United States Patent [19]

Wolf

[11] **3,800,093**

[45] **Mar. 26, 1974**

[54] **METHOD OF DESIGNING ORTHOGONAL FILTERS**

[76] Inventor: **Alfred A. Wolf**, RFD 3, Box 327, Annapolis, Md. 21403

[22] Filed: **Oct. 20, 1971**

[21] Appl. No.: **191,003**

[52] **U.S. Cl.**..... **179/15 BC**, 179/15.55 R, 333/70 R

[51] **Int. Cl.** **H04b 1/66**

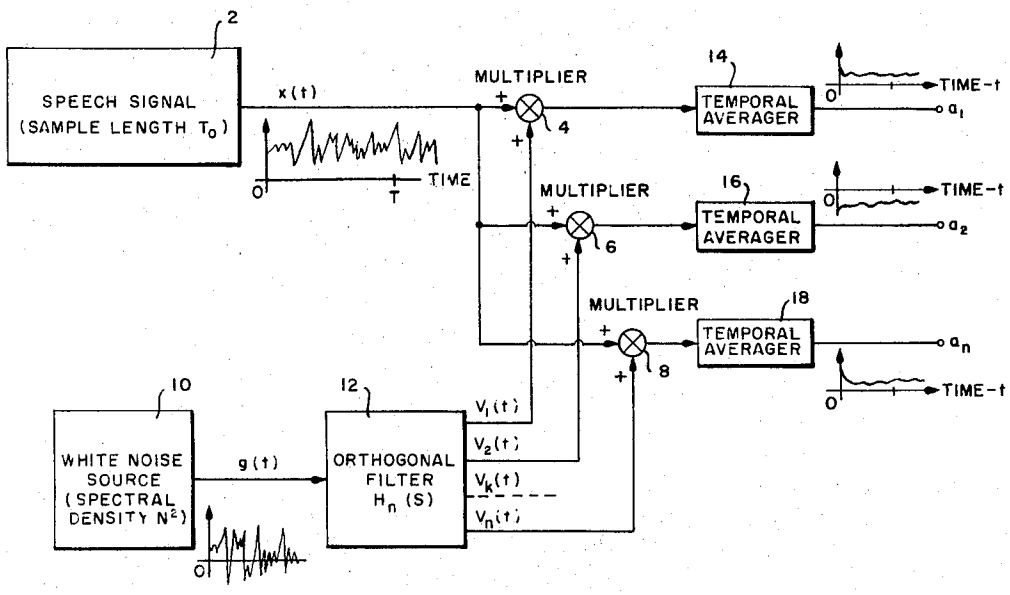
[58] **Field of Search** ..... 179/15 A, 15.55, 15 BC; 333/20, 70 R

Primary Examiner—Kathleen H. Claffy
Assistant Examiner—David L. Stewart
Attorney, Agent, or Firm—R. S. Sciascia; Q. E. Hodges

[57] **ABSTRACT**

A method of designing an orthogonal filter for use in a speech synthesizing system wherein the filter is constructed from a series of linear filter sections. The poles in the complex plane of any section of the filter are cancelled by the zeros of the next section.

4 Claims, 5 Drawing Figures



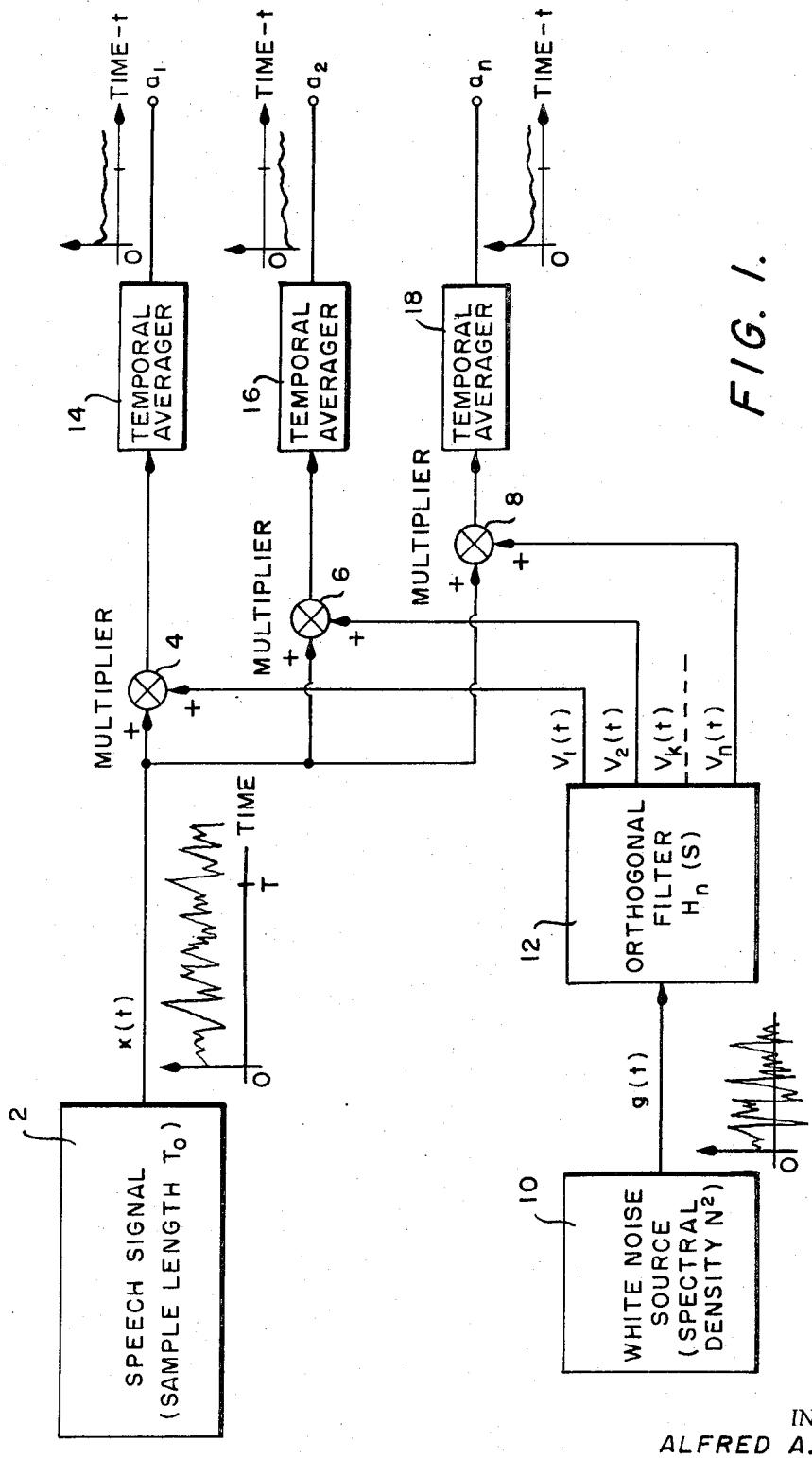


FIG. 1.

INVENTOR.
ALFRED A. WOLF

BY

J. C. Hodges
ATTORNEY

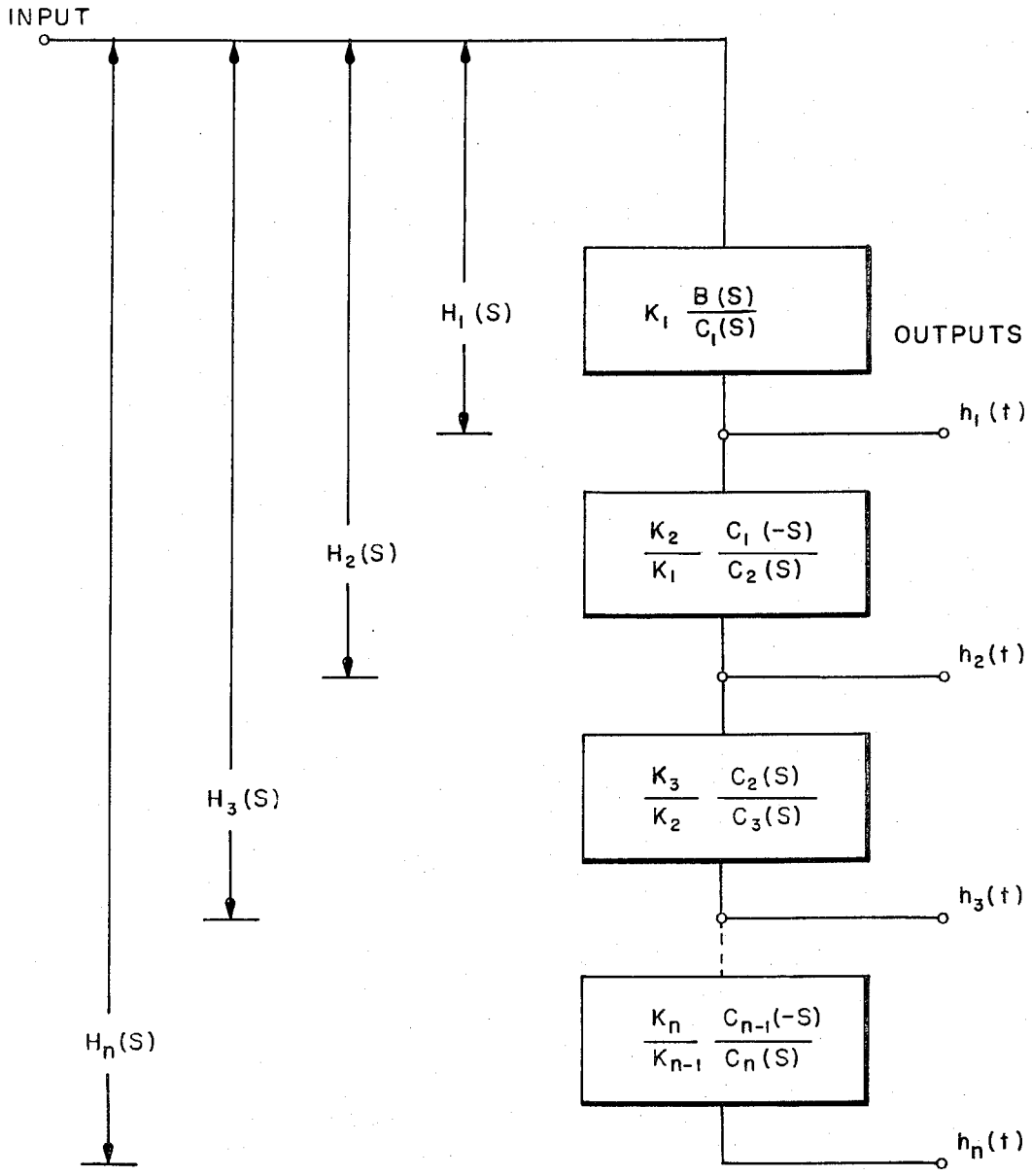


FIG. 2.

INVENTOR.
ALFRED A. WOLF

BY

ATTORNEY

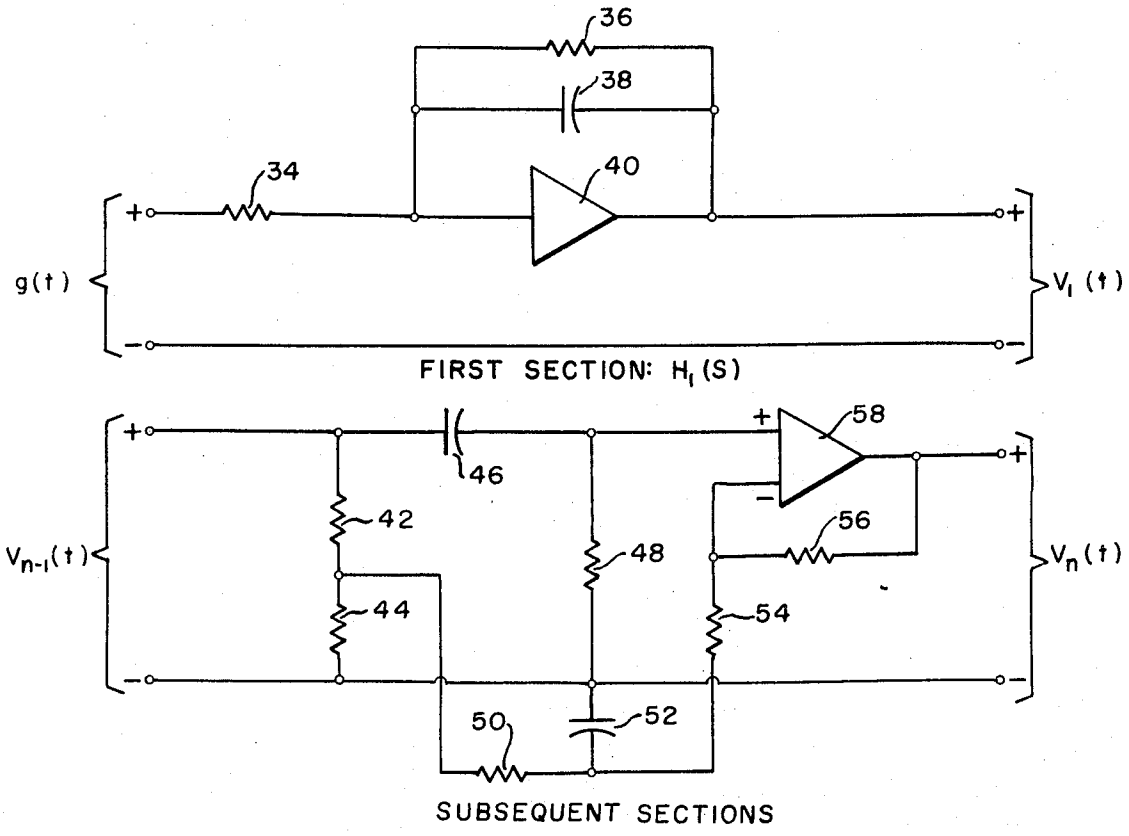


FIG. 3.

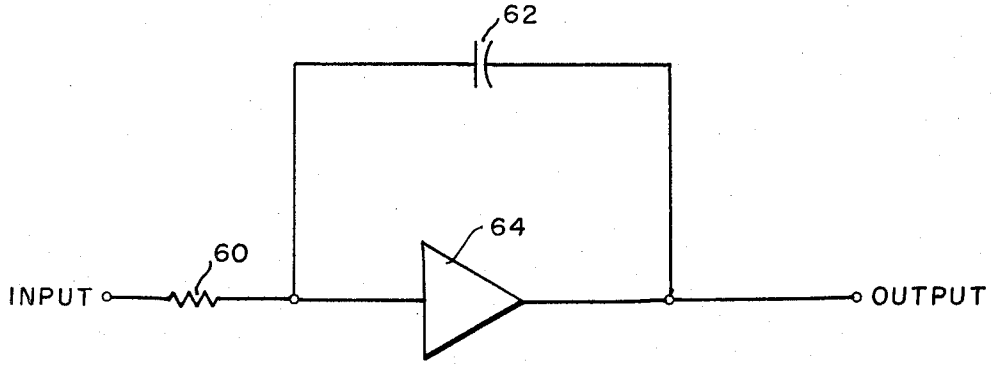


FIG. 4.

INVENTOR.
ALFRED A. WOLF
BY
ATTORNEY

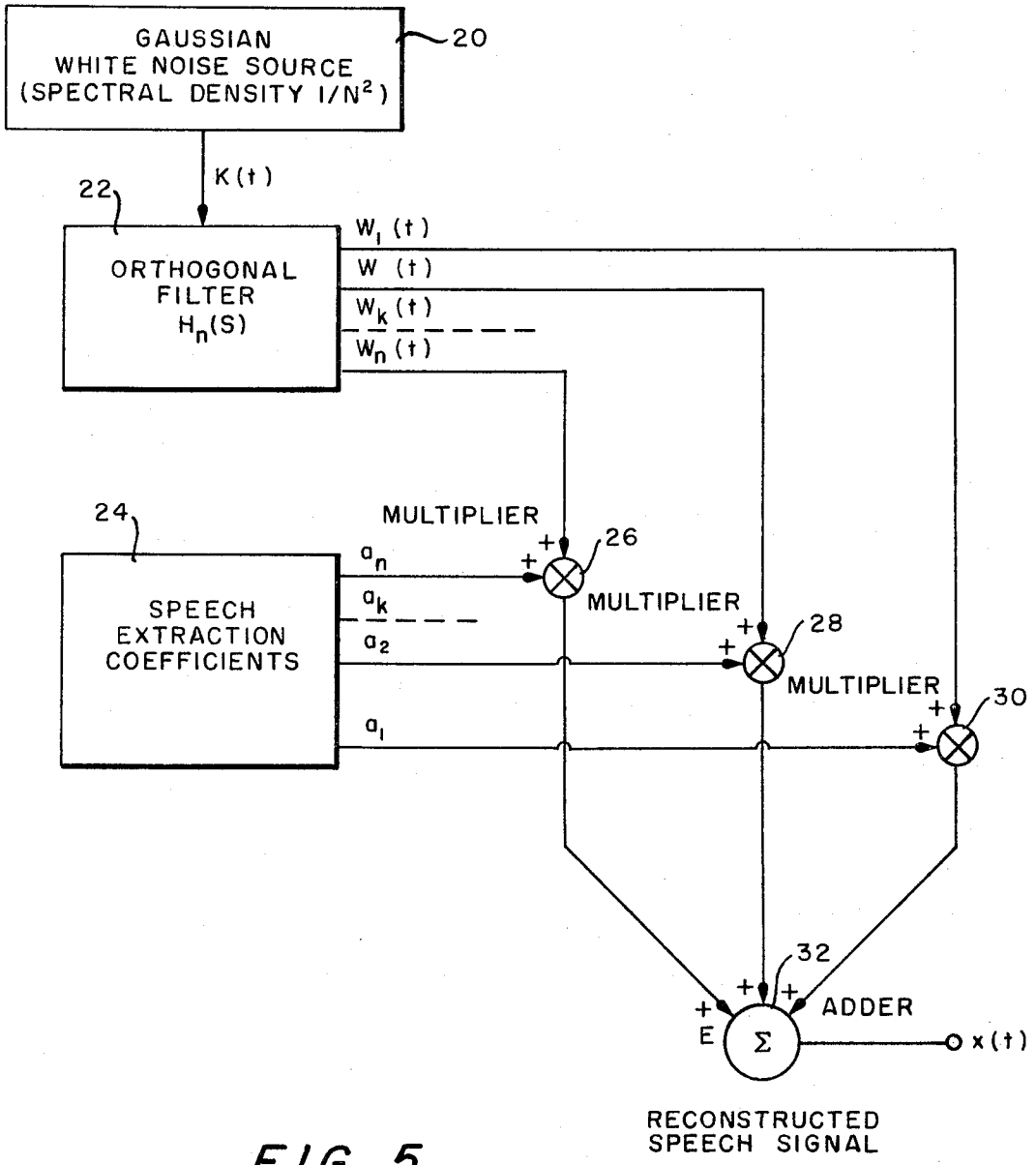


FIG. 5.

INVENTOR.
ALFRED A. WOLF

BY

ATTORNEY

METHOD OF DESIGNING ORTHOGONAL FILTERS

This application is a continuation-in-part of an application filed on June 23, 1971 bearing the Ser. No. 155,988 now U.S. Pat. No. 3,746,791, issued July 17, 1973.

The invention described herein may be manufactured and used by or for the Government of the United States of America for governmental purposes without the payment of any royalties thereon or therefor.

FIELD OF INVENTION

An orthogonal filter is a filter whose output is a set of orthogonal functions related by their coefficients to the input signal to said filter. If the input signal is random white noise, then the filter's output is a set of random orthogonal functions. If the coefficients for this set of random functions are properly selected, any voice signal may be reproduced by temporally averaging the resulting set of orthogonal functions.

With the advent of systems in which there is a man/-machine interface such as computer systems or control systems, it is desirable for a man to be able to communicate with the machine as easily as possible. The ideal situation would be to have the man speak to the machine and to have the machine reply in human language. Several attempts have been made to produce such a system; however, all such schemes to date are generally very complicated and have a limited vocabulary.

DESCRIPTION OF THE PRIOR ART

In the prior art there is no known general method of designing an orthogonal filter. The U.S. Pat. to Norbert Wiener, Nos. 2,024,900 and 2,128,257 describe the design of a specific orthogonal filter which was empirically developed for his situation. This orthogonal filter was the only known orthogonal filter from the date of the Wiener patents to the present time. The present method was developed while designing an orthogonal filter for use in a speech synthesizing system and the method is most clearly described in combination with the speech system.

One of the most popular man/machine interface schemes in which the man merely talks to a machine employs the fact that in each language there are a few basic sounds that make up the words in that language. These basic sounds are called phonemes. In more precise language, phonemes are the sound features which are common to all speakers of a given speech form and which are exactly reproduced in repetition. In any language there are a definite and small number of phonemes. In the English language there are 46 phonemes. These phonemes are known soundwise, and any system employing phonemes records the basic phoneme sounds on magnetic tape or some other recording means. A computer program is then written to connect the proper phonemes to produce words that convey speech information in the form of recurrent patterns. In such a system one can type information into, say a computer, and have it speak back to the operator.

To date such systems have the disadvantage that, as the vocabulary of the system increases the computer programs become more complicated. Thus the speech vocabulary of such systems is usually very limited. Another disadvantage of the system is that the words spoken by the computer are generally of poor fidelity and difficult to understand. Also, such a system is not flexible since the information that can be conveyed by the computer depends on the extent of the vocabulary of the computer, that is, the complexity of the computer program.

Another type of speech synthesizer system is known as the Vocoder (Voice Coder). The Vocoder dates back to the 1920's. The standard Vocoder is a spectrum/channel vocoder. It consists of an analyzer which produces a signal proportional to the short term amplitude spectrum of the fundamental frequency of the speech input, and the synthesizer consists of devices that reconstruct speech by means of electrical signals appearing at the analyzer output. In both the analyzer and the synthesizer, signals are generated that are proportional to both the voiced and unvoiced sounds and the pitch of the sounds. There are, of course, other speech analyzer/synthesizer systems which will not be dealt with here.

Suffice it to say that in each of these systems the method of speech used by the human is imitated in one way or another or the linguistic properties of speech are capitalized upon.

It is an object of the present disclosure to provide a general method of designing orthogonal filters.

It is another object of the present disclosure to provide an orthogonal filter for use in a speech synthesizing system.

These and other objects of the present invention are set forth in the following disclosure.

SUMMARY OF THE INVENTION

The present speech system capitalizes on the fact that speech is a stochastic process. Speech is stochastic because long samples of speech convey information which is probabilistic in time. The present system employs a gaussian white noise source; the output of which is passed through an orthogonal filter to produce a set of random orthogonal functions which when multiplied by a speech signal and averaged produces coefficients which can be used to synthesize speech information at a later time. This method is somewhat analogous to the use of generalized Fourier coefficients to define and reproduce a periodic function.

DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of the system for obtaining the speech extraction coefficients.

FIG. 2 is a block diagram of a linear orthogonal filter as a cascaded chain of linear filter sections.

FIG. 3 is a schematic drawing of an RC orthogonal filter.

FIG. 4 is a schematic drawing of the temperal averager.

FIG. 5 is a block diagram of the speech reconstruction system.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention relates to a general method for designing orthogonal filters. In particular an orthogonal filter design is disclosed which may be used in a speech synthesizing and codifying system.

It should be noted that the following mathematical terminology is standard in the art and can be found in the textbook published by the Addison-Wesley Publish-

ing Company, entitled: "Principles of Feedback Control," written by Charles H. Wilts, as well as other textbooks.

This invention makes it possible to transform a sample of speech of length, T_o into a set of speech coefficients designated by $a_1, a_2, \ldots a_n$, each of which depends on time according to the sample length, T_o of the speech sample, which is explained further below. These coefficients $a_1, a_2, \ldots, n$ can be thought of as a set of generalized Fourier coefficients defined on a set of orthogonal noise sample functions obtained from a white noise source. If the waveform of the speech sample, of length T_o, is denoted by $x(t)$, then by making use of the fact that $x(t)$ is a sample function from a stochastic process, it can be decomposed into an infinite series of orthogonal sample functions taken from a white noise source. Each orthogonal sample function is weighted by a coefficient the value of which depends on the speech sample under consideration. The set of coefficients $\{a_n: n = 1, 2, \ldots\}$ thus contains the necessary information from which the original speech sample can be reconstructed by weighting the orthogonal sample functions $\{w_n(t): n = 1, 2, \ldots\}$ (see FIG. 5) with the appropriate corresponding coefficient, $\{a_n: n = 1, 2, \ldots\}$ in which T_o is a strip of time running to infinity and t is a given instant of time. The invention makes possible the maximum compression of speech by using the coefficients $\{a: n = 1, 2, \ldots\}$ to convey information. Using these coefficients, it is now possible for man to communicate with computers and to have them in turn communicate with man by means of speech. This invention opens up the possibility of new, unforeseen innovations in machines and systems in which there is a speech communication interface with man. One such example, in addition to the possibility of talking to a computer is the possibility of talking to a typewriter.

DESCRIPTION OF THE COEFFICIENT OR CODING GENERATOR

FIG. 1 is a block diagram of the system for obtaining the speech extraction coefficients: The speech signal **2**, denoted by the sample functions $\{x(t)\}$ of sample length T_o is multiplied (instant by instant) by each orthogonal noise signal denoted respectively by the sample functions $v_1(t), v_2(t), v_3(t), \ldots, v_n(t)$, which are derived as the set of outputs of an orthogonal filter **12**, described below when the white noise signal **10**, denoted by the sample function $\{g(t)\}$, of spectral density N^2 watts per cycle, is applied to the orthogonal filter's input. The sample functions that result from the multiplication of the set of orthogonal noise sample functions $v_1(t), \ldots, v_n(t)$ with the sample function $x(t)$ in the multipliers **4, 6,** and **8,** are the speech extraction samples of length T_o seconds as denoted by $\{v_n(t)\,x(t): n = 1, 2, \ldots\}$. This set of product sample functions $v_1(t)\,x(t), v_2(t)\,x(t), v_3(t)\,x(t), \ldots, v_n(t)\,x(t)$ is averaged temporally in the temporal averages 14, 16, and 18 give rise to the corresponding set of cofficients. $a_1, a_2, a_3, \ldots, a_n$, where mathematically these averages are formally given by the equations:

$$a_1 = \frac{1}{T}\int_0^T v_1(t)x(t)dt \qquad (1)$$

$$a_2 = \frac{1}{T}\int_0^T v_2(t)x(t)dt \qquad (2)$$

$$a_3 = \frac{1}{T}\int_0^T v_3(t)x(t)dt \qquad (3)$$

and so forth to

$$a_n = \frac{1}{T}\int_0^T v_n(t)x(t)dt \qquad (4)$$

where T is the averaging time and is at least equal to the sample length, T_o of the speech sample. The integral over (o,T) of each of the product sample functions, $v_n x$ divided by T is the temporal average of those sample functions. The resulting speech extraction coefficients $a_1, a_2, \ldots, a_n$ are dependent only on the sample length, T_o and are the coefficients that represent the essential information needed to reconstruct the speech signal, $x(t)$. A rule of thumb for the value of T, the averaging time is that about ten times the sample length, T_o of the speech sample.

GENERATION OF THE ORTHOGONAL NOISE FUNCTIONS

$$v_1(t), v_2(t), \ldots, v_n(t)$$

In the generation of the speech code, $a_1, a_2, a_3, \ldots, a_n$, it is necessary to generate a corresponding set of random sample functions $v_1(t), v_2(t), v_3(t), \ldots, v_n(t)$, that are pairwise orthogonal. By pairwise orthogonality we roughly mean that the information contained in each of the sample functions $v_1(t), v_2(t), \ldots, v_n(t)$ is unique. To put this another way, no overlap in information exists between any two sample functions $v_1(t), v_2(t), \ldots, v_n(t)$. For descriptive purposes, the generation of the orthogonal random sample functions is achieved by means of an orthogonal filter when white noise of power spectral density N^2 is applied to the input of the orthogonal filter.

The orthogonal filter is a linear filter with one pair of input terminals and many pairs of output terminals. The filter, which is described below in detail, may be roughly likened to an ideal prism on which white light is incident. The incident light may be thought of as the input to the prism. The output of an ideal prism is essentially the complementary colors in response to the incident white light upon it. The complementary colors are of course pairwise orthogonal. In our case, however, the orthogonal functions are random within a band of frequencies whereas in the prism case the complementary functions are in principle roughly single frequency sinusoids with random amplitudes and random phases. The latter is due to the randomness of the white light.

CONSTRUCTION OF THE ORTHOGONAL FILTERS

Linear orthogonal filters can be constructed as a chain of linear filter sections in which the poles of any section of the filter, in the complex plane, is essentially cancelled by the zeros of the next immediate section. FIG. 2 shows this general scheme for constructing an orthogonal filter. The complex function $H_1(s)$ is the transfer function of the first section of the filter as a function of s, the complex frequency variable in which

$$s = \alpha + j\omega \qquad (5)$$

where

$\omega = 2\pi f$,

f = frequency in cycles per second

α = decrement in amplitude in cycles per second

$j = \sqrt{-1}$, unit of the imaginary number system.

The complex function $H_2^0(s)$ is the transfer function from the input of the second section of the orthogonal filter to its output. Similarly, $H_n^0(s)$ is the transfer function from the input of the n-th section of the filter to the n-th output terminal pair of the orthogonal filter. The process for realizing the actual construction of an orthogonal filter is given by the following formula:

Let $H_1(s)$ be the transfer function of the first filter section. Then

$$H_1(s) = K_1(B(s)/C_1(s)) \tag{6}$$

The numerator $B(s)$ is selected to satisfy the equation

$$B(s) = b_0 + b_1 s + \ldots + b_k s^k \tag{7}$$

and the denominator $C(s)$ is selected to satisfy the equation

$$C_1(s) = c_{10} + c_{11} s + \ldots + c_{1m} s^m \tag{8}$$

where the constants $b_0, b_1, \ldots, b_k$ and the constants $c_{10}, c_{11}, \ldots, c_{1m}$ are chosen such that $H(s)$ is physically realizable. In a like way the transfer function

$$H_2^0(s) = (K_2/K_1)\, C_1(-s)/C_2(s) \tag{9}$$

is designed such that the numerator $C_1(-s)$ is exactly like the denominator of the first section, $H_1(s)$ with s replaced by $(-s)$. This makes the zeros of $H_2(s) = H_2^0(s)H_1(s)$ occupy the same positions as the poles of $H_1(s)$ except that they are reflected about the $j\omega$-axis. Hence, the zeros of one transfer functio in effect cancel the poles of the next section giving rise to the orthogonality of the pair. In general, the n-th transfer function of the orthogonal filter from the input of the filter is given by the formula

$$H_n(s) = K_n B(s) \frac{\prod_{k=1}^{n-1} C_k(-s)}{\prod_{k=1}^{n} C_k(s)} \tag{10}$$

for n ≤ 2. Where the symbol, $\prod_{k=1}^{n}$ is the product of all

is the product of all factors in the expansion of Equation 10 from 1 to n. For this speech device, a filter is developed using this process of construction of orthogonal filters having simple poles and zeros along the axis of the reals. This gives rise to simple resistance-capacitance (RC) network with differential amplifiers. The RC orthogonal filter is shown in FIG. **3.** The first section of the filter includes an input resistor **34,** feedback resistor **36** designated as R in Equation 13, feedback capacitor **38,** designated as C in Equation 13, and operational amplifier **40.** The subsequent section of the filter includes inputs resistors **42, 44,** designated as r_{1n} and r_{2n} in Equation 15 respectively, resistors **48, 50, 54,** and **56;** capacitors **46, 52;** and differential amplifier **58.** Resistors **48** and **50** are designated as R_n in Equation 17, resistors **54** and **56** are designated as R_{1n} and R_{fn} respectively in Equation 16 and capacitors **46** and **52** are designated as C_n in Equation 17. For these filters the transfer function of the first section is given by

$$H(s) = (\sqrt{2s_1}/s+s_1) \tag{11}$$

and for $n \geq 2$

$$H_n(s) = (s_n/s_n)\,(s - s_{n-1}/s+s_n)\, H_{n-1(s)} \tag{12}$$

where for design purposes the first pole is placed at s_1 along the real axis where

$$s_1 = 1/RC \tag{13}$$

and

$$\sqrt{2s_1} = 1/rC_1 \tag{14}$$

for the first section and for the subsequent tandem sections, the poles $s_2, s_3, \ldots$ are placed according to the formulae

$$(s_{n-1}/s_n) = (r_{2n}/r_{1n} + r_{2n}) \tag{15}$$

$$(s_n/s_{n-1}) = (R_{fn}/R_{1n}) \tag{16}$$

and

$$s_n = 1/R_n\, C_n \tag{17}$$

In the ordinary orthogonal filter, the poles $s_1, s_2, s_3, \ldots$ can be placed arbitrarily along the real axis. But for the speech synthesizer this procedure will fail to work. It is important to place the poles $s_1, s_2, s_3, \ldots$ in such way as to make the reconstructed signal converge to the desired speech signal. In the process of constructing an orthogonal filter for the speech synthesizer the first pole is selected approximately equal to the bandwidth of the speech signal. Since speech signals can cover a bandwidth of the order of 200 Hz to 2,500 Hz, depending on the speaker, s_1 can be selected to be a frequency within this range. To make the reconstructed signal converge to the original speech signal in the mean squared sense, the remaining poles are chosen according to the formula

$$s_n = s_1/n^2;\ n \geq 1 \tag{18}$$

If the remaining poles are placed according to the formula

$$s_n = s_1/n^3;\ n \geq 1 \tag{19}$$

then the reconstructed signal will converge absolutely almost everywhere. In general

$$s_n = s_1/n^p;\ n \geq 1 \tag{20}$$

where $p \geq 2$. The higher p, the faster the convergence and the less the number of terms in the series that will be needed for convergence. There is a practical difficulty however, which causes a resolution problem in placing the poles at distinct positions along the real axis.

SUMMARY OF THE GENERAL PROCESS FOR PRODUCING ORTHOGONAL FILTERS

1. Select the nature of the poles of the orthogonal filter to be, i.e., whether they are simple or complex, from the application for which the filters are to be used.

2. From Step 1 above, the degree of $B(s)$ in Eq. (7) is determined and hence the degree of $C_1(s)$ is determined from physical realizability considerations, i.e., if k is the degree of $B(s)$ and m is the degree of $C_1(s)$ then $m = k+1$.
3. Check to see if

$$\int_{-\infty}^{\infty} \frac{\log H_n(\omega)\, d\omega}{1+\omega^2} < \infty \qquad (21)$$

is satisfied where: $H_n(\omega) = |H_n(j\omega)|$ is the amplitude characteristic of the transfer function $H_n(s)$.
4. From each transfer function, i.e., $H_1(s)$ and $H_2(s)$, etc. given by

$$H_1(s) = K_1(B(s)/C_1(s)) \qquad (22)$$

and

$$H_2(s) = H_2(s)\ H_1(s) \qquad (23)$$

where

$$H_2(s) = (K_2/K_1)\ (C_1(-s)/C_2(s)) \qquad (24)$$

etc. the circuit is synthesized according to standard procedure noting that the placement of the poles must follow the requirements of convergence if for example a speech signal is to be reconstructed.
5. Once a design of $H_1(s)$ and $H_2^0(s)$ are made, the orthogonal filter is made by cascading $H_1(s)$ and $H_2^0(s)$ together and then adding to this cascade as many sections as needed according to the requirements of the application, in which subsequent sections are identical to $H_2^0(s)$ in terms of the circuit used.

The individual sections following the second section are designed according to the formula

$$H_n(s) = (K_n/K_{n-1})\ (C_{n-1}(s)/C_n\ (s)) \qquad (25)$$

It is therefore clear that in terms of the circuit all the $H_2^0(s)$ for $n \geq 2$ are the same with the exception of the circuit parameters which depend only on n, the number of the section. Hence, the additional requirements on the poles depend on the convergence of the reconstruction process or on some other physical requirement. The resulting chain is now an orthogonal filter.
6. The design of the transfer functions $H_1(s)$, $H_2^0(s)$, $H_3^0(s)$, . . . depend on the placement of the poles $s_1, s_2, s_3, \ldots$. These in turn define the constants b_0, $b_1, b_2, \ldots$, and $C_{n0}, C_{n1}, C_{n2}, \ldots$, where n is the number of the section. For an orthogonal filter used in reconstruction of signals, the poles s_n, $n = 1, 2, \ldots$ must be placed according to the rules:
   a. s_1 selected equal to or greater than the bandwidth of the signal under consideration.
   b. For convergence in the mean-square case

$$S_n = S_1/n^2$$

   c. For convergence almost everywhere

$$S_n = S_1/n^3$$

   d. For other types of convergence

$$S_n = S_1/n^p;\ p \geq 4$$

7. The kind of convergence or the method of pole placement depends on the application intended and once that is decided, the resolution of pole placement is determined and the resulting error in convergence can be estimated.
8. Once the transfer function is designed according to the process given in steps 1 through 7 above, the hardware is obtained by standard analog computer techniques or by factoring the transfer functions and then determining the equivalent partial fraction expansion using amplifiers and differential amplifiers to take account of the negative signs that may result from the partial fraction constants. The process described above is the one used to obtain the hardware configuration of the orthogonal filter shown in FIG. **3.**

SIGNAL MULTIPLIERS

The signal multipliers shown in FIG. **1** for multiplying the orthogonal noise components of the white noise source with the speech signal are standard devices. In the frequency range of interest, i.e., bandwidths up to 20 KHz, the quarter Gauss square method was used. Other types can easily be used instead.

The quarter Gauss square multiplier is one that makes use of the identity

$$(A+B)^2 - (A-B)^2 = 4\ AB \qquad (26)$$

TEMPORAL AVERAGER

The temporal averager is a simple RC low pass filter with a very long time constant compared to the highest frequency component in the speech signal. If f_0 denotes this frequency, then

$$RC >> (1/f_0)$$

FIG. **4** gives an example of the configuration of the temporal averager. The temporal averager consists of input resistor **60,** feedback capacitor **62** and amplifier **64.**

GAUSSIAN WHITE NOISE SOURCE

The Gaussian white noise source is a standard noise generator of the diode type.

SYSTEM FOR THE RECONSTRUCTION OF THE SPEECH SIGNAL FROM SPEECH EXTRACTION COEFFICIENTS AND WHITE NOISE

Once the speech extraction coefficients, a_1, a_2, a_3, a_n are obtained from the Speech Extraction Code Generator, shown in FIG. **1,** it is possible to reconstruct from these coefficients the original speech sample, $x(t)$. The reconstruction of $x(t)$ is carried out by multiplying each coefficient $\{a_k : k = 1, 2, \ldots, n\}$ by a corresponding set of orthogonal noise functions which we shall denote by the symbols $\{w_k(t) : K = 1, 2, \ldots, n$ and $0 \leq t < \infty\}$ and then summing the resulting set of products. Thus, we form the set of products $a_1 w_1(t)$, $a_2 w_2(t)$, $a_3 w_3(t)$, . . ., $a_n w_n(t)$ and then sum to obtain the reconstructed speech signal $x(t)$ as the summation

$$s(t) \approx \sum_{k=1}^{n} a_k w_k(t) \qquad (27)$$

It is evident by analogy that the coefficients $a_1, a_2, \ldots, a_n$ correspond to the coefficients of a Fourier series and the orthogonal noise functions $w_1(t), w_2(t), w_3(t), \ldots, w_n(t)$ correspond to the orthogonal set of trigonometric functions. Equation (27) is the partial sum of a

generalized random orthogonal series which converges as $n \to \infty$ to the speech sample $x(t)$ in some probabilistic sense. If the series converges rapidly then by truncating the series at some finite number n the partial sum will still approximate the speech sample with a predictable error. The mean squared error,

$$\epsilon_n^2$$

can be calculated from the equation

$$\overline{\epsilon_n^2} = \overline{x^2} - \sum_{k=1}^{n} a_k^2 \qquad (28)$$

For a speech sample $\overline{x^2}$ will be fixed. Then $\overline{\epsilon^2}_n$ is smallest when

$$\sum_{k=1}^{n} a_k^2$$

is largest.

THE PHYSICAL RECONSTRUCTION SYSTEM

FIG. 5 is a block diagram of the speech reconstruction system. Using the coding coefficients $a_1, a_2, \ldots, a_n$ and a Gaussian white noise source of spectral power density of $1/N^2$ watts per cycle the speech sample function, $x(t)$ is reconstructed according to the description given above.

In FIG. 5, the white noise signal, $k(t)$ of spectral density of $1/N^2$ watts per cycle is applied to an orthogonal filter which has a transfer function, $H_n(s)$ identical to the orthogonal filter of the coding generator shown in FIG. 1. The response of this orthogonal filter is the set of orthogonal noise sample functions $w_1(t)$, $w_2(t)$, $w_3(t)$, . . ., $w_n(t)$. Each of these is multiplied respectively by the coefficients $a_1, a_2, \ldots, a_n$ derived from the coding generator (FIG. 1). The products thus formed in each of the multipliers is summed in the adder to give

$$x(t) \approx \sum_{k=1}^{n} aw(t) \qquad (29)$$

It will be noted that the elements of the reconstruction system are quite similar to the coefficient code generator system. The orthogonal filter is the exact same design as in the code generator case. The white noise source of the reconstruction system differs from the coding generator system in that the spectral density of one is the reciprocal of the other.

The multiplier is a signal multiplier the same as described previously above. The speech extraction coefficients can be stored in a memory such as a tape as voltages of appropriate value.

THE ADDER

The summing of the products $a_1w_1(t)$, $a_2w_2(t)$, . . ., $a_nw_n(t)$ are accomplished in a conventional adder.

RELATIONS BETWEEN SPECTRAL DENSITIES

Since the white noise source of the generation and reconstruction systems have reciprocal spectral density functions, the sample functions are related according to

$$g(t) = N^2k(t) \qquad (30)$$

SUMMARY

In the method presented here the speech analysis and synthesis technique capitalizes on the fact that speech is a random signal that can be decomposed into a generalized orthogonal series something like the Fourier series. This means that a speech signal can be represented by a set of coefficients which depend only on the nature of the speech information and on the length of the speech sample. This speech synthesizing system is an electronic system for accomplishing the generation of these speech coefficients or speech extraction parameters and for utilizing them to reconstruct the speech into a spoken signal.

The novelty of the system consists in the utilization of orthogonal filters, white noise and temporal averaging devices connected in the unique arrangement shown in FIG. 1 that gives rise to the speech extraction coefficients, $a_1, a_2, \ldots, a_n$. It should be noted that the speech extraction parameters are very narrow band signals. These signals are nearly constant for $T >> T_o$. This also means that the system can be used to greatly compress speech.

Obviously many modifications and variations of the present invention are possible in the light of the above teachings. It is therefore to be understood that withing the scope of the appended claims the invention may be practiced otherwise than as specifically described.

What is claimed is:

1. A method for making an orthogonal filter having sections 1 through n comprising the steps of:

selecting a first transfer function as

$$H_1(s) = K_1(B(s)/C_1(s))$$

wherein the numerator $B(s)$ is selected to satisfy the relationship:

$$B(s) = b_o + b_1s + \ldots b_ks^k$$

and the denominator C_1 is selected to satisfy the following relationship:

$$C_1(s) = c_{10} + c_{11}s + \ldots + c_{1m}s^m$$

wherein the constants $b_o, b_1, \ldots, b_k$ and the constants $c_{10}, c_{11}, \ldots, c_{1m}$ are chosen such that $H_1(s)$ is capable of being manufactured from realizable components;

constructing a first filter section having as its transfer function said first transfer function;

placing said first filter section first serially in a group of n filter sections;

selecting a second transfer function as

$$H_2{}^o(s) = (K_2/K_1)\ (C_1(-s)/C_2(s))$$

electing the numerator of the second equation $C_1(-s)$ and the denominator of the first equation $C_1(s)$ of the transfer function so that the zeros of the second transfer function occupy the positions of the poles of the first transfer function and are reflected about the $j\omega$-axis, and with the effect that poles of the first transfer function cancel the zeros of the second transfer function thereby producing orthogonality;

constructing a second filter section having as its transfer function said second transfer function;

placing said second filter section second serially after said first filter section in a group of n filter sections;

arranging each successive section following section 2 such that the zeros of their respective trnsfer functions occupy the positions of the poles of the preceeding transfer functions reflected about the $j\omega$ axis and with the zeros of the successive transfer function cancelling the poles of the preceeding

transfer function thereby giving rise to orthogonality; and,

arranging the n-th successive filter section of the orthogonal filter such that the transfer function of each filter in a series including n filters is described by the relationship:

$$H_n(s) = K_nB(s) \frac{\prod_{k=1}^{n-1} C_k(s)}{\prod_{K=1}^{n} C_k(s)}$$

wherein k signifies all the filters in the series from the first filter to the n-th filter, inclusively.

2. A method as defined in claim **1** wherein the orthogonal filter is to be used in a speech synthesizer including the steps of:

selecting s_1 to be within the frequency range of 200 Hertz to 2,500 Hertz, and

selecting the remaining poles in accordance with the formula

$$s_n = (s_1/n^3) \text{ for } n \geq 1.$$

3. An orthogonal filter having sections 1 through n comprising:

a first section characterized by the transfer function

$$H_1(s) = K_1(B(s)/C_1(s))$$

wherein the numerator $B(s)$ is selected to satisfy the relationship:

$$B(s) = b_o + b_1s + \ldots + b_ks^k$$

and the denominator C_1 is selected to satisfy the following relationship:

$$C_1(s) = c_{10} + c_{11}s + \ldots + c_{1m}s^m$$

wherein the constants $b_o, b_1, \ldots b_k$ and the constant $c_{10}, c_{11}, \ldots c_{1m}$ are chosen such that $H_1(s)$ is capable of being manufactured from realizable components;

a second section characterized by the transfer function

$$H_2{}^o(s) = (K_2/K_1)\ (C_1\ (-s)/C_2\ (s))$$

wherein the numerator of the second equation $C_1(-s)$ and the denominator of the first equation $C_1(s)$ of the transfer function are selected such that the zeros of the second transfer function occupy the positions of the poles of the first transfer function and are reflected about the $j\omega$-axis, and with the effect that poles of the first transfer function cancel the zeros of the second transfer function thereby producing orthogonality;

arranging each successive filter section following section 2 such that the zeros their respective transfer function occupy the positions of the poles of the preceeding transfer function reflected about the $j\omega$ axis and with the zeros of the transfer function of the successive filter section cancelling the poles of the transfer function of the preceeding section thereby giving rise to orthogonality; and

an n-th section characterized by the transfer function

$$H_n(s) = K_nB(s) \frac{\prod_{k=1}^{n-1} C_k(-s)}{\prod_{K-1}^{n} C_k(s)}$$

wherein k denotes all the filter sections in the series from the first filter to the n-th filter inclusive.

4. An orthogonal filter as defined in claim **3** for use in a speech synthesizing system which further includes:

selecting s_1 from within the range of 200 Hertz to 2,500 Hertz, and

selecting the remaining poles in accordance with the formula

$$s_n = (s_1/n^3) \text{ for } n \geq 1.$$

* * * * *