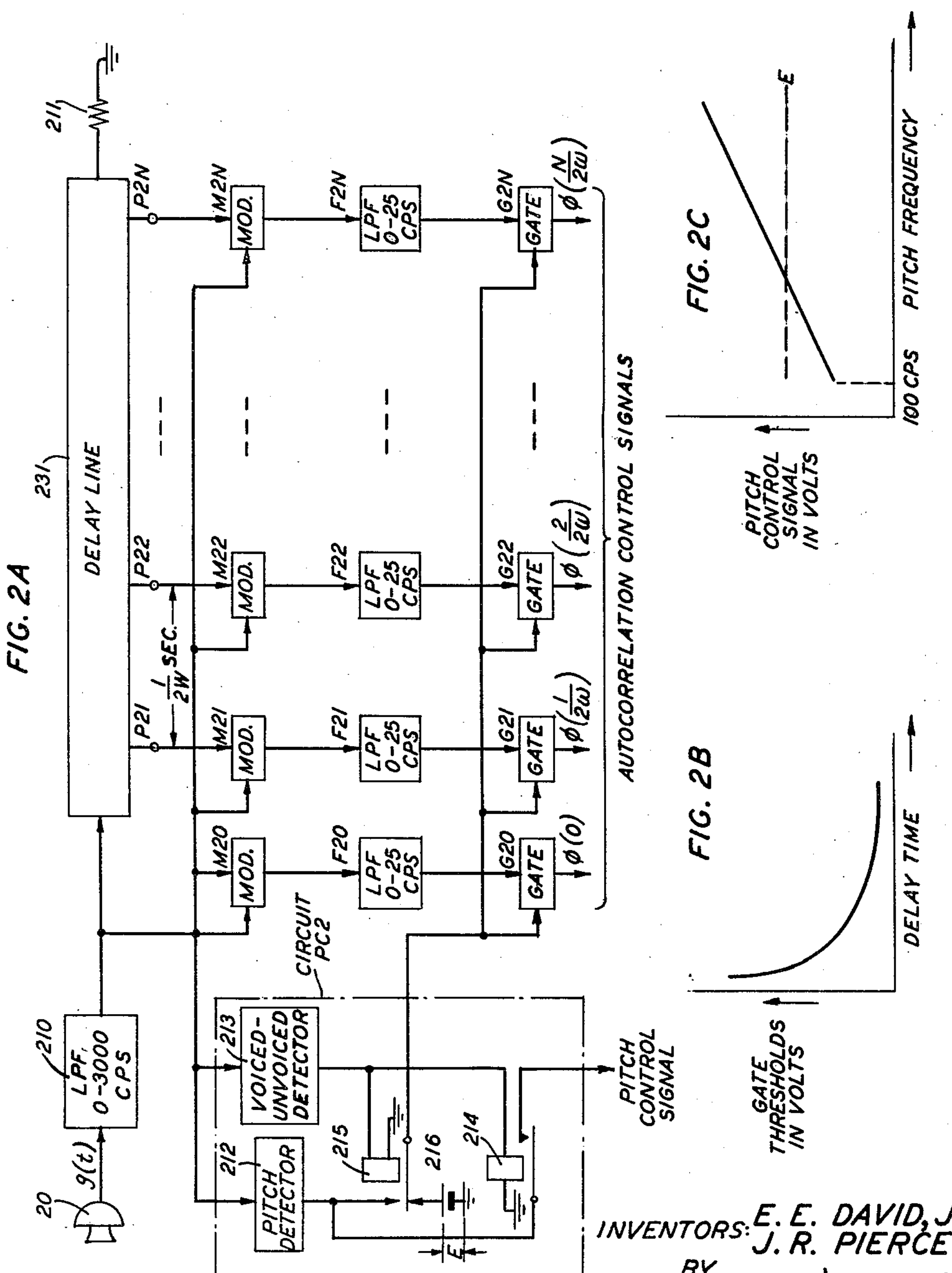
INVENTORS: E. E. DAVID, JR.
J. R. PIERCE

BY A. E. Hirsch Jr.

ATTORNEY

PITCH SYNCHRONOUS AUTOCORRELATION VOCODER

Filed Aug. 9, 1960 11 Sheets-Sheet 3



INVENTORS: E. E. DAVID, JR.
J. R. PIERCE

BY G. E. Hirsch, Jr.

ATTORNEY

PITCH SYNCHRONOUS AUTOCORRELATION VOCODER

Filed Aug. 9, 1960 11 Sheets-Sheet 4

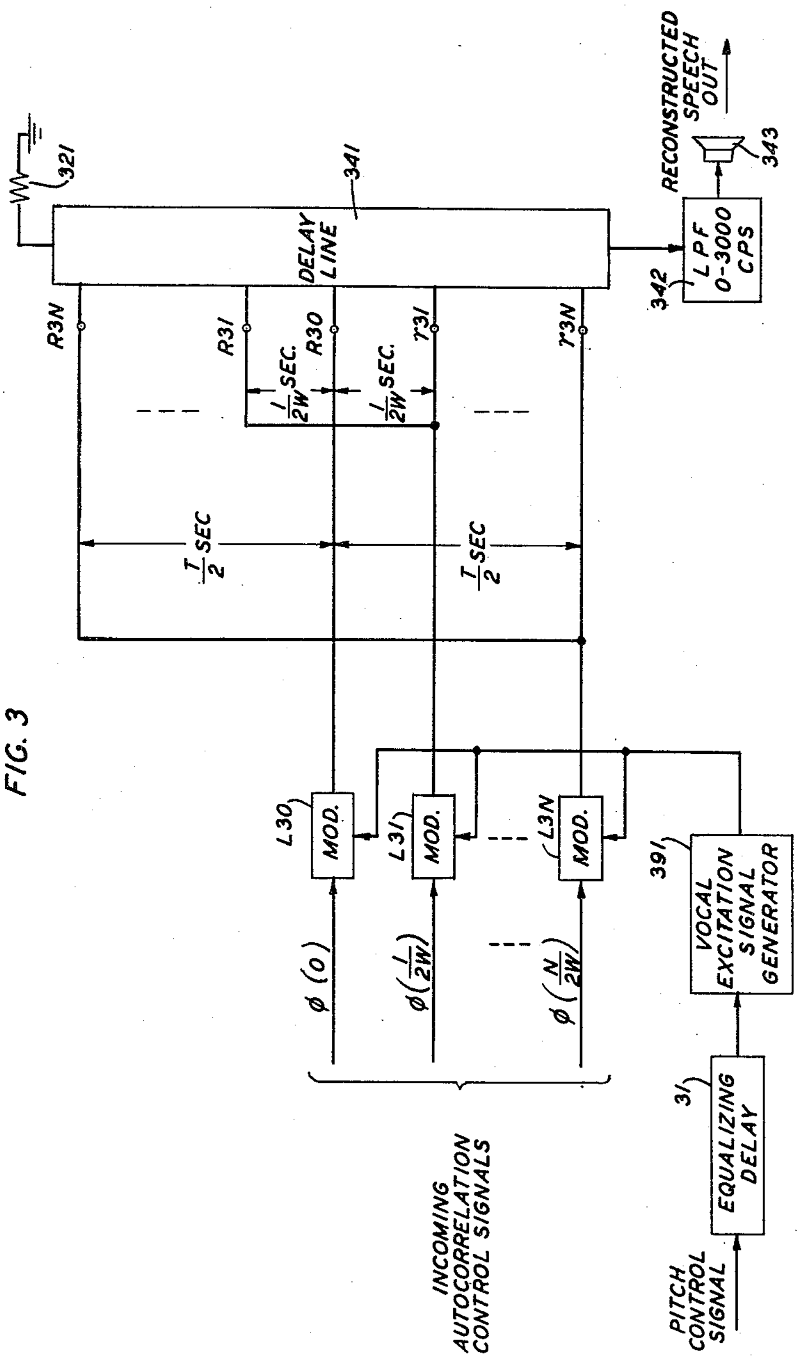


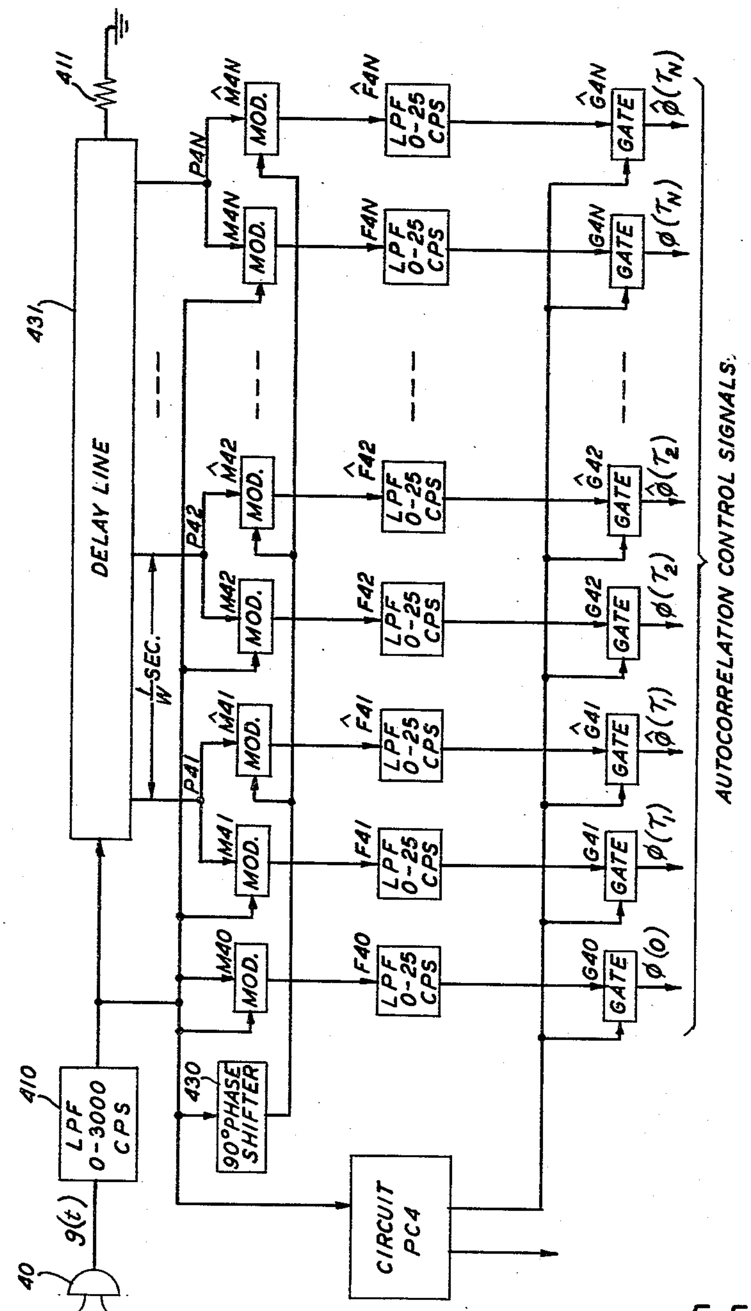
FIG. 3

INVENTORS: E. E. DAVID, JR.
J. R. PIERCE

BY A. E. Hirsch Jr.

ATTORNEY

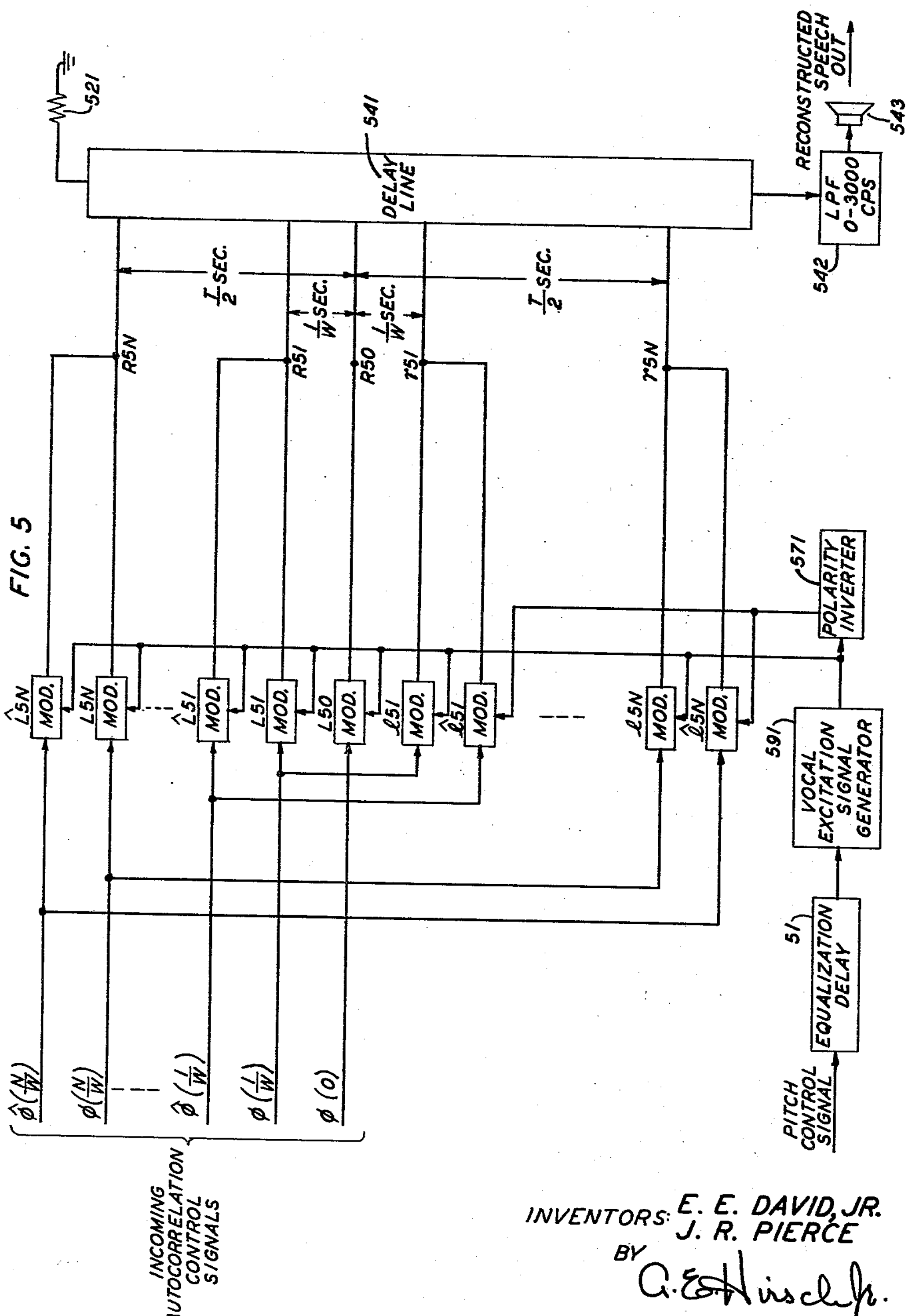
FIG. 4



INVENTORS: E. E. DAVID, JR.
J. R. PIERCE

BY A. E. Hirsch, Jr.

ATTORNEY



INVENTORS: E. E. DAVID, JR.
J. R. PIERCE
BY A. E. Hirsch Jr.
ATTORNEY

PITCH SYNCHRONOUS AUTOCORRELATION VOCODER

Filed Aug. 9, 1960

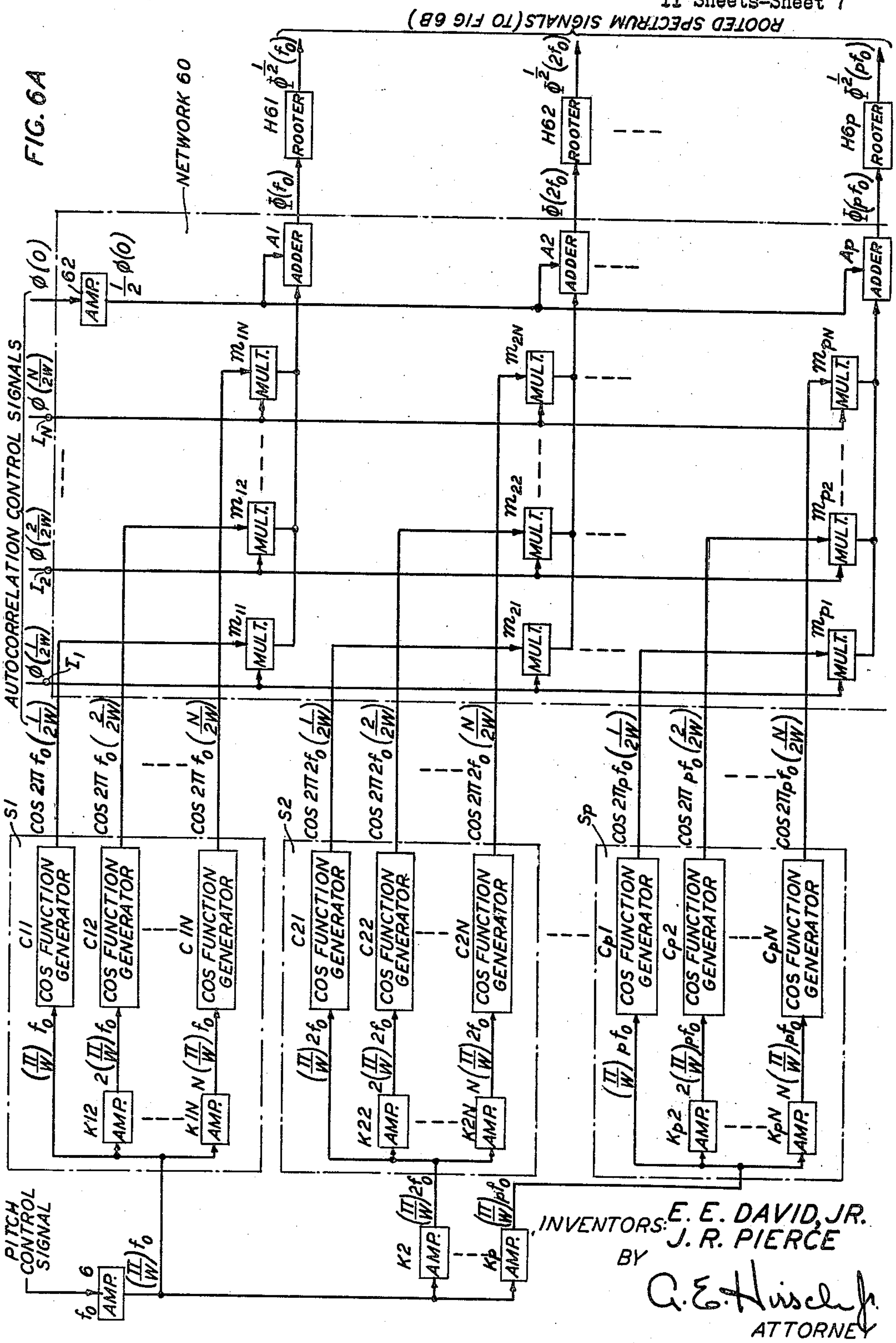


FIG. 6A

INVENTORS: E. E. DAVID, JR.
J. R. PIERCE

BY G. E. Hirsch, Jr.

ATTORNEY

Oct. 29, 1963 E. E. DAVID, JR., ET AL 3,109,070

PITCH SYNCHRONOUS AUTOCORRELATION VOCODER

Filed Aug. 9, 1960 11 Sheets-Sheet 8

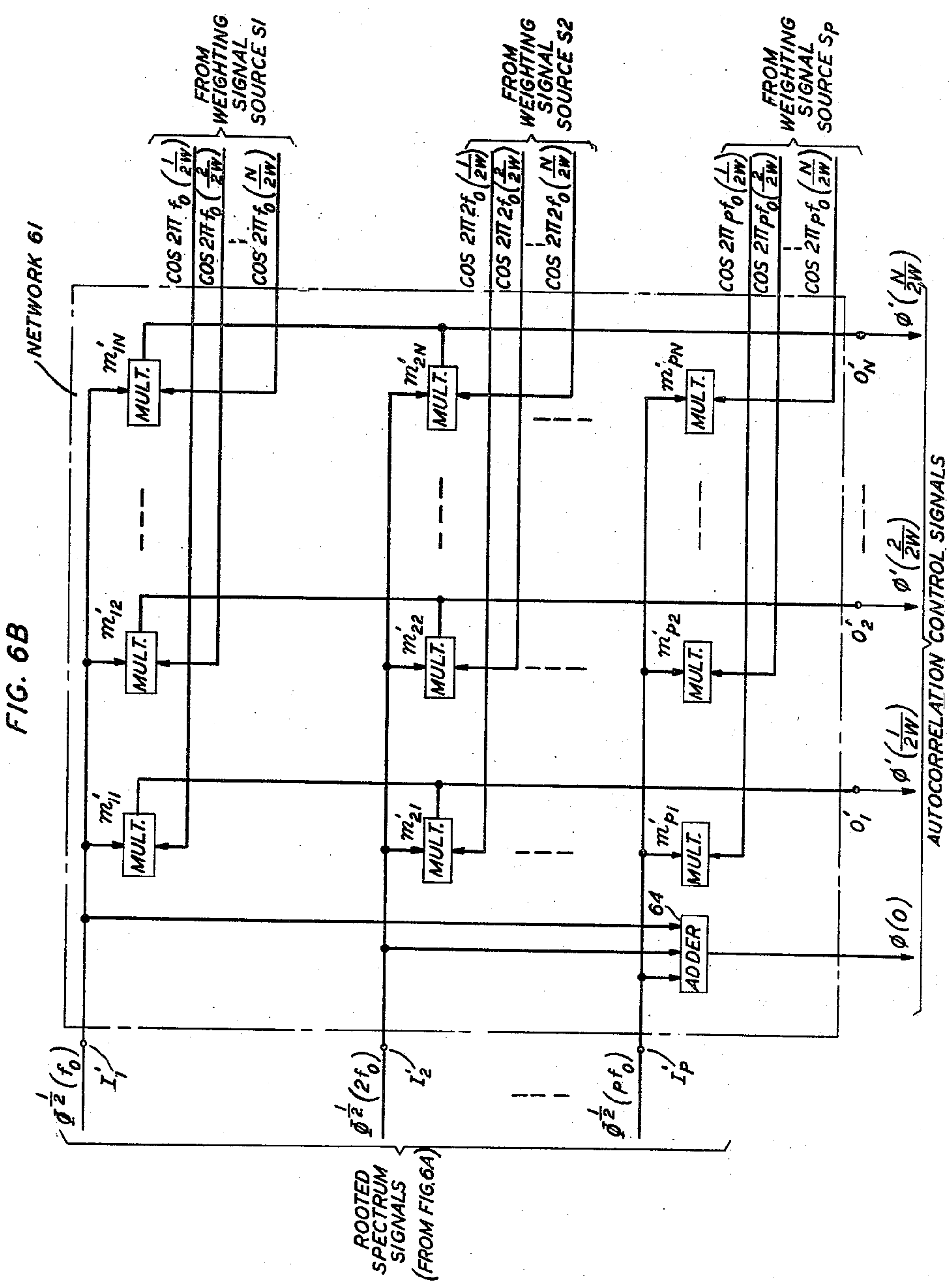
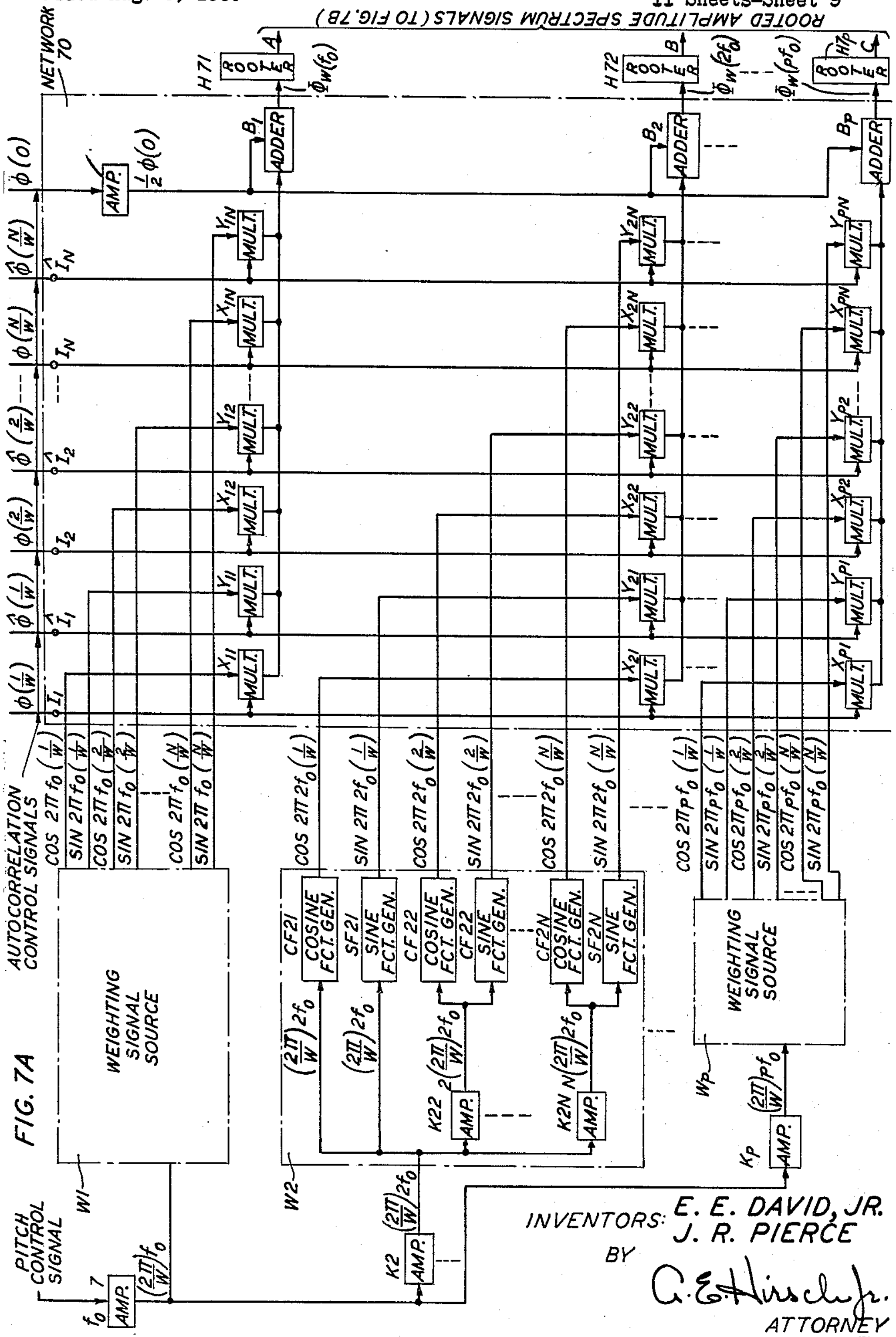


FIG. 6B

INVENTORS: E. E. DAVID, JR.
J. R. PIERCE

BY G. E. Hirsch, Jr.

ATTORNEY



INVENTORS: E. E. DAVID, JR.
J. R. PIERCE
BY
G. E. Hirsch, Jr.
ATTORNEY

PITCH SYNCHRONOUS AUTOCORRELATION VOCODER

Filed Aug. 9, 1960

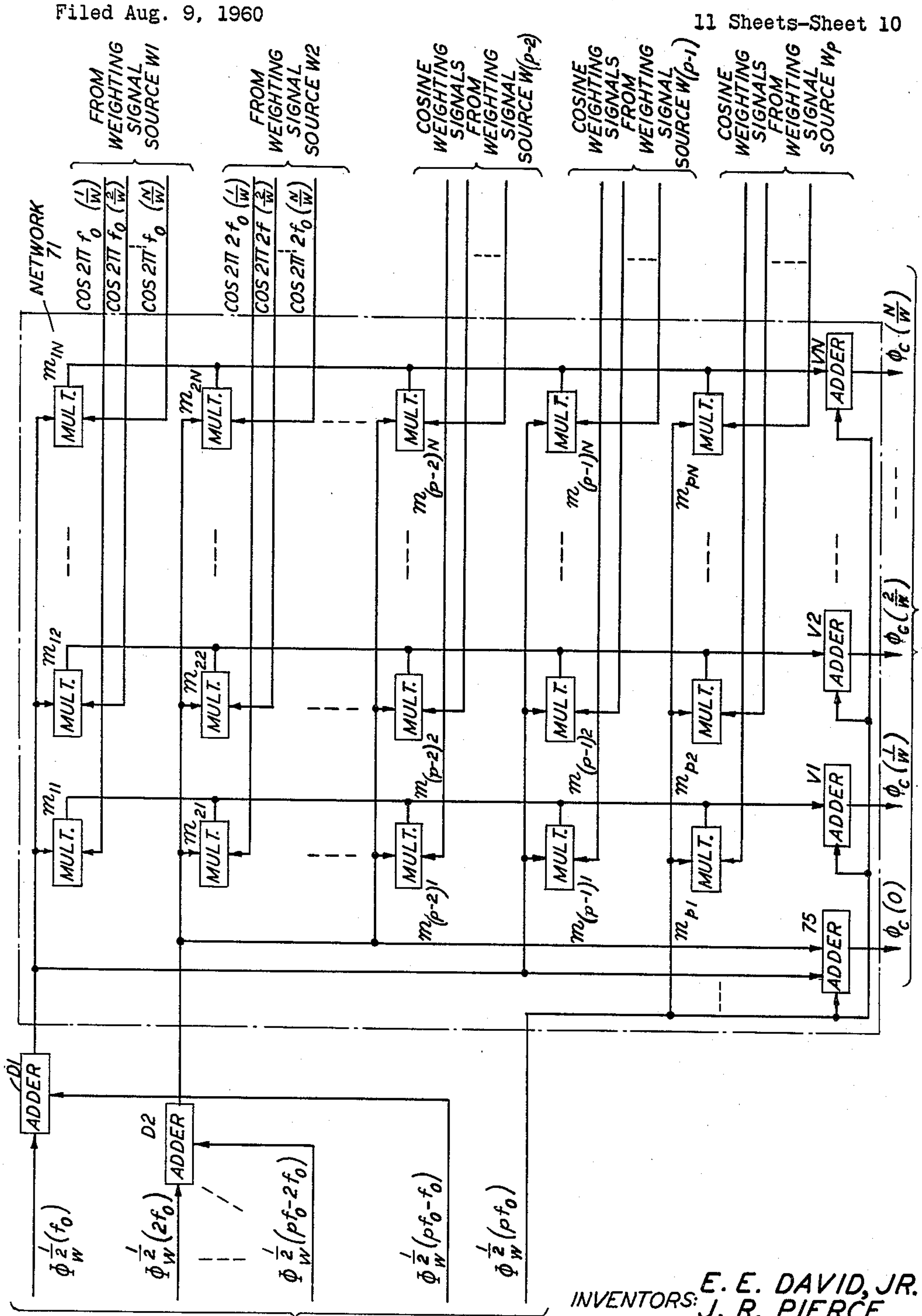


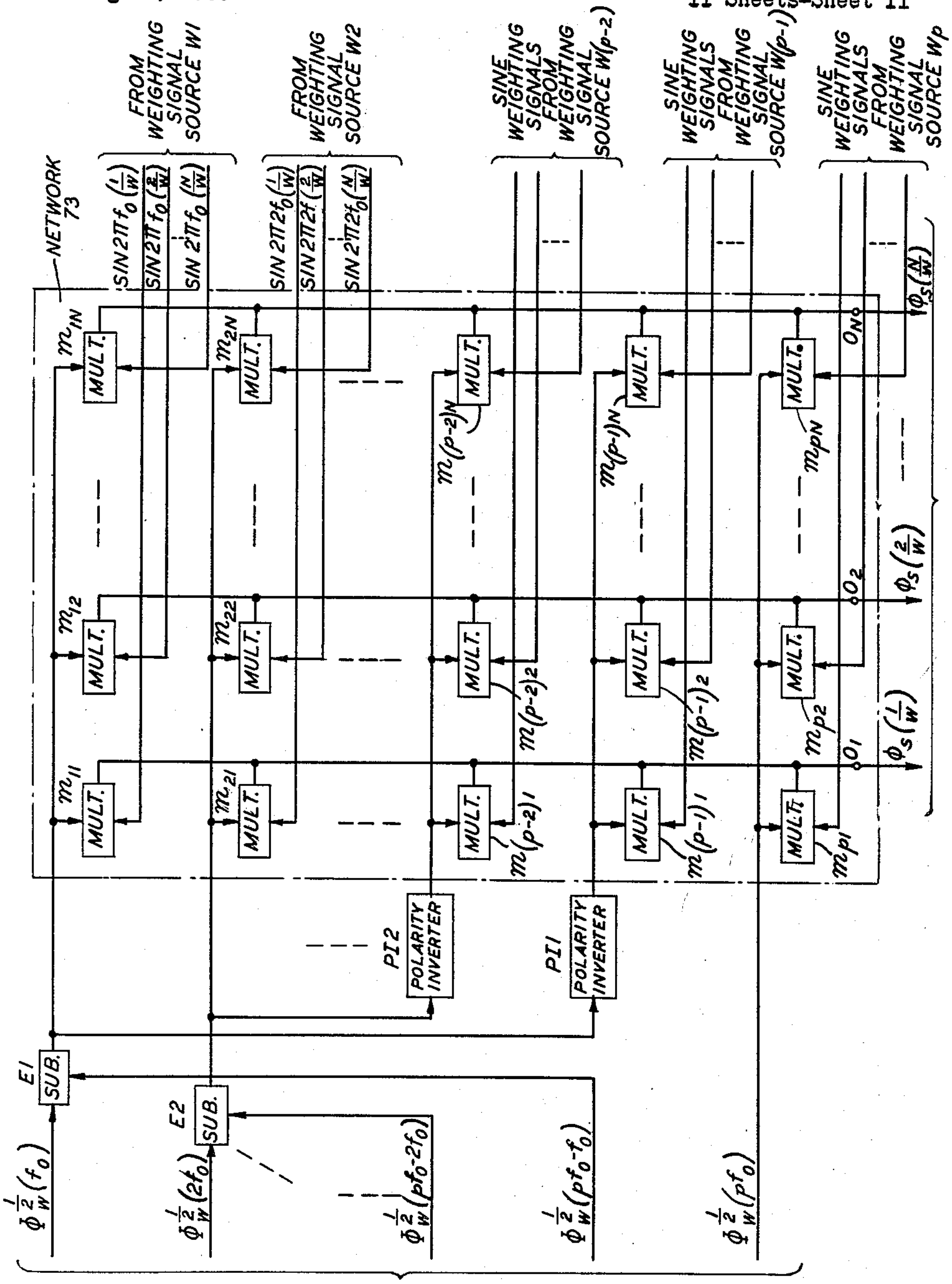
FIG. 7B

INVENTORS: E. E. DAVID, JR.
J. R. PIERCE
BY
A. E. Hirsch, Jr.
ATTORNEY

PITCH SYNCHRONOUS AUTOCORRELATION VOCODER

Filed Aug. 9, 1960

FIG. 7C



ROOTED AMPLITUDE SPECTRUM SIGNALS FROM FIG 7A

INVENTORS: E. E. DAVID, JR.
J. R. PIERCE

BY C. E. Hirsch Jr.

ATTORNEY

3,109,070
PITCH SYNCHRONOUS AUTOCORRELATION VOCODER

Edward E. David, Jr., and John R. Pierce, Berkeley Heights, N.J., assignors to Bell Telephone Laboratories, Incorporated, New York, N.Y., a corporation of New York

Filed Aug. 9, 1960, Ser. No. 48,423
12 Claims. (Cl. 179—15.55)

This invention relates to speech transmission systems, and particularly to the transmission of speech over narrow-frequency band channels in terms of autocorrelation functions.

Among speech coding systems for the conservation of transmission channel bandwidth, one of the best known is the so-called channel vocoder described in H. W. Dudley Patent 2,151,091, issued March 21, 1939. At the transmitter terminal of the channel vocoder, the amplitude spectrum of an incoming speech wave is divided into frequency bands by a group of band-pass filters, and the energy contained within each band is represented by the magnitude of a low-frequency control signal. After transmission over a reduced bandwidth channel to a receiver terminal, the control signals adjust the energies of frequency bands of an excitation spectrum generated at the synthesizer. The energy-adjusted frequency bands are then combined to produce synthetic speech.

Synthetic speech produced by the Dudley vocoder is distorted by the inherent limitations of the band-pass filters that it employs. For ideal reproduction of speech, the speech amplitude spectrum should be represented by a group of points; as practiced by the Dudley vocoder, however, the finite widths of the band-pass filters produce "points" that are in fact averages of many points within the spectral bands passed by the filters.

It is a specific object of this invention to eliminate the need for band-pass filters and to reduce distortion in synthesized speech by representing speech in terms of nearly ideal points on its autocorrelation functions.

In this invention, an incoming speech wave is first band-limited and then correlated with itself at the analyzer to obtain a number of samples of each period of the speech autocorrelation function. These samples of the autocorrelation function constitute a group of control signals. Each control signal occupies a relatively narrow-frequency band because the autocorrelation function changes very little from period to period; further, the entire group of control signals occupies a much smaller band of frequencies than does the original speech wave, thereby yielding a substantial saving in transmission channel bandwidth. At the synthesizer, the control signals adjust the amplitude of an excitation signal generated at the synthesizer, and a synthetic speech wave is reconstructed from the amplitude-adjusted excitation signal.

Another autocorrelation function representation of speech is obtained in the present invention by correlating the incoming speech wave with both itself and its Hilbert transform to obtain a number of autocorrelation samples composed in equal parts of samples of periods of the autocorrelation function and samples of periods of the ninety-degree phase-shifted autocorrelation function. These samples of the two autocorrelation functions also constitute a group of narrow-band control signals.

In order to reconstruct an accurate replica of the speech autocorrelation function at the synthesizer, the sampling theorem requires that the samples of each period of the autocorrelation function be spaced at the Nyquist interval. Since the number of samples per period depends upon the length of the fundamental period, which is the reciprocal of the variable fundamental pitch

frequency of the original speech wave, variations in the fundamental pitch frequency must be accompanied by changes in the number of autocorrelation samples if distortion-free speech is to be reproduced. Accordingly, it is a specific object of this invention to reproduce synthetic speech with accuracy by varying the number of autocorrelation samples in synchrony with the instantaneous fundamental pitch frequency.

The number of autocorrelation samples is varied in synchrony with the instantaneous fundamental pitch frequency by first obtaining a fixed number of samples corresponding to the longest pitch period to be accurately reproduced. The fixed number of samples is then passed through a bank of gates that are set to close at various predetermined thresholds in response to a variable magnitude pitch control signal. As the talker's pitch changes, the pitch control signal changes the number of gates in the open condition, thereby varying the number of autocorrelation samples passed by the gates. The samples passed by the gates constitute narrow-band control signals from which distortion-free synthetic speech may be reproduced.

Transmission channel bandwidth is further conserved in this invention by taking advantage of the symmetry of autocorrelation function. This is achieved by obtaining samples of half periods of the symmetrical autocorrelation function at the analyzer, and by deriving from each control signal transmitted to the synthesizer two equal amplitude, symmetrically located points on each period of the reconstructed speech wave. Similarly, half periods of the antisymmetrical ninety-degree phase-shifted autocorrelation function are sampled at the analyzer, and at the synthesizer, two equal amplitude, antisymmetrically located points on each period of the reconstructed wave are derived from each half-period sample. By thus reducing the number of control signals by a factor of two, the bandwidth necessary to transmit the control signals is also reduced by a factor of two.

An inherent source of distortion in the representation of most speech sounds by autocorrelation functions lies in the quadratic character of the amplitude spectra of autocorrelation functions generally, as compared with the amplitude spectra of the waves from which they are derived. It is a specific object of this invention to eliminate the quadratic distortion inherent in the autocorrelation function representation of speech by performing a square-root-taking operation upon the amplitude spectrum of the speech autocorrelation function.

In this invention, the autocorrelation control signals are converted into signals representative of discrete values of the autocorrelation amplitude spectrum by a network composed of an array of multiplier elements supplied with weighting signals. The control signals and the weighting signals are combined by the array of multipliers in accordance with a Fourier transformation to produce autocorrelation amplitude spectrum signals. The spectrum signals are then applied to a group of square-root-taking circuits which produce rooted spectrum signals representing an amplitude spectrum of the same shape as the amplitude spectrum of the original speech wave. Artificial speech free of spectrum squaring distortion is reproduced from the autocorrelation function counterparts of the rooted spectrum signals, the counterparts being generated by a second array of multipliers supplied with weighting signals in accordance with an inverse Fourier transformation.

An important feature of this invention is the variation of the magnitudes of the output signals of the multiplier arrays in synchrony with changes in the fundamental pitch frequency of the original speech wave. This synchronism is required for accurate reproduction of speech, since both the autocorrelation function and its amplitude spectrum are functions of the variable fundamental pitch frequency. Synchronization is achieved by varying the magnitudes of the weighting signals in synchrony with the instantaneous fundamental pitch frequency.

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

FIGS. 1A, 1B, 1C, 1D, 1E, and 1F are waveform diagrams of assistance in explaining the operation of the apparatus of this invention;

FIG. 1G is a schematic block diagram showing a complete speech transmission system based upon the principles of this invention;

FIG. 2A is a schematic block diagram showing apparatus for representing speech in terms of autocorrelation control signals;

FIG. 2B is a graph showing the variation of the threshold voltages of the gates of FIG. 2A with increasing delay times of the input signals to the gates;

FIG. 2C is a graph showing the variation of pitch control signal voltage from circuit PC2 of FIG. 2A as a function of the fundamental pitch frequency of the incoming speech wave;

FIG. 3 is a schematic block diagram showing apparatus for reconstructing artificial speech from autocorrelation control signals;

FIG. 4 is a schematic block diagram showing apparatus alternative to that of FIG. 2A;

FIG. 5 is a schematic block diagram showing apparatus for reconstructing artificial speech from autocorrelation control signals produced by the apparatus of FIG. 4;

FIG. 6A is a schematic block diagram showing apparatus for unsquaring the amplitude spectrum of the control signals produced by the apparatus of FIG. 2A;

FIG. 6B is a schematic block diagram showing apparatus for generating autocorrelation control signals from the rooted spectrum signals produced by the apparatus of FIG. 6A;

FIG. 7A is a schematic block diagram showing apparatus for unsquaring the amplitude spectrum of the control signals produced by the apparatus of FIG. 4; and

FIGS. 7B and 7C are schematic block diagrams showing apparatus for generating autocorrelation control signals from the rooted spectrum signals produced by the apparatus of FIG. 7A.

Mathematical Foundations

A periodic speech wave $g(t)$ with period T, fundamental frequency

$$f_0=\frac{1}{T}$$

and bandwidth W may be expanded in a Fourier series

$$g(t)=\sum_n G(nf_0)\cos(2\pi nf_0t+\psi_n) \qquad (1)$$

where the coefficients $G(nf_0)$ constitute the amplitude spectrum of $g(t)$, and ψ_n are the phase angles.

The autocorrelation function of $g(t)$ is defined as

$$\varphi(\tau)\equiv\frac{1}{T}\int_0^T g(t)\cdot g(t-\tau)dt \qquad (2)$$

where $\varphi(\tau)$ has the same period as $g(t)$, and the variable τ represents the amount of time by which the speech wave is delayed before being multiplied together with the undelayed speech wave.

Substituting (1) into (2) and performing the indicated integration produces

$$\varphi(\tau)=\sum_n \Phi(nf_0)\cos(2\pi nf_0\tau) \qquad (3)$$

where

$$\Phi(nf_0)=\Phi(-nf_0) \qquad (3a)$$

and

$$\Phi(nf_0)=G^2(nf_0) \qquad (3b)$$

that is, $\Phi(nf_0)$ is an even function, the amplitude spectrum of $\varphi(\tau)$ is the square of the amplitude spectrum of $g(t)$, and all of the phase angles of $\varphi(\tau)$ are zero.

The Hilbert transform of $g(t)$ is defined as

$$\hat{g}(t)\equiv\sum_n G(nf_0)\sin(2\pi nf_0t+\psi_n) \qquad (4)$$

and by correlating $\hat{g}(t)$ and $g(t)$ in accordance with Equation 2, there is produced the ninety-degree phase-shifted speech autocorrelation function

$$\hat{\varphi}(\tau)=\frac{1}{T}\int_0^T \hat{g}(t)\cdot g(t-\tau)dt \qquad (5)$$

where $\hat{\varphi}(\tau)$ has the same period as $g(t)$. Substituting (1) and (4) in (5) and performing the indicated integration, Equation 5 becomes

$$\hat{\varphi}(\tau)=\sum_n \hat{\Phi}(nf_0)\sin(2\pi nf_0\tau) \qquad (6)$$

where

$$\hat{\Phi}(nf_0=\Phi(nf_0),\text{ for } nf_0>0 \qquad (7)$$

and

$$\hat{\Phi}(nf_0)=-\Phi(nf_0),\text{ for } nf_0<0 \qquad (7a)$$

that is, $\hat{\Phi}(nf_0)$ is an odd function with the same absolute amplitude as $\Phi(nf_0)$.

FIG. 1A shows several periods of a typical speech wave, and FIGS. 1B and 1C show several periods of its autocorrelation functions $\varphi(\tau)$ and $\hat{\varphi}(\tau)$, respectively.

FIG. 1D shows the amplitude spectrum of a typical voiced sound, and FIG. 1E shows the amplitude spectrum of the corresponding speech autocorrelation function. A comparison of FIG. 1D with FIG. 1E reveals that squaring the amplitude spectrum as given in Equation 3*b* produces a doubling, in decibels, of the amplitude differences between peaks of the spectral envelope, thereby suppressing the relatively small peaks.

FIG. 1F illustrates the amplitude spectrum $\hat{\Phi}(nf_0)$ of $\hat{\varphi}(\tau)$ corresponding to the voiced sound of FIG. 1D. It is observed that $\hat{\Phi}(nf_0)$ is antisymmetrical about $f=0$, whereas both $G(nf_0)$ and $\Phi(nf_0)$ are symmetrical about $f=0$.

Since $\Phi(nf_0)$ is symmetrical over the interval $[-W,W]$ centered about $f=0$, it may be expanded in a Fourier cosine series over this interval

$$\Phi(nf_0)=\frac{1}{2TW}\sum_{j=-TW}^{TW}\varphi\left(\frac{j}{2W}\right)\cos\left[2\pi nf_0\left(\frac{j}{2W}\right)\right] \qquad (8)$$

where T is the period and W is the bandwidth of both $g(t)$ and $\varphi(\tau)$. Since the half-period samples of both $\varphi(\tau)$ and the cosine function from

$$-\frac{T}{2}\text{ to } 0$$

are mirror images of the half-period samples from

$$0\text{ to }\frac{T}{2}$$

$\Phi(nf_0)$ is uniquely determined by either set of half-period samples; hence Equation 8 may be rewritten

$$\Phi(nf_0)=\frac{1}{TW}\left[\frac{1}{2}\varphi(0)+\sum_{j=1}^{N=TW}\varphi\left(\frac{j}{2W}\right)\cos 2\pi nf_0\left(\frac{j}{2W}\right)\right] \qquad (9)$$

The inverse of Equation 9, corresponding to Equation 3, is

$$\varphi\left(\frac{j}{2W}\right)=\sum_{n=1}^{p=\frac{W}{f_0}}\Phi(nf_0)\cos 2\pi nf_0\left(\frac{j}{2W}\right) \tag{9a}$$

The symmetry of the values of $\Phi(nf_0)$ itself over the interval $[-W,W]$ implies that $\Phi(nf_0)$ is uniquely determined by its values over either of the half intervals, $[-W,0]$ or $[0,W]$. Considering the nonsymmetrical, one-sided spectrum $\Phi_W(nf_0)\equiv\Phi(nf_0)$, $0\leq nf_0\leq W$, over the half interval $[0,W]$, it may be decomposed into an even function $\Phi_c(nf_0)$ and an odd function $\Phi_s(nf_0)$,

$$\Phi_W(nf_0)=\Phi_c(nf_0)+\Phi_s(nf_0) \tag{10}$$

where

$$\Phi_c(nf_0)=\tfrac{1}{2}[\Phi(nf_0)+\Phi(nf_0-W)] \tag{11}$$

and

$$\Phi_s(nf_0)=\tfrac{1}{2}[\Phi(nf_0)-\Phi(nf_0-W)] \tag{12}$$

From Equations 7 and 7*a*, Equation 12 may be written

$$\Phi_s(nf_0)=\tfrac{1}{2}[\hat{\Phi}(nf_0)+\hat{\Phi}(nf_0-W)] \tag{13}$$

Expanding $\Phi_W(nf_0)$ in a Fourier series, the even part $\Phi_c(nf_0)$ gives rise to cosine terms whose coefficients are

$$\int_0^W \frac{1}{2}[\Phi(nf_0)+\Phi(nf_0-W)]\cos 2\pi nf_0\left(\frac{j}{W}\right)d(nf_0)=\varphi\left(\frac{j}{W}\right) \tag{14}$$

and the odd part $\Phi_s(nf_0)$ gives rise to sine terms whose coefficients are

$$\int_0^W \frac{1}{2}[\hat{\Phi}(nf_0)+\hat{\Phi}(nf_0-W)]\sin 2\pi nf_0\left(\frac{j}{W}\right)d(nf_0)=\hat{\varphi}\left(\frac{j}{W}\right) \tag{15}$$

Therefore, the Fourier series expansion of $\hat{\Phi}_W(nf_0)$ is

$$\Phi_W(nf_0)=\frac{1}{TW}\sum_{j=-\frac{TW}{2}}^{\frac{TW}{2}}\varphi\left(\frac{j}{W}\right)\cos 2\pi nf_0\left(\frac{j}{W}\right)+\hat{\varphi}\left(\frac{j}{W}\right)\sin 2\pi nf_0\left(\frac{j}{W}\right) \tag{16}$$

or, from symmetry considerations,

$$\Phi_W(nf_0)=\frac{2}{TW}\left[\frac{1}{2}\varphi(0)+\sum_{j=1}^{N=\frac{TW}{2}}\varphi\left(\frac{j}{W}\right)\cos 2\pi nf_0\left(\frac{j}{W}\right)+\hat{\varphi}\left(\frac{j}{W}\right)\sin 2\pi nf_0\left(\frac{j}{W}\right)\right] \tag{17}$$

The speech correlation function

$$\varphi_W\left(\frac{j}{W}\right)$$

corresponding to $\Phi_W(nf_0)$ may also be expanded in a Fourier series

$$\varphi_W\left(\frac{j}{W}\right)=\sum_{n=0}^{p=\frac{W}{f_0}}\Phi_c(nf_0)\cos 2\pi nf_0\left(\frac{j}{W}\right)+\Phi_S(nf_0)\sin 2\pi nf_0\left(\frac{j}{W}\right) \tag{18}$$

By definition, the coefficients of the sine and cosine terms in Equation 18 may be written in terms of the one-sided spectrum,

$$\Phi_c(nf_0)=\tfrac{1}{2}[\Phi_W(nf_0)+\Phi_W(W-nf_0)] \tag{19}$$

$$\Phi_s(nf_0)=\tfrac{1}{2}[\Phi_W(nf_0)-\Phi_W(W-nf_0)] \tag{19a}$$

substituting (19) and (19*a*) into (18), the correlation function

$$\varphi_W\left(\frac{j}{W}\right)$$

may be expressed in terms of the one-sided spectrum,

$$\varphi_W\left(\frac{j}{W}\right)=\frac{1}{2}\sum_{n=0}^{p=\frac{W}{f_0}}[\Phi_W(nf_0)+\Phi_W(W-nf_0)]\cos 2\pi nf_0\left(\frac{j}{W}\right)+[\Phi_W(nf_0)-\Phi_W(W-nf_0)]\sin 2\pi nf_0\left(\frac{j}{W}\right) \tag{20}$$

Complete Speech Transmission System

Referring first to FIG. 1G, an incoming speech wave from source **80** is applied in parallel to pitch circuit **81** and to speech autocorrelation function analyzer **82**. Analyzer **82** derives samples of the speech autocorrelation function from the incoming speech wave, and, under the control of a pitch signal from circuit **81**, varies the number of samples in synchrony with the instantaneous fundamental pitch frequency of the speech wave. The details of source **80**, circuit **81**, and analyzer **82** are described below, as are the details of the other elements of FIG. 1G.

The amplitude spectrum of the autocorrelation samples produced by analyzer **82** is unsquared by apparatus composed of weighting signal sources **83**, Fourier transformation networks **84** and **86**, and square-root-taking circuits **85**. A speech wave is reconstructed from the unsquared autocorrelation samples by speech synthesizer **88**, utilizing an excitation signal from generator **87**, and artificial speech is reproduced from the reconstructed speech wave by reproducer **89**.

Pitch circuit **81** and analyzer **82** are located at a transmitter terminal, while generator **87** and synthesizer **88** are located at a receiver terminal. Elements **83**, **84**, **85**, and **86**, however, may be located at either the transmitter terminal or the receiver terminal.

Analyzer

Referring now to FIG. 2A, an incoming speech wave $g(t)$ from source **20**, for example, a microphone, is passed through low-pass filter **210**, which may be of any well-known variety. Filter **210** serves to band-limit the speech wave by removing all frequency components of $g(t)$ greater than W cycles per second, where, for example, the passband of filter **210** may be chosen so as to remove all frequency components greater than $W=3{,}000$ cycles per second.

The band-limited speech wave output of filter **210** is applied simultaneously to a tapped delay line, for example, a tapped acoustic delay line **231**, and to one of the input terminals of a bank of multipliers, for example, a bank of well-known modulators M21, M22, . . ., M2N, each provided with two input terminals and one output terminal.

Delay line **231**, which is terminated in a matched impedance **211** to prevent reflection, is provided with N taps P21, P22, . . ., P2N, at which appear signals proportional to samples of the variously delayed speech wave, $g(t-\tau_1)$, $g(t-\tau_2)$, . . ., $g(t-\tau_N)$, where

$$\tau_1=\frac{1}{2W}\text{ seconds, }\tau_2=\frac{2}{2W}\text{ seconds, }\ldots,\ \tau_N=\frac{N}{2W}\text{ seconds}$$

are the various delay times corresponding to taps **P21**, **P22**, . . ., **P2N**, respectively.

Modulators M21, M22, . . ., M2N, in addition to receiving the undelayed speech wave at one of their input terminals, have their second input terminals connected to delay line taps P21, P22, . . ., P2N, respectively, to receive the variously delayed speech wave samples as second input signals.

The modulators develop at their output terminals signals proportional to the products $g(t).g(t-\tau_1)$, $g(t)\cdot g(t-\tau_2)$, . . ., $g(t)\cdot g(t-\tau_N)$, which are then passed to a bank of averaging devices, for example, conventional low-pass filters F21, F22, . . ., F2N, each having a cut-off of 25 cycles per second. The magnitudes of the output signals of filters F21, F22, . . ., F2N are proportional to time averages of the product signals received from the bank of modulators; hence from Equation 2, it is seen that the magnitudes of these signals are proportional to samples of the autocorrelation function at various delay times,

$$\varphi\left(\frac{1}{2W}\right), \varphi\left(\frac{2}{2W}\right), \ldots, \varphi\left(\frac{N}{2W}\right)$$

In addition, the incoming speech wave is correlated with itself to form a signal proportional to $\varphi(0)$. This is achieved by applying the undelayed speech wave to both input terminals of modulator M20, and by passing the product output signal of M20 through low-pass filter F20.

As pointed out in FIG. 1B, the samples of $\varphi(\tau)$ over each half period from

$$\frac{1}{2W} \text{ to } \frac{T}{2}$$

seconds delay are mirror images of the samples of $\varphi(\tau)$ over the half period from

$$-\frac{1}{2W} \text{ to} -\frac{T}{2}$$

seconds delay; hence $\varphi(\tau)$ is uniquely represented by TW samples, in addition to $\varphi(0)$, starting at

$$\frac{1}{2W}$$

seconds delay and ending at

$$\frac{T}{2}$$

seconds delay, with a delay interval of

$$\frac{1}{2W}$$

seconds between successive samples. By reducing the number of control signals by a factor of two, the bandwidth necessary to transmit the control signals is also reduced by a factor of two. For example, if $\varphi(\tau)$ has a period of $T=10$ milliseconds, then for a bandlimit of $W=3{,}000$ cycles per second, $\varphi(\tau)$ is uniquely determined by providing delay line **231** with $TW=30$ taps, where the delay time at tap **P21** is

$$\frac{1}{2W}=\frac{1}{6}$$

millisecond, the delay time at tap **P22** is

$$\frac{2}{2W}=\frac{1}{3}$$

millisecond, the delay time at tap **P2N** is

$$\frac{T}{2}=5$$

milliseconds, and the delay interval between adjacent taps is ⅙ millisecond.

Obtaining the correct number of samples of the speech autocorrelation function is complicated by variations in the period T of the speech autocorrelation function in synchrony with variations in the fundamental pitch frequency,

$$f_0=\frac{1}{T}$$

of the original speech wave. In order to obtain the proper number of samples of each half period of the speech autocorrelation function despite these variations, delay line **231** of FIG. 2A is provided with that number of taps needed to obtain at the output terminals of filters F20, F21, F22, . . ., F2N a sufficient number of half-period autocorrelation samples of the longest period that it is desired to represent fully. Filters **F20, F21, F22,** . . ., F2N are then connected to conventional gates **G20, G21,** G22, . . ., G2N, respectively, and when shorter periods occur, excess samples occurring at points past the

$$\frac{T}{2}$$

seconds or half-period sample point are blocked by closing the appropriate gates.

Gates G20, G21, G22, . . ., G2N are set to close at various predetermined thresholds in response to a pitch control signal from circuit **PC2** whose magnitude varies in synchrony with the fundamental pitch frequency, as illustrated in FIG. 2C. Correspondingly, as shown in FIG. 2B, the thresholds of the gates are set at progressively lower voltages as the delay times of the autocorrelation samples applied to them increase. The solid line in FIG. 2C indicates that the voltage of the pitch control signal applied to the gates increases as the pitch periods become shorter than the maximum period, and the pitch control signal closes those gates that are supplied with autocorrelation samples occurring at delay times past the half-period sample point.

For example, if the longest pitch period that it is desired to represent fully is 10 milliseconds (corresponding to a deep bass voice of fundamental pitch frequency 100 cycles per second), and if frequency components greater than $W=3{,}000$ cycles per second are removed by filter **210** of FIG. 2A, then in addition to $\varphi(0)$, 30 half-period autocorrelation samples are required to represent the period fully. By providing delay line **231** with 30 taps, and by providing 30 modulators and 30 filters (in addition to modulator **M20** and filter **F20** for $\varphi(0)$), the longest autocorrelation function periods are adequately sampled. To the output terminal of each filter there is connected a gate, and when shorter pitch periods occur, for example, a pitch period of 5 milliseconds (corresponding to a fundamental pitch frequency of 200 cycles per second), then only 15 half-period samples are required, besides $\varphi(0)$, and the remaining 15 samples, which occur at delay times past the half-period sample point,

$$\frac{T}{2}=2\frac{1}{2}$$

milliseconds, are eliminated by closing the 15 gates receiving the excess samples from their respective filters.

In order to derive the pitch control signal for gates **G20, G21, G22,** . . ., **G2N**, the band-limited speech wave output of filter **210** is applied simultaneously to pitch detector **212** and voiced-unvoiced detector **213** of circuit **PC2**. Pitch detector **212**, which may be of any well-known construction, develops at its output terminal a signal whose magnitude is directly proportional to the instantaneous fundamental pitch frequency, f_0, of voiced sounds, as illustrated by the solid line in FIG. 2C. The output terminal of pitch detector **212** is connected to the control terminals of gates **G20, G21, G22,** . . ., **G2N** by relay **215** when voiced-unvoiced detector **213**, of well-known design, detects the presence of voiced sounds. When unvoiced sounds are present, however, voiced-unvoiced detector **213** causes relay **215** to connect the output terminal of energy source **216** to the control terminals of gates G20, G21, G22, . . ., G2N. Energy source **216** produces a signal of constant magnitude E, as shown by the dashed line in FIG. 2C, which permits sufficient autocorrelation function samples to be passed by gates **G20, G21,** G22, . . ., G2N to specify unvoiced sounds.

The signals from filters F20, F21, F22, . . . F2N, passed by gates G20, G21, G22, . . . G2N, constitute a set of control signals that represent half periods of the speech autocorrelation function. It has been observed that the speech autocorrelation function changes very little from period to period, hence the variation of the control signals is very small. As a result, each control signals occupies a relatively narrow-frequency band, on the order of 25 cycles per second, and the entire group of control signals may be transmitted, if desired, over a much narrower frequency band than is required for transmission of the original speech wave. It is to be understood, however, that the control signals may be transmitted by means of any well-known transmission method, for example, pulse-code modulation, to meet the requirements of a particular application of this invention.

The autocorrelation control signals transmitted to the synthesizer terminal must be supplemented with a signal that indicates whether the instantaneous speech sound is voiced or unvoiced, and if voiced, its fundamental pitch frequency. At the synthesizer, there is derived from this supplementary signal an excitation signal which is closely correlated with the vocal characteristics of the original speech wave, and the artificial speech reconstructed from the excitation signal thus preserves the vocal characteristics of the original speech. As shown in circuit PC2 of FIG. 2A, the output terminal of voiced-unvoiced detector **213** is connected to relay **214** so as to operate the relay when sounds detected by microphone **20** are voiced, and to disconnect the relay when the sounds are unvoiced. When the contacts of relay **214** are closed in response to the presence of voiced sounds, the output signal from pitch detector **212** serves as a supplementary pitch control signal to be transmitted together with the autocorrelation control signals.

Synthesizer

Speech is reproduced from the transmitted control signals at a synthesizer, a preferred embodiment of which is shown in FIG. 3. The synthesizer reconstructs full periods of a synthetic speech wave from the half-period control signals by using each control signal to form a pair of symmetrically disposed samples of the reconstructed periods. The artificial periods thereby reconstructed duplicate the symmetry of the original autocorrelation periods from which they are derived.

Referring now to the apparatus of FIG. 3, the incoming control signals

$$\varphi(0), \varphi\left(\frac{1}{2W}\right), \ldots, \varphi\left(\frac{N}{2W}\right)$$

are applied to the control terminals of a bank of modulators L30, L31, . . . L3N, respectively. The output terminal of modulator L30 is connected to the center tap R30 of delay line 341, which is terminated in a matched impedance 321 to prevent reflection. The output terminal of each of the other modulators L31, . . . L3N is connected to delay line 341 via two taps which are disposed at equal intervals about center tap R30. Thus, for example, the output terminal of modulator L31 is connected to taps R31, r31 spaced at equal

$$\frac{1}{2W}$$

second intervals about tap R30, respectively, and the output terminal of modulator L3N is connected to taps R3N, r3N spaced at equal

$$\frac{T}{2}$$

second intervals about tap R30, respectively.

The incoming control signals applied to the control terminals of the modulators adjust the amplitude of a vocal excitation signal supplied to the modulators from vocal excitation signal generator **391**. The excitation signal is derived by generator **391** from the incoming pitch control signal, which is first passed through an equalizing delay **31** to synchronize the pitch control signal

with the control signals. Excitation generator **391**, which may comprise conventional buzz and hiss sources, operates to provide an excitation signal to the modulators that is closely correlated with the voiced-unvoiced and fundamental pitch frequency characteristics of the original speech wave. Artificial speech reconstructed under the influence of the control signals from the excitation signal supplied by generator **391** thus preserves the naturalness as well as the intelligibility of the original speech wave.

The amplitude-adjusted excitation signals formed by the modulators at taps R31, . . . R3N are symmetrically located with respect to the signals formed at taps r31, . . ., r3N, thereby duplicating the symmetry of $\varphi(\tau)$ about the center of each period. The variously delayed signals appearing at the lower terminal of delay line **341** are smoothed to form an artificial wave by low-pass filter **342**, proportioned to eliminate all frequencies greater than 3,000 cycles per second. The output terminal of filter **342** is connected to a conventional reproducer **343**, which converts the artificially reconstructed electrical speech wave into audible and intelligible speech sounds.

Alternate Analyzer

Alternative apparatus for representing speech in terms of the one-sided amplitude spectrum, $\Phi_W(nf_0)$, is shown in FIG. 4. As given by Equation 17, $\Phi_W(nf_0)$ may be expressed in terms of half-period samples of both $\varphi(\tau)$ and $\varphi(\tau)$, and the apparatus of FIG. 4 is based upon this relationship.

Referring now to FIG. 4, an incoming speech wave $g(t)$ from microphone **40** is band-limited by low-pass filter **410** to remove all frequencies greater than W cycles per second, for example, 3,000 cycles per second. The band-limited speech wave is applied simultaneously to tapped delay line **431**, of a suitable variety, and to one input point of modulators M41, M42, . . ., M4N.

Delay line **431** is terminated in a matched impedance **411** to prevent reflection, and is provided with taps P41, P42, . . ., P4N, at which appear signals proportional to values of the variously delayed speech wave, $g(t-\tau_1)$, $g(t-\tau_2)$, . . ., $g(t-\tau_N)$, where

$$\tau_1=\frac{1}{W}\text{ seconds, }\tau_2=\frac{2}{W}\text{ seconds, }\ldots, \tau_N=\frac{N}{W}\text{ seconds}$$

are the various delay times corresponding to taps P41, P42, . . ., P4N, respectively.

The signals appearing at the taps of delay line **431** are passed to the second input points of modulators M41, M42, . . ., M4N, and the modulators develop at their output terminals signals whose magnitudes are proportional to the products $g(t)\cdot g(t-\tau_1)$, $g(t)\cdot g(t-\tau_2)$, . . ., $g(t)\cdot g(t-\tau_N)$, respectively. The output terminals of these modulators are connected to averaging devices, for example, low-pass filters F41, F42, . . ., F4N, respectively, which are constructed to pass frequencies equal to or less than 25 cycles per second, thereby forming output signals proportional to samples of the autocorrelation function at various delay times,

$$\varphi\left(\frac{1}{W}\right), \varphi\left(\frac{2}{W}\right), \ldots, \varphi\left(\frac{N}{W}\right)$$

In addition, the incoming speech wave is correlated with itself to form a signal proportional to the $\varphi(0)$ term in Equation 17. This is accomplished by applying the undelayed speech wave to both input terminals of modulator M40, and by passing the product output signal $g(t)\cdot g(t)$ of M40 through low-pass filter F40.

The band-limited speech wave output of filter **410** is also applied to ninety-degree phase-shifter **430** which develops a signal proportional to the Hilbert transform $\hat{g}(t)$ of the speech wave. The use of ninety-degree phase-shifter **430** to obtain the Hilbert transform of the speech wave is based upon the well-known quadrature relationship between a function and its Hilbert transform, a proof of which is found in S. Goldman, Information Theory, page 332 (1953). Since functions that are in quadrature with each other differ solely in phase by ninety degrees, element **430** may be one of several equally suitable ninety-degree phase-shifting devices for obtaining the Hilbert transform $\hat{g}(t)$ of the input signal $g(t)$.

The Hilbert transform output of phase-shifter **430** is applied in parallel to one input point of modulators $\hat{M}$**41**, $\hat{M}$**42**, . . . , $\hat{M}$**4N**, and together with the delayed signals from taps **P41**, **P42**, . . . , **P4N** applied to the second input points of these modulators, there is formed at their output terminals signals whose magnitudes are proportional to the products $\hat{g}(t)\cdot g(t-\tau_1), \hat{g}(t)\cdot g(t-\tau_2), \ldots, \hat{g}(t)\cdot g(t-\tau_N)$. The output terminals of modulators $\hat{M}$**41**, $\hat{M}$**42**, . . . , $\hat{M}$**4N** are connected to low-pass filters $\hat{F}$**41**, $\hat{F}$**42**, . . . , $\hat{F}$**4N**, respectively, of a construction similar to that of filters **F41**, **F42**, . . . , **F4N**, thereby developing output signals whose magnitudes are proportional to values of the ninety-degree phase-shifted speech autocorrelation function,

$$\hat{\varphi}\left(\frac{1}{W}\right), \hat{\varphi}\left(\frac{2}{W}\right), \ldots, \hat{\varphi}\left(\frac{N}{W}\right)$$

in accordance with Equation 5.

From Equation 17, the one-sided amplitude spectrum is determined by

$$\left(\frac{TW}{2}+1\right)$$

half-period samples of $\varphi(\tau)$ and

$$\frac{TW}{2}$$

half-period samples of $\hat{\varphi}(\tau)$. It is understood, however, that the half-period samples of $\hat{\varphi}(\tau)$ from

$$\frac{1}{W} \text{ to } \frac{T}{2}$$

seconds delay are inverted images of the half-period samples from

$$-\frac{1}{W} \text{ to} -\frac{T}{2}$$

seconds delay.

As in the case of the representation of speech by samples of $\varphi(\tau)$ alone, variations in the period T of both $\varphi(\tau)$ and $\hat{\varphi}(\tau)$ make it necessary to vary the number of samples of both $\varphi(\tau)$ and $\hat{\varphi}(\tau)$ in synchrony with changes in the fundamental pitch frequency,

$$f_0=\frac{1}{T}$$

of the incoming speech wave. This is achieved by providing delay line **431** with a fixed number of taps sufficient to sample the longest pitch period to be fully represented. For pitch periods shorter than the maximum to be represented fully, filters **F40**, **F41**, **F42**, . . . , **F4N**, and $\hat{F}$**41**, $\hat{F}$**42**, . . . , $\hat{F}$**4N** are connected to gates **G40**, **G41**, **G42**, . . . , **G4N**, and $\hat{G}$**41**, $\hat{G}$**42**, . . . , $\hat{G}$**4N**, respectively, which block excess samples occurring at points past the half-period sample point of both $\varphi(\tau)$ and $\hat{\varphi}(\tau)$. The gates shown in FIG. 4 operate in the same fashion as the gates shown in FIG. 2A, being set to close at various predetermined thresholds by a variable magnitude pitch control signal from circuit PC4. Circuit PC4, which is identical in both structure and operation to circuit PC2 of FIG. 2A, produces from the band-limited output signal of filter **410** a pitch control signal whose magnitude is proportional to the instantaneous fundamental pitch frequency.

The autocorrelation samples passed by the gates constitute a set of control signals representing half periods of each of the speech autocorrelation functions, $\varphi(\tau)$ and $\hat{\varphi}(\tau)$. Each of these signals occupies a relatively narrow-frequency range, on the order of **25** cycles per second, and the entire group of signals may be transmitted over a substantially narrower band of frequencies than that required for transmission of the original speech wave.

As in the case of speech representation by samples of $\varphi(\tau)$ alone, the control signals derived by the analyzer apparatus of FIG. 4 must be supplemented with a signal indicative of the voiced-unvoiced and fundamental pitch frequency characteristics of the instantaneous speech sound. This supplementary or pitch control signal is derived from the incoming speech wave from source **40** by circuit PC4. The pitch control signal and the autocorrelation control signals constitute the compressed frequency representation of speech from which artificial speech is reconstructed.

The autocorrelation control signals and the supplementary signal produced by the apparatus of FIG. 4 may be transmitted over a narrow-band channel by any suitable method, for example, by pulse-code modulation. After transmission, speech is reproduced from the signals by a synthesizer, a preferred embodiment of which is shown in FIG. 5.

Alternate Synthesizer

The synthesizer shown in FIG. 5 simultaneously reconstructs samples of symmetrical periods from the half-period $\varphi(\tau)$ control signals and samples of antisymmetrical periods from the half-period $\hat{\varphi}(\tau)$ control signals. The two sets of samples are added and filtered to form a speech wave whose one-sided amplitude spectrum $\Phi_W(nf_0)$ contains the same information as the two-sided amplitude spectrum $\Phi(nf_0)$ of $\varphi(\tau)$.

Referring now to FIG. 5, control signal $\varphi(0)$ is applied to the control terminal of modulator **L50**, the output terminal of which is connected to the center tap **R50** of delay line **541**, terminated in a matched impedance **521** to prevent reflection. Each of the other control signals is applied simultaneously to the control terminals of two modulators whose output terminals are connected to delay line **541** through two taps disposed at equal intervals above and below center tap **R50**. Thus control signal

$$\varphi\left(\frac{1}{W}\right)$$

for example, is applied to the control terminals of modulators **L51**, ***l*51**, the output terminals of which are connected to taps **R51**, ***r*51**, respectively, located at equal

$$\frac{1}{W}$$

second intervals about center tap **R50**.

In addition, the output terminals of each pair of modulators receiving a sample of $\varphi(\tau)$ and of $\hat{\varphi}(\tau)$ at the same delay time are connected to the same tap of delay line **541**, in accordance with Equation 16. Hence, the output terminals of modulators **L51**, $\hat{L}$**51** are both connected to tap **R51**, and the output terminals of modulators ***l*51**, $\hat{l}$**51** are both connected to tap ***r*51**.

The control signals adjust the amplitude of an excitation signal supplied to the input terminals of the modulators from vocal excitation signal generator **591**. Generator **591**, which is preceded by equalizing delay **51**, is identical in construction and operation to generator **391** of FIG. 3, deriving from the incoming pitch control signal an excitation signal closely correlated with the voiced-unvoiced and fundamental pitch frequency characteristics of the original speech wave.

The excitation signal is applied directly to the input terminals of all modulators supplied with $\varphi(\tau)$ control signals, thereby developing at their respective delay line taps samples of a symmetrical wave duplicating the original symmetry of $\varphi(\tau)$. For modulators supplied with $\hat{\varphi}(\tau)$ control signals, however, the excitation signal is applied directly to the input terminals of half of these modulators, for example, modulators $\hat{L}51, \ldots, \hat{L}5N$, connected to taps above center tap R50, but is reversed in polarity by polarity inverter 571 before being applied to the input terminals of modulators $\hat{l}51, \ldots, \hat{l}5N$, connected to taps below center tap R50. Thus the modulators supplied with $\hat{\varphi}(\tau)$ control signals develop at their respective taps samples of an antisymmetrical wave duplicating the original antisymmetry $\hat{\varphi}(\tau)$.

The reconstructed signal formed at the output terminal of delay line 541 is a linear combination of the amplitude-adjusted excitation signals formed at the various taps, and the information content of the spectrum of this reconstructed signal equals the information content of the spectrum of the original speech autocorrelation function. The reconstructed output signals of delay line 541 are smoothed by low-pass filter 542, having a passband of 0 to 3,000 cycles per second, to form a synthetic speech wave. Artificial speech is reconstructed from the output wave of filter 542 by reproducer 543, which may be of any well-known construction.

Unsquaring Apparatus

The amplitude spectrum of artificial speech synthesized directly from the autocorrelation control signals produced by the analyzer apparatus shown in either FIG. 2A or FIG. 4 is the square of the amplitude spectrum of the original speech wave, in accordance with Equations 3*b* and 7. This squaring of the amplitude spectrum impairs the intelligibility of most artificial speech sounds reconstructed from the autocorrelation control signals, due to changes in relative spectral values. In order for synthetic speech to be free of the spectrum squaring defects inherent in the autocorrelation function representation of speech, the present invention subjects the control signals produced by the above analyzers to a square-root-taking operation prior to speech synthesis.

Apparatus for removing the square law dependence of the autocorrelation amplitude spectrum is shown in FIGS. 6A, 6B, 7A, 7B, and 7C. The apparatus of FIGS. 6A and 6B is designed to operate upon the control signals produced by the analyzer shown in FIG. 2A, and the apparatus of FIGS. 7A, 7B, and 7C is designed to operate upon the control signals produced by the analyzer shown in FIG. 4. The apparatus shown in FIGS. 6A, 6B, 7A, 7B, and 7C may be located at either the analyzer station or the synthesizer station, or, if desired, the component parts may be conveniently divided between the two stations.

As shown in FIG. 6A, the incoming autocorrelation control signals derived from an analyzer of the type illustrated in FIG. 2A are applied to the input terminals of network 60. The function of network 60 is to transform the control signals into signals representative of the autocorrelation amplitude spectrum, in accordance with the Fourier transformation given by Equation 9. The amplitude spectrum signals developed at the output terminals of network 60 are then passed through a bank of rooters H61, H62, . . ., $H6_p$, which perform a square-root-taking operation, thereby producing a new set of signals representing an amplitude spectrum of the same shape as the amplitude spectrum of the original speech wave. The rooted spectrum signals are then passed through network 61 of FIG. 6B, which operates in accordance with the inverse Fourier transformation given by Equation 9*a* to produce a new set of autocorrelation control signals from which artificial speech free of spectrum squaring distortion may be reconstructed by a synthesizer of the type shown in FIG. 3.

Recalling Equation 9, the autocorrelation amplitude spectrum is derived from samples of the autocorrelation function in the following fashion, ignoring the constant factor

$$\frac{1}{TW}$$

$$\Phi(f_0)=\frac{1}{2}\varphi(0)+\varphi\left(\frac{1}{2W}\right)\cos 2\pi f_0\left(\frac{1}{2W}\right)+\varphi\left(\frac{2}{2W}\right)\cos 2\pi f_0\left(\frac{2}{2W}\right)+\cdots+\varphi\left(\frac{N}{2W}\right)\cos 2\pi f_0\left(\frac{N}{2W}\right)$$

$$\Phi(2f_0)=\frac{1}{2}\varphi(0)+\varphi\left(\frac{1}{2W}\right)\cos 2\pi 2f_0\left(\frac{1}{2W}\right)+\varphi\left(\frac{2}{2W}\right)\cos 2\pi 2f_0\left(\frac{2}{2W}\right)+\ldots+\varphi\left(\frac{N}{2W}\right)\cos 2\pi 2f_0\left(\frac{N}{2W}\right)$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\Phi(pf_0)=\left(\frac{1}{2}\right)\varphi(0)+\varphi\left(\frac{1}{2W}\right)\cos 2\pi pf_0\left(\frac{1}{2W}\right)+\varphi\left(\frac{2}{2W}\right)\cos 2\pi pf_0\left(\frac{2}{2W}\right)+\cdots+\varphi\left(\frac{N}{2W}\right)\cos 2\pi pf_0\left(\frac{N}{2W}\right) \quad (21)$$

From Equation 21 it is seen that signals proportional to discrete values of autocorrelation amplitude spectrum may be derived by various linear combinations of the autocorrelation control signals, after weighting the control signals by signals proportional to selected cosine functions.

Referring to the apparatus shown in FIG. 6A, network 60 consists of an array of *p.N* multipliers m_{ij}, $i=1, 2, \ldots, p$, $j=1, 2, \ldots, N$, of conventional construction, arranged in *p* rows and N columns. Each row of multipliers corresponds to one series in Equation 21, and each multiplier corresponds to one term containing a cosine factor in a particular series. Each multiplier is provided with two input terminals and one output terminal, and the multipliers in the jth column have one of their input terminals connected to a common input point, I_j, to which point the jth autocorrelation control signal

$$\varphi\left(\frac{j}{2W}\right)$$

is applied. The second input terminals of the multipliers in each row are connected to the output terminals of one of the weighting signal sources S1, S2, . . ., S*p*, each of which generates N weighting signals proportional to the N cosine factors in the particular series of Equation 21 to which each row of multipliers corresponds.

From the autocorrelation control signals and the weighting signals, the multipliers in each row develop at their output terminals weighted autocorrelation signals proportional to the individual weighted terms of a particular series of Equation 21. The output terminals of the multipliers in the ith row are connected to one input terminal of adder A*i*, and the control signal $\varphi(0)$, after a reduction in amplitude achieved by passing it through amplifier 62 with a gain constant of ½, is applied to the other input terminal of the ith adder. The linear combination of weighted control signals formed at the output terminal of each adder is proportional to a discrete value of the autocorrelation amplitude spectrum, in accordance with Equation 21.

It is observed in Equations 9 and 21 that the autocorrelation amplitude spectrum is a cosine function of the fundamental pitch frequency, f_0, of the original speech wave. The variable nature of the pitch frequency has been noted and discussed earlier in connection with the gating apparatus of FIG. 2A. In the present discussion of network 60, the weighting signals supplied to the multipliers from sources S1, S2, . . ., S*p* must vary with changes in f_0 in order to obtain signals that accurately represent the amplitude spectrum. A signal that changes in synchrony with variations in f_0 is already available in the form of the pitch control signal output of circuit PC2 of FIG. 2A, and from this signal, after appropriate modification, sources S1, S2, . . ., Sp derive the various weighting signals utilized in network 60.

It is observed in Equation 21 that in addition to the variable f_0 the arguments of the cosine weighting factors contain three constants, denoted

$$a=\frac{\pi}{W},\ b=1, 2, \ldots, p, \text{ and } c=1, 2, \ldots, N$$

It is necessary, therefore, that before generating the cosine weighting signals, the magnitude of the pitch control signal must be appropriately amplified in order to form the proper argument for the cosine signals.

Since the first constant

$$a=\left(\frac{\pi}{W}\right)$$

appears in the arugment of every weighting factor, the pitch control signal is first passed through amplifier 6, for example, a conventional voltage amplifier, having a gain constant

$$\left(\frac{\pi}{W}\right)$$

The second constant $b=1, 2, \ldots, p$ is the same in the argument of each cosine factor appearing in a single series of (21). Since each source generates signals corresponding to the cosine factors in a single series, each source S2, . . ., Sp (except S1, where $b=1$) is preceded by an amplifier K2, . . ., Kp having a gain constant 2, . . . , p, respectively, appropriate to the value of b for that source. Thus the output signal of amplifier 6 is applied in parallel to the input terminals of source S1 and amplifiers K2, . . ., Kp thereby producing at the input terminals of sources S1, 2, . . ., Sp signal proportional to

$$\left(\frac{\pi}{W}\right)f_0,\ \left(\frac{\pi}{W}\right)2f_0,\ \ldots,\ \left(\frac{\pi}{W}\right)pf_0$$

respectively.

The third constant $c=1, 2, \ldots$, N differs from term-to-term within a single series of (21) and each source is therefore provided with a bank of N−1 amplifiers Ki2, . . ., KiN, $i=1, 2, \ldots, p$, with gain constants 2, . . ., N respectively, it being understood that no amplifier is necessary for the constant $c=1$.

In addition, each source contains a bank of N cosine function generators, of any well-known variety, each of which develops at its output terminal a signal whose magnitude is proportional to the cosine of the magnitude of the signal applied to its input terminal. For example, in the generation of weighting signals by source S2, the output signal of amplifier K2 is applied in parallel to cosine generator C21, for $c=1$, and to amplifiers K22, . . ., K2N, having gain constants 2, . . ., N, respectively. Hence the signals appearing at the input terminals of cosine function generators C21, C22, . . ., C2N are proportional to

$$\left(\frac{\pi}{W}\right)2f_0,\ 2\left(\frac{\pi}{W}\right)2f_0,\ \ldots,\ N\left(\frac{\pi}{W}\right)2f_0$$

respectively, and the signals developed at the output terminals of these generators are proportional to

$$\cos 2\pi 2f_0\left(\frac{1}{2W}\right),\ \cos 2\pi 2f_0\left(\frac{2}{2W}\right),\ \ldots,\ \cos 2\pi 2f_0\left(\frac{N}{2W}\right)$$

respectively.

The number of weighting signal sources utilized depends upon the frequency resolution desired in the amplitude spectrum of the reconstructed speech wave. For maximum resolution of voices having a fundamental pitch frequency of 100 cycles per second and a bandwith of 3,000 cycles per second, for example,

$$p=\frac{W}{f_0}=30$$

sources are required.

The number of cosine function generators contained in each source is determined by the maximum number of control signals to be weighted. As previously discussed, the maximum number of control signals depends upon the longest period which it is desired to represent fully. Thus for full representation of 10 millisecond periods, a maximum of 30 control signals must be weighted and each weighting signal source must contain 30 cosine function generators.

Having obtained signals proportional to discrete values of the autocorrelation amplitude spectrum at the output points of network 60, the signals are applied to a bank of rooters H61, H62, . . ., H6p, of any suitable variety. The rooters develop at their respective output points signals proportional to the square roots of the individual spectrum values,

$$\Phi^{1/2}(f_0),\ \Phi^{1/2}(2f_0),\ \ldots,\ \Phi^{1/2}(pf_0)$$

From Equation 3b, it is apparent that the rooters produce a group of signals representing an amplitude spectrum that is of the same shape as the amplitude spectrum of the original speech wave, thereby removing the spectrum squaring distortion inherent in the autocorrelation function representation of speech.

In order to reconstruct speech in a synthesizer of the type shown in FIG. 3, the rooted spectrum signals must be converted into their autocorrelation function counterparts. This is achieved by passing the rooted spectrum signals through the apparatus of FIG. 6B, which converts the rooted spectrum signals into autocorrelation samples in accordance with the inverse Fourier transformation of Equation 9a.

From Equation 9a, the autocorrelation function samples corresponding to the rooted spectrum signals are given by the following group of cosine series:

$$\begin{aligned}
\varphi'(0)&=\Phi^{1/2}(f_0)+\Phi^{1/2}(2f_0)+\ldots+\Phi^{1/2}(pf_0)\\
\varphi'\left(\frac{1}{2W}\right)&=\Phi^{1/2}(f_0)\cos 2\pi f_0\left(\frac{1}{2W}\right)+\Phi^{1/2}(2f_0)\cos 2\pi 2f_0\left(\frac{1}{2W}\right)+\ldots+\Phi^{1/2}(pf_0)\cos 2\pi pf_0\left(\frac{1}{2W}\right)\\
\varphi'\left(\frac{2}{2W}\right)&=\Phi^{1/2}(f_0)\cos 2\pi f_0\left(\frac{2}{2W}\right)+\Phi^{1/2}(2f_0)\cos 2\pi 2f_0\left(\frac{2}{2W}\right)+\ldots+\Phi^{1/2}(pf_0)\cos 2\pi pf_0\left(\frac{2}{2W}\right)\\
&\quad\cdot\qquad\cdot\qquad\cdot\\
\varphi'\left(\frac{N}{2W}\right)&=\Phi^{1/2}(f_0)\cos 2\pi f_0\left(\frac{N}{2W}\right)+\Phi^{1/2}(2f_0)\cos 2\pi 2f_0\left(\frac{N}{2W}\right)+\ldots+\Phi^{1/2}(pf_0)\cos 2p\pi f_0\left(\frac{N}{2W}\right)
\end{aligned}\tag{22}$$

It is observed in Equation 22 that half-period samples of the autocorrelation function $\varphi'(\tau)$ corresponding to the rooted spestrum $\Phi^{1/2}(nf_0)$ may be derived from various linear combinations of rooted spectrum signals, after weighting the rooted spectrum signals by selected cosine function signals.

A comparison of Equations 21 and 22 reveals that the transformation of the rooted spectrum signals into autocorrelation signals may be accomplished by apparatus similar in structure to that shown in FIG. 6A. Referring now to FIG. 6B, network 61 consists of an array of $p\cdot N$ multipliers m'_{ij}, $i=1, 2, \ldots, p$, $j=1, 2, \ldots$, N, arranged in p rows and N columns, each column of multipliers corresponding to one cosine series in Equation 22, and each multiplier in each column corresponding to one term in a particular cosine series. Each multiplier is provided with two input terminals and one output terminal, and the multipliers in the ith row have one of their input terminals connected to a common input point, I'_i, to which point the ith rooted spectrum signal, $\Phi^{1/2}(if_0)$, is applied. The second input terminals of the multipliers in the ith row are supplied with weighting signals from the ith source Si of FIG. 6A. The sources supply weighting signals proportional to the cosine factors in Equation 22, and the p multipliers in each column develop at their output terminals weighted spectrum signals proportional to the p individual cosine terms of a particular series of Equation 22. The output terminals of the multipliers in the jth column are connected to a common output point $0'_j$, and the linear combination of weighted spectrum signals formed at each common output point is proportional to a specific value of the corresponding autocorrelation function $\varphi'(\tau)$ in accordance with Equation 22.

The value of $\varphi'(\tau)$ for $\tau=0$ is derived in accordance with the first series of Equation 22 by applying the rooted spectrum signals to the input terminals of an adder 64 of conventional construction, and the output signal developed by adder 64 is proportional to $\varphi'(0)$.

Artificial speech is reconstructed from the auto-correlation signals produced by network **61** by applying them together with the pitch control signal from circuit PC**2** of the analyzer apparatus of FIG. 2A to a synthesizer such as that shown in FIG. 3 of this application. Artificial speech reconstructed from these signals has an amplitude spectrum of the same shape as the amplitude spectrum of the original speech wave; hence the artificial speech is free of the distortion caused by spectrum squaring.

FIGS. 7A, 7B, 7C illustrate apparatus for unsquaring the amplitude spectrum of control signals produced by an analyzer of the type shown in FIG. 4. The apparatus shown in FIGS. 7A, 7B, and 7C operates upon the same principles as that of FIGS. 6A and 6B to transform the incoming autocorrelation control signals into one-sided amplitude spectrum signals, to take the square root of the spectrum signals, and to convert the rooted spectrum signals into a new set of autocorrelation control signals.

Referring to Equation 17, the one-sided spectrum may be derived from the half-period samples of $\varphi(\tau)$ and $\varphi(\tau)$ by the following transformation, ignoring the constant factor $\frac{2}{TW}$

$$\Phi_W(f_0)=\frac{1}{2}\varphi(0)+\sum_{j=1}^{N=\frac{TW}{2}}\varphi\left(\frac{j}{W}\right)\cos 2\pi f_0\left(\frac{j}{W}\right)+\sum_{j=1}^{N=\frac{TW}{2}}\hat{\varphi}\left(\frac{j}{W}\right)\cos 2\pi 2f_0\left(\frac{j}{W}\right)$$

$$\Phi_W(2f_0)=\frac{1}{2}\varphi(0)+\sum_{j=1}^{N=\frac{TW}{2}}\varphi\left(\frac{j}{W}\right)\cos 2\pi 2f_0\left(\frac{j}{W}\right)+\sum_{j=1}^{N=\frac{TW}{2}}\hat{\varphi}\left(\frac{j}{W}\right)\sin 2\pi 2f_0\left(\frac{j}{W}\right)$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\Phi(pf_0)=\frac{1}{2}\varphi(0)+\sum_{j=1}^{N=\frac{TW}{2}}\varphi\left(\frac{j}{W}\right)\cos 2\pi pf_0\left(\frac{j}{W}\right)+\sum_{j=1}^{N=\frac{TW}{2}}\hat{\varphi}\left(\frac{j}{W}\right)\sin 2\pi pf_0\left(\frac{j}{W}\right) \quad (23)$$

It is apparent from Equation 23 that discrete values of the one-sided spectrum may be obtained by forming various linear combinations of the two sets of autocorrelation samples, after weighting the $\varphi(\tau)$ samples by selected cosine function signals and weighting the $\hat{\varphi}(\tau)$ samples by selected sine function signals.

Referring now to the apparatus shown in FIG. 7A, network **70** consists of an array of $p.2N$ multipliers x_{ij}, y_{ij}, $i=1,2,\ldots,p$, $j=1,2,\ldots,N$, arranged in p rows and 2N columns, each row of multipliers corresponding to one series in Equation 23, and each multiplier in each row corresponding to one term in a particular series. Each multiplier is provided with two input terminals and one output terminal, and the multipliers x_{ij}, y_{ij} in the jth columns have one of their input terminals connected to common input points, I_j, $\hat{I}_j$, to which points the jth autocorrelation signals

$$\varphi\left(\frac{j}{W}\right),\ \hat{\varphi}\left(\frac{j}{W}\right)$$

respectively, are applied, in accordance with Equation **23**. The second input terminals of the multipliers in each row are connected to the output terminals of one of the weighting signals source W**1**, W**2**, . . ., Wp. Each source generates N signals proportional to the cosine factors and N signals proportional to the sine factors in the particular series of Equation 23 to which each row of multipliers corresponds.

From the autocorrelation control signals and the weighting signals the multipliers in each row develop at their output terminals weighted autocorrelation signals proportional to the terms of a particular series of Equation 23. The output terminals of the multipliers in the ith row are connected to one input point of adder Bi, and the control signal $\varphi(0)$, after a reduction in magnitude achieved by passing it through amplifier **72**, having a gain constant of

$$\frac{1}{2}$$

is applied to the other input terminal of the ith adder. The linear combination of weighted control signals formed at the output terminal of each adder is proportional to a discrete value of the one-sided amplitude spectrum, in accordance with Equation 23.

It is noted in Equation 23 that the one-sided amplitude spectrum is also a function of the fundamental pitch frequency, f_0, of the original speech. In order to reproduce the spectrum accurately, the weighting signals generated by sources W**1**, W**2**, . . ., Wp, must vary with changes in f_0. Sources W**1**, W**2**, . . ., Wp derive the various weighting signals from the pitch control signal produced by circuit PC**4** of FIG. 4, thereby producing weighting signals whose magnitudes vary in synchrony with variations in f_0.

It is observed in Equation 23, as in Equation 21, that in addition to the variable f_0 the arguments of both the sine and cosine weighting factors contain three constants, denoted

$$a=\left(\frac{2\pi}{W}\right),\ b=1, 2, \ldots, p, \text{ and } c=1, 2, \ldots, N$$

As in the weighting signal sources of the apparatus of FIG. 6A, the pitch control signal must be appropriately modified in order to form the proper argument for the weighting signals.

Since the first constant

$$a=\left(\frac{2\pi}{W}\right)$$

appears in the argument of every weighting factor, the pitch control signal is first passed through amplifier **7**, having a gain constant

$$\left(\frac{2\pi}{W}\right)$$

Within a single series of (23) the constant $b=1,2, \ldots, p$ is the same in the argument of each sine and cosine factor. Since each source generates signals corresponding to the sine and cosine factors in a single series each source W2, . . ., Wp (except W1, where $n=1$) is preceded by an amplifier K2, . . ., Kp having a gain constant 2, . . ., p, respectively, appropriate to the value of b for that source. Thus the output signal of amplifier 7 is applied in parallel to the input terminals of source W1 and amplifiers K2, . . ., Kp, and the signals appearing at the input terminals of sources W1, W2, . . ., Wp are proportional to

$$\left(\frac{2\pi}{W}\right)f_0, \left(\frac{2\pi}{W}\right)2f_0, \ldots, \left(\frac{2\pi}{W}\right)pf_0$$

respectively.

The third constant $c=1,2, \ldots,$ N differs from one pair of cosine and sine factors to the next pair in a single series of Equation 23, consequently each source is provided with a group of amplifiers whose gain constants are adjusted to the appropriate values of c. In FIG. 7A, only source W2 is shown in detail, the other sources being similar in structure.

As shown in FIG. 7A, source W2 contains N−1 amplifiers K22, . . ., K2N, with gain constants 2, . . ., N, respectively, it being understood that no amplifier is necessary for the case $c=1$. In addition, source W2 contains N pairs of sine and cosine function generators, of any well-known variety, each pair of which is connected in tandem with the appropriate amplifier. Thus the modified pitch control signal

$$\left(\frac{2\pi}{W}\right)2f_0$$

is applied in parallel to the input terminals of generators CF21, SF21, and to the input terminals of amplifiers K22, . . ., K2N. The output terminals of K22, . . ., K2N are connected to the input terminals of the pairs of generators CF22, SF22, . . ., CF2N, SF2N, respectively, and the weighting signals developed at the output terminals of the generators are proportional to the cosine and sine factors in the second series of Equation 23, as shown in FIG. 7A.

The number of weighting signal sources utilized depends upon the frequency resolution desired in the one-sided amplitude spectrum. For example, maximum resolution of voices having a fundamental pitch frequency of 100 cycles per second and a bandwidth of 3,000 cycles per second requires

$$p=\frac{W}{f_0}=30$$

sources. Further, the number of cosine and sine function generators contained in each source is determined by the maximum number of control signals to be weighted. For example, 15 samples of $\varphi(\tau)$, in addition to $\varphi(0)$, and 15 samples of $\hat{\varphi}(\tau)$, require 15 pairs of cosine and sine function generators in each source.

The one-sided amplitude spectrum signals appearing at the output terminals of network 70 are applied to a bank of rooters H71, H72, . . ., H7p, of a suitable construction. The rooters develop at their respective output points signals proportional to the square roots of the individual spectrum values,

$$\Phi_W^{1/2}(f_0),\ \Phi_W^{1/2}(2f_0), \ldots, \Phi_W^{1/2}(pf_0)$$

The rooted spectrum signals represent an amplitude spectrum having the same shape as the amplitude spectrum of the original speech wave, and artificial speech reconstructed from the rooted spectrum signals is free of the distortion associated with amplitude spectrum squaring.

Before reconstructing artificial speech from the rooted spectrum signals in a synthesizer of the type shown in FIG. 5, the rooted spectrum signals must be converted into autocorrelation function samples. The conversion is performed by passing the rooted spectrum output signals of the rooters through network 71 of FIG. 7B, in accordance with the inverse Fourier transformation of Equation 20.

The Fourier transformation of Equation 20 applied to the rooted spectrum signals may be rewritten in the following form, disregarding the constant factor

$$\frac{1}{2}$$

$$\varphi'_W\left(\frac{j}{W}\right)=\sum_{n=0}^{p}\left[\Phi_W^{1/2}(nf_0)+\Phi_W^{1/2}(pf_0-nf_0)\right]\cos 2\pi nf_0\left(\frac{j}{W}\right)+\left[\Phi_W^{1/2}(nf_0)-\Phi_W^{1/2}(pf_0-nf_0)\right]\sin 2\pi nf_0\left(\frac{j}{W}\right) \quad (24)$$

$j=1,2, \ldots,$ N. Let

$$\varphi_c\left(\frac{j}{W}\right)$$

denote the even part of

$$\varphi'_W\left(\frac{j}{W}\right)$$

and let

$$\varphi_s\left(\frac{j}{W}\right)$$

denote the odd part; then Equation 24 may be rewritten

$$\varphi'_W\left(\frac{j}{W}\right)=\varphi_c\left(\frac{j}{W}\right)+\varphi_s\left(\frac{j}{W}\right) \quad (25)$$

where

$$\begin{aligned}\varphi_c\left(\frac{j}{W}\right)&=\sum_{n=0}^{p}\left[\Phi_W^{1/2}(nf_0)+\Phi_W^{1/2}(pf_0-nf_0)\right]\cos 2\pi nf_0\left(\frac{j}{W}\right)\\&=\Phi_W^{1/2}(pf_0)+\left[\Phi_W^{1/2}(f_0)+\Phi_W^{1/2}(pf_0-f_0)\right]\cos 2\pi f_0\left(\frac{j}{W}\right)\\&\quad+\cdots+\Phi_W^{1/2}(pf_0)\cos 2\pi pf_0\left(\frac{j}{W}\right)\end{aligned} \quad (26)$$

and

$$\begin{aligned}\varphi_s\left(\frac{j}{W}\right)&=\sum_{n=0}^{p}\left[\Phi_W^{1/2}(nf_0)-\Phi_W^{1/2}(pf_0-nf_0)\right]\sin 2\pi nf_0\left(\frac{j}{W}\right)\\&=\left[\Phi_W^{1/2}(f_0)-\Phi_W^{1/2}(pf_0-f_0)\right]\sin 2\pi f_0\left(\frac{j}{W}\right)\\&\quad+\cdots+\Phi_W^{1/2}(pf_0)\sin 2\pi pf_0\left(\frac{j}{W}\right)\end{aligned} \quad (27)$$

Apparatus for reconstructing samples of the even part of

$$\varphi'_W\left(\frac{j}{W}\right)$$

is shown in FIG. 7B, based upon Equation 26. The rooted spectrum signals from rooters H71, H72, . . ., H7p are added in adders D1, D2, . . ., in the order given by Equation 26, and the output signals of the adders, which are proportional to the coefficients in Equation 26, are applied to network 71 in order to form the weighted cosine terms of Equation 26.

It is noted in Equation 26 that the spectrum coefficients from

$$f_0 \text{ to } \frac{pf_0}{2}$$

are identical to the coefficients from

$$pf_0 \text{ to } \frac{pf_0}{2}$$

hence the output signal of each adder D1, D2, . . ., may be used as the input signal for two input terminals of network 71, as shown in FIG. 7B. The cosine weighting signals from sources W1, W2, . . ., Wp of FIG. 7A are also applied to network 71, which consists of an array of $p \cdot N$ multipliers m_{ij}, $i=1, 2, \ldots, p$, $j=1, 2, \ldots, N$, arranged in p rows and N columns.

Each column of multipliers corresponds to a particular value of

$$\varphi_c\left(\frac{j}{W}\right)$$

and each multiplier in each column corresponds to a cosine term on the right-hand side of Equation 26. Each multiplier is provided with two input terminals and one output terminal, and the multipliers in the ith row have one of their input terminals connected to one of the output terminals of weighting source Wi of FIG. 7A.

The output terminals of the multipliers in the jth column are connected to one input terminal of adder Vj, and the rooted spectrum signal

$$\Phi_W^{1/2}(pf_0)$$

is applied to the other input terminal of the jth adder. The linear combination of weighted spectrum signals developed at the output terminal of each adder is proportional to a discrete value of

$$\varphi_c\left(\frac{j}{W}\right)$$

in accordance with Equation 26. The signal proportional to $\varphi_c(0)$ is formed by combining the outputs of adders D1, D2, . . ., in adder 75.

Apparatus for reconstructing samples of the odd part of

$$\varphi_w'\left(\frac{j}{W}\right)$$

is shown in FIG. 7C, based upon Equation 27. The rooted spectrum signals from rooters H71, H72, . . ., H7p of FIG. 7A are subtracted in elements E1, E2, . . ., of any well-known construction, in the order given by Equation 27, thereby forming a group of signals proportional to the coefficients of Equation 27. It is noted in Equation 27 that the coefficients from

$$f_0 \text{ to } \frac{pf_0}{2}$$

are antisymmetrical with respect to the coefficients from

$$\frac{pf_0}{2} \text{ to } pf_0$$

hence it is only necessary to perform half of the subtractions indicated in Equation 27, the remaining coefficients being obtained by reversing the polarities of the output signals of the subtracting elements. As shown in FIG. 7C, the output signals of elements E1, E2, . . ., are applied directly to half of the input terminals of network 73, but are reversed in polarity by polarity inverters PI1, PI2, . . ., prior to applying the signals to the remaining input terminals of network 73.

Network 73 consists of an array of $p \cdot N$ multipliers in an arrangement identical to that of network 71 of FIG. 7B. The multipliers in each column, supplied with the output signals from elments E1, E2, . . ., and with the sine weighting signals from the appropriate weighting signal source of FIG. 7A, develop at their output terminals signals proportional to the individual terms of Equation 27 for a particular value of j. The output terminals of the multipliers in the jth column are connected to a common output point, 0_j, and the linear combination of signals formed at each common output point is proportional to a particular value of

$$\hat{\varphi}\left(\frac{j}{W}\right)$$

The output signals of networks 71 and 73 constitute a set of control signals from which artificial speech may be reconstructed in apparatus of the type shown in FIG. 5.

Referring now to FIG. 5, the output signals of network 71 are applied to the control terminals of the modulators to which samples of $\hat{\varphi}(\tau)$ are shown as being applied, since

$$\varphi_c\left(\frac{j}{W}\right)$$

is symmetrical about the center of each period. Similarly, the output signals of network 73 are applied to the control terminals of those modulators which are shown supplied with samples of $\hat{\varphi}(\tau)$, since

$$\varphi_s\left(\frac{j}{W}\right)$$

is antisymmetrical about the center of each period.

Artificial speech reconstructed from the output signals of networks 71 and 73 has an amplitude spectrum of the same shape as the amplitude spectrum of the original speech, thus the artificial speech produced by the apparatus of FIG. 5 is free of distortion due to squaring of the amplitude spectrum.

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 system for the narrow-band transmission of speech, the combination that comprises a source of an incoming speech wave, means for obtaining from said speech wave a signal representative of the instantaneous pitch characteristic of said speech wave, means for correlating said speech wave with itself to produce speech autocorrelation function samples, means under the control of said pitch signal for gating said speech autocorrelation function samples to obtain a group of narrow-band control signals which vary in number in synchrony with changes in said pitch characteristic, means for transmitting said pitch signal and said control signals to a receiver station, and, at said receiver station, means for reconstructing an artificial speech wave from said transmitted pitch signal and said transmitted control signals.

2. In a system for the narrow-band transmission of speech, the combination that comprises means connected to a source of an incoming speech wave for obtaining a signal indicative of the instantaneous pitch characteristic of said speech wave, means for correlating said speech wave with itself to obtain samples of the speech autocorrelation function, means responsive to said pitch signal for gating said speech auto-correlation function samples to obtain a group of narrow-band control signals which vary in number in synchrony with changes in said pitch characteristic, means for unsquaring the amplitude spectrum of said control signals, means for transmitting said pitch signal and said unsquared control signals to a receiver station, and, at said receiver station, means for generating an excitation signal from said pitch signal, and means responsive to said unsquared control signal for synthesizing an artificial speech wave from said excitation signal.

3. In a system for the narrow-band transmission of speech, the combination that comprises means connected to a source of an incoming speech wave for obtaining a signal indicative of the instantaneous pitch characteristic of said speech wave, means for correlating said speech wave with itself to obtain a group of signals representative of a speech autocorrelation function, means under the control of said pitch signal for varying the number of said autocorrelation signals to derive a group of control signals representative of half of each period of said speech autocorrelation function, means for transmitting said pitch signal and said control signals to a receiver station, and,

at said receiver station, means for obtaining an excitation signal from said pitch signal, and means under the influence of said control signals for synthesizing full periods of an artificial speech wave from said excitation signal.

4. Apparatus as defined in claim 3 wherein said means for varying the number of said autocorrelation signals comprises a plurality of gating means in one-to-one correspondence with said autocorrelation signals, wherein each of said gating means is provided with a preselected threshold and a control terminal, an input terminal, and an output terminal, means for applying each autocorrelation signal to the input terminal of its corresponding gating means, and means for applying said pitch signal to the control terminals of all of said gating means to obtain at the output terminals of said gating means a group of control signals representative of half of each period of said autocorrelation function.

5. Speech transmission apparatus that comprises a source of an incoming speech wave, means for obtaining from said speech wave a signal representative of the instantaneous pitch characteristic of said speech wave, means for correlating said speech wave with itself to obtain speech autocorrelation function samples, means under the control of said pitch signal for gating said speech autocorrelation function samples to obtain a group of narrow-band control signals representative of half periods of said autocorrelation function, means for transmitting pitch signal and said control signals to a receiver station, and, at said receiver station, means responsive to said pitch signal for unsquaring the amplitude spectrum of said control signals, means for deriving an excitation signal from said pitch signal, and means responsive to said unsquared control signals for reconstructing full periods of an artificial speech wave from said excitation signal.

6. Apparatus as defined in claim 5 wherein said means for unsquaring the amplitude spectrum of said control signals comprises a first amplifier having a selected gain constant, means for applying said pitch signal to said first amplifier, a plurality of second amplifiers having selected gain constants, means for connecting the output terminal of said first amplifier in parallel to the input terminals of said second amplifiers, a plurality of weighting signal sources wherein each of said weighting signal sources comprises a plurality of third amplifiers having selected gain constants connected in tandem with a plurality of cosine function generators, means for connecting the output terminals of said second amplifiers to the input terminals of said weighting signal sources, a first array of multipliers arranged in rows and columns, wherein each multiplier is provided with two input terminals and one output terminal, one of the input terminals of each multiplier in each of said columns is connected to a common input point for each column, the other input terminal of each of said multipliers is connected to one of said cosine function generators, and the output terminals of the multipliers in each row are connected to the input point of an adder for each row, a plurality of square-root-taking devices, the input terminal of each of which is connected to the output point of the adder for each row of said first array of multipliers, a second array of multipliers arranged in rows and columns, wherein each of said multipliers is provided with two input terminals and one output terminal, one of the input terminals of each multiplier in each of said rows is connected to a common input point to which is connected the output terminal of one of said square-root-taking devices, each of the other input terminals of said multipliers is connected to one of said cosine function generators, and the output terminals of the multipliers in each column are connected to a common output point for each row.

7. In a system for the transmission of speech, the combination that comprises a source of an incoming speech wave, means for shifting the phase of said incoming speech wave by ninety degrees to obtain its Hilbert transform, means for correlating said speech wave with itself and with its Hilbert transform to obtain speech autocorrelation function samples, means for developing from said speech wave a pitch signal indicative of the instantaneous pitch characteristic of said speech wave, means actuated by said pitch signal for gating said autocorrelation function samples to obtain a group of narrow-band control signals representative of half periods of said autocorrelation function, means for transmitting said pitch signal and said control signals to a receiver station, and, at said receiver station, means responsive to said pitch signal for unsquaring the amplitude spectrum of said control signals, means supplied with said pitch signal for generating an excitation signal, and means under the influence of said unsquared control signals for reconstructing full periods of an artificial speech wave from said excitation signal.

8. Apparatus as defined in claim 7 wherein said means for unsquaring the amplitude spectrum of said control signals comprises a first amplifier having a selected gain constant, means for applying said pitch signal to said first amplifier, a plurality of second amplifiers having selected gain constants, means for connecting the output terminal of said first amplifier in parallel to the input terminals of said second amplifiers, a plurality of weighting signal sources wherein each of said weighting signal sources comprises a plurality of third amplifiers having selected gain constants, a plurality of cosine function generators paired with a plurality of sine function generators, and means for connecting the output terminal of each of said third amplifiers in parallel to the input terminals of one of said pairs of function generators, means for connecting the output terminals of said second amplifiers to the input terminals of said weighting signal sources, a first array of multipliers arranged in horizontal rows and vertical columns, wherein each multiplier is provided with two input terminals and one output terminal, one of the input terminals of each multiplier in each of said columns is connected to a common input point, the other input terminal of each of said multipliers is connected to one of said function generators, and the output terminals of the multipliers in each row are connected to the input point of an adder for each row, a plurality of square-root-taking devices, the input terminal of each of which is connected to the output point of the adder for each row of said first array of multipliers, adding means connected to the output terminals of said square-root-taking devices, a second array of multipliers arranged in horizontal rows and vertical columns, wherein each multiplier is provided with two input terminals and one output terminal, one of the input terminals of each multiplier in each of said rows is connected to a common input point to which is connected the output terminal of one of said adding means, the other input terminal of each of said multipliers is connected to one of said cosine function generators, and the output terminals of the multipliers in each column are connected to the input terminal of an adder for each column, subtracting elements connected to the output terminals of said square-root-taking devices, a third array of multipliers arranged in horizontal rows and vertical columns, wherein each multiplier is provided with two input terminals and one output terminal, one of the input terminals of each multiplier in each of said rows is connected to a common input point to which is connected the output terminal of one of said subtracting elements, the other input terminal of each of said multipliers is connected to one of said sine function generators, and the output terminals of the multipliers in each column are connected to a common output point for each column.

9. Apparatus for the synthesis of artificial speech which comprises a source of incoming control signals representative of half periods of a speech autocorrelation function, a signal propagation device provided with a plurality of taps symmetrically disposed about a center

tap one end of said propagation device being terminated in a matched impedance, the other end of said propagation device being provided with an output terminal, means under the control of an incoming pitch control signal for generating an excitation signal, means for employing said control signals and said excitation signal to synthesize at the taps of said propagation device a group of artificial signals, and means for reproducing speech from the artificial signals appearing at the output terminal of said propagation device.

10. Apparatus for the synthesis of speech which comprises a source of incoming control signals representative of half periods of a symmetrical autocorrelation function derived from a speech wave, a source of an excitation signal, a plurality of modulating means in one-to-one correspondence with said control signals, wherein each of said modulating means is provided with a control terminal, an input terminal and an output terminal, a signal propagation device provided with a plurality of taps disposed in symmetrically located pairs about a center tap, wherein one end of said propagation device is terminated in a matched impedance, and the other end of said device is provided with an output terminal, means for applying said incoming control signals to the control terminals of said corresponding modulating means, means for applying said excitation signal in parallel to the input terminals of said modulating means, means for connecting the output terminal of one of said modulating means to the center tap of said propagation device, means for connecting the output terminal of each of said other modulating means to a pair of symmetrically located taps of said propagation device, and means connected to the output terminal of said propagation device for reproducing artificial speech.

11. Apparatus for the synthesis of speech which comprises a first source of incoming control signals representative of half periods of a symmetrical speech autocorrelation function, a second source of incoming control signals representative of half periods of an antisymmetrical speech autocorrelation function, a signal propagation device provided with a plurality of taps symmetrically located in equal numbers above and below a center tap, wherein one end of said propagation device is terminated in a matched impedance, and the other end of said propagation device is provided with an output terminal, means under the control of an incoming pitch signal for generating an excitation signal, means connected to the output terminal of said excitation signal generating means for reversing the polarity of said excitation signal, a plurality of first modulating means in one-to-one correspondence with the taps of said propagation device, wherein each of said first modulating means is provided with a control terminal, an input terminal, and an output terminal, means for connecting the output terminal of one of said first modulating means to the center tap of said propagation device, means for connecting the output terminal of each of said other first modulating means to its corresponding tap of said propagation device, means for applying one of said control signals from said first source to the control of the modulating means connected to the center tap of said propagation device, means for applying each of said other control signals from said first source to the control terminals of two of said first modulating means which are connected to symmetrically located taps of said propagation device, means for applying said excitation signal in parallel to the input terminal of each of said first modulating means, a plurality of second modulating means in one-to-one correspondence with the taps of said propagation device, wherein each of said second modulating means is provided with a control terminal, an input terminal, and an output terminal, means for connecting the output terminal of each of said second modulating means to its corresponding tap of said propagation device, means for applying each of said control signals from said second source to the control terminals of two of said second modulating means which are connected to symmetrically located taps of said propagation device, means for applying said excitation signal in parallel to the input terminals of those of said second modulating means which are connected to taps above the center tap of said propagation device, means for applying said reversed polarity excitation signal to the input terminals of those of said second modulating means which are connected to taps below the center tap of said propagation device, and means connected to the output terminal of said propagation device for reproducing artificial speech.

12. In a speech producing system, a source of an excitation signal, a source of incoming control signals representative of half periods of a speech autocorrelation function, means supplied with said control signals for unsquaring the amplitude spectrum of said control signals, and means responsive to said unsquared control signals for reconstructing full periods of an artificial speech wave from said excitation signal.

References Cited in the file of this patent

UNITED STATES PATENTS

2,908,761	Raisbeck	Oct. 13, 1959
2,953,645	Schroeder	Sept. 20, 1960