Synthesis of 4–2–4 Matrix Recording Systems

JAMES V. WHITE

CBS Technology Center, Stamford, CT 06905

Linear time-invariant 4–2–4 matrix recording systems are defined mathematically. Mono and stereo compatibility objectives are stated; the class of compatible encoders is then synthesized that meets these objectives. To each of these encoders there corresponds a decoder that yields the smallest possible mean-square error in its quadraphonic outputs. An explicit formula for this minimum-error decoder is derived, which shows that the minimum-error decoder is time varying (uses "logic" to track the correlations of the quadraphonic signals). The optimal time-invariant (no logic) decoder, which minimizes the mean-square error for worst case inputs, is then synthesized, and several of its most important properties are discussed. The conditions for optimum time-invariant decoding are then combined with the conditions for compatible encoding; this yields the class of optimized compatible recording systems, which is termed the SQ family. Two well-known members of this family are basic SQ and forward-oriented SQ.

INTRODUCTION: For the purpose of this paper a 4–2–4 matrix recording system is defined as a linear time-invariant system with four input signals (LF, RF, RB, LB) and three sets of outputs, termed quadraphonic, stereo, and mono. There are four quadraphonic output signals (LF′, RF′, RB′, LB′), two stereo output signals (LT, RT), and one mono signal (LT+RT). The system is shown in Fig. 1. For convenience, we represent the set of inputs by the matrix x:

$$x \triangleq \begin{bmatrix} \mathrm{LF} \\ \mathrm{RF} \\ \mathrm{RB} \\ \mathrm{LB} \end{bmatrix} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \qquad (1)$$

where $\triangleq$ means "equal be definition".

In similar manner, we also define the output matrices x' and y as shown in Fig. 1. The elements of these matrices are complex functions of time, called analytic signals [1]. For harmonic signals these functions are simply phasors with exp $(j\omega t)$ factors included. More information about analytic signals may be found in the Appendix.

The encoder is a linear time-invariant system with the transfer-function matrix E, which is a 2×4 array of complex numbers. The stereo output is therefore given as

$$y = Ex. \qquad (2)$$

Note that y is the matrix of recorded signals.

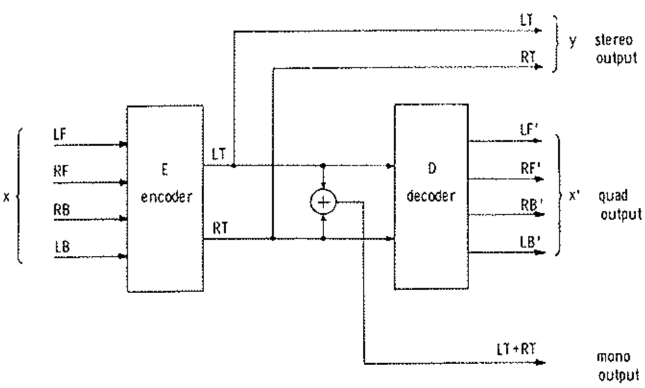

Fig. 1. Structure of a 4–2–4 matrix recording system.

The decoder is also a linear time-invariant system. Its transfer-function matrix D is a 4×2 array of complex numbers, and the quadraphonic output x' is given as

$$x' = Dy. \tag{3}$$

The problem of mathematical synthesis of 4–2–4 matrix systems may be stated as follows: Given a class of important input signals and a set of performance objectives for the resulting outputs, determine analytically the set of all pairs of encoders and decoders (E, D) that meet the objectives.

REVIEW OF THE LITERATURE

Cooper and Shiga [2] were the first to describe a mathematical procedure for synthesizing matrix recording systems to meet specified performance objectives. Their formulation of the synthesis problem differs from ours in the following ways. They place no restrictions on the number of input signals, recorded signals, or output signals, whereas our synthesis is specialized to 4–2–4 systems. They place no mono or stereo compatibility requirements on their synthesis (although they consider compatibility when they interpret their results) whereas we prescribe four compatibility requirements and use these as the starting point in our synthesis. They stipulate that their recording systems should be azimuthally nonoriented, whereas the front sound stage is given a special status in our formulation, which yields systems with azimuthal orientation. They synthesize their systems so that the mean-square error in the output signals is minimized, provided the input signals to the encoder meet certain conditions regarding spatial harmonic content. In contrast, we synthesize systems so that the mean-square error in the quadraphonic output signals is minimized, provided compatibility requirements are met by the encoder, and the input signals meet certain correlation conditions.

Gerzon's paper [3] on the synthesis of periphonic systems generalizes the approach of Cooper and Shiga to signals distributed on the sphere. Many other papers have been published that analyze the mathematical and acoustical properties of matrix systems, but none of these papers describe true synthesis procedures, in which specific performance objectives are prescribed, and then mathematical analysis is used to deduce the structures of all systems that meet these objectives.

PLAN

In this paper we use the following plan to synthesize a class of time-invariant matrices that meet specific stereo and mono compatibility requirements. We start by defining a large class of important input signals for commercial recording. We then impose performance objectives on the stereo and mono outputs produced by these inputs. The class of all encoders that meet these objectives is identified; this is the class of compatible encoders. Next, a quadraphonic performance objective is imposed on the system: the decoder is to be chosen to maximize the mean-square accuracy of the quadraphonic outputs. The decoder that does this is shown to use "logic," which means that the minimum-error decoder utilizes information about the correlations in the signals to modify its structure and thereby minimize decoding errors. (Logic decoders are time-varying systems that track the changing correlations in the signals.) We then impose the condition that the decoder be time invariant (no logic) and derive the equations for optimal time-invariant decoders.

OVERVIEW OF ASSUMPTIONS AND RESULTS

Every synthesis is based on a set of prescribed performance objectives. These prescriptions should accurately describe as many of the important engineering objectives as possible. This means that secondary details should be omitted, lest the mathematics become intractible. It also means that questionable prescriptions that have no clear physical significance should also be omitted, lest the conclusions reached be inconclusive to experienced engineers. In other words, the prescriptions should be clean and lean.

With this in mind, we prescribe a set of four performance objectives that we believe a compatible matrix system should satisfy. These are considered important necessary performance objectives that meet the "clean-lean" criterion. A precise statement of these prescriptions requires mathematics, which is reserved for a later section in this paper. For the purposes of giving a general overview, the gist of these assumptions is described below.

One mono capability objective is prescribed for the encoder.

1) The mono output power should equal the total four-channel input power for arbitrary uncorrelated inputs.

Three stereo compatibility objectives are prescribed for the encoder.

1) The front sound stage of the four-channel input should be reproduced in stereo without errors of any kind.

2) The stereo output power should equal the total four-channel input power for arbitrary uncorrelated inputs.

3) The total stereo output power should be divided equally between the left and right channels when the four input signals are uncorrelated and of equal power.

All encoders that satisfy these performance objectives are identified in this paper; they form the class of compatible encoders. For these encoders there are optimum decoders, and the final step in our synthesis is to identify these. For this purpose, we must define what we mean by "optimum," i.e., we must prescribe a set of quadraphonic performance objectives for the decoders. In keeping with our desire to keep things "clean and lean" the following two prescriptions are used to define optimal time-invariant decoders.

1) The decoder should be as accurate as possible for worst-case inputs in the sense that the decoding errors should contain the smallest possible total power. (Worst-case inputs are assumed to consist of four uncorrelated signals of equal power.)

2) The decoder should be robust in the sense that its performance should be as insensitive as possible to small errors in its gains.

The analysis in this paper proves that all the above performance objectives will be met if and only if the encoder E and its optimum time-invariant decoder D have the following forms:

$$E = \begin{bmatrix} 1 & 0 & A & B \\ 0 & 1 & pc_1A & -pB/c_1 \end{bmatrix} \quad (4)$$

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ A^* & -pc_1A^* \\ B^* & pB^*/c_1 \end{bmatrix} \quad (5)$$

where A and B are any complex numbers such that $|A|^2 + |B|^2 = 1$, $p = \pm j$, $c_1 = |B|/|A|$, and the asterisks denote complex conjugates. The "if and only if" part of this result means that the prescribed objectives for compatibility and optimality are met by any matrices with the forms given in Eqs. (4) and (5), but only by matrices with these forms.

This set of compatible optimum matrices is termed the SQ[1] family. Two members of this family are well known. These are basic SQ [9]–[11] (which corresponds to $p = +j$, $A = 0.707$, and $B = -0.707j$) and forward-oriented SQ [10] (which corresponds to $p = +j$ and the reversed assignment $A = -0.707j$ and $B = 0.707$).

INPUTS

The first step in our synthesis procedure is to describe accurately the kinds of inputs that are mentioned in the performance objectives. We will start with a specific example: the class of uncorrelated unit-power input signals. This class contains signals in which there are equal amounts of power in each of the four input channels, and these four signals are uncorrelated with respect to each other. An example of such an input is

$$x \triangleq \begin{bmatrix} \mathrm{LF} \\ \mathrm{RF} \\ \mathrm{RB} \\ \mathrm{LB} \end{bmatrix} \triangleq \begin{bmatrix} \exp(j\omega_1 t) \\ \exp(j\omega_2 t) \\ \exp(j\omega_3 t) \\ \exp(j\omega_4 t) \end{bmatrix}. \quad (6)$$

Each input channel is driven with unit power at a different frequency ($\omega_i = \omega_j$, $i \neq j$). This is only one example; there are an infinite number of other signals that are also uncorrelated and of unit power. A compact way of describing the entire class of such signals is to use the correlation matrix K, which is defined as

$$K \triangleq \overline{xx^+}. \quad (7)$$

The overbar denotes a time average, which does not need to be defined explicitly for this analysis, and x^+ denotes the adjoint of x, which is the complex conjugate transpose matrix.

The correlation matrix for x in Eq. (6) is therefore the time average of the following matrix product:

$$\begin{bmatrix} \exp(j\omega_1 t) \\ \exp(j\omega_2 t) \\ \exp(j\omega_3 t) \\ \exp(j\omega_4 t) \end{bmatrix} \times$$

$$[\exp(-j\omega_1 t) \quad \exp(-j\omega_2 t) \quad \exp(-j\omega_3 t) \quad \exp(-j\omega_4 t)].$$

The element in the ith row and kth column of this time average is

$$K_{ik} = \overline{\exp(j\omega_i t)\exp(-j\omega_k t)} = \begin{cases} 1, & i = k \\ 0, & i \neq k. \end{cases} \quad (8)$$

Thus

$$K = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \triangleq I_4. \quad (9)$$

This states that K is the 4×4 identity matrix I_4.

Correlation matrices have a physical interpretation[2] that will now be described. K is a 4×4 array of complex numbers; the element of K in the ith row and jth column is $K_{ij} = \overline{x_i x_j^*}$, where x_i is the ith element of x, and the asterisk denotes the complex conjugate. The diagonal element K_{ii} equals the average power in x_i, while the off-diagonal element K_{ij} equals the complex correlation between x_i and x_j. The real part, $\mathrm{Re}(K_{ij})$, is the real correlation; this number is equal to the time average of the product of the physical signals in channels i and j. The imaginary part, $\mathrm{Im}(K_{ij})$, is the quadrature correlation; this number is equal to the time average of the product of the physical signals in channels i and j after the ith signal has been lagged 90° in phase with respect to the jth signal.

With this compact notation established, we now give precise definitions of the classes of inputs needed for our synthesis procedure. The class of *uncorrelated unit-power* input signals is defined as the class for which K is the 4×4 identity matrix. Another important class of inputs is the class of *uncorrelated inputs,* with no restriction placed on the power in each x_i. This is the class for which K is any diagonal matrix with nonnegative elements. The last class of signals we need is called the class of inputs with *zero power in the back*. This is the class for which $x_3 \equiv x_4 \equiv 0$.

PERFORMANCE OBJECTIVES FOR STEREO COMPATIBILITY

The next step in the synthesis is to state the prescribed stereo performance objectives in mathematical form.

1) *Accurate Reproduction of Front Sounds*. For the class of inputs with zero power in the back, the stereo outputs must correctly reproduce the left and right front inputs, i.e., LT ≡ LF and RT ≡ RF. In terms of matrix elements this objective states that $y_1 \equiv x_1$ and $y_2 \equiv x_2$ for any input x such that $x_3 \equiv x_4 \equiv 0$.

2) *Conservation of Power*. For the class of uncorrelated inputs, the stereo output power must equal the input power, i.e., $\overline{|\mathrm{LT}|^2} + \overline{|\mathrm{RT}|^2} \equiv \overline{|\mathrm{LF}|^2} + \overline{|\mathrm{RF}|^2} + \overline{|\mathrm{RB}|^2} + \overline{|\mathrm{LB}|^2}$. This may be written more compactly as $\overline{y^+y} \equiv \overline{x^+x}$ for any input x such that K is diagonal.

3) *Left–Right Symmetry*. For the class of uncorrelated unit-power inputs, the left-channel and right-channel powers must be equal in the stereo output, i.e., $\overline{|y_1|^2} \equiv \overline{|y_2|^2}$ for any input x such that K is the identity matrix.

[1] SQ is a trademark of CBS, Inc.

[2] This interpretation is rigorous only for "sufficiently long" time averaging.

CONSEQUENCES OF THE STEREO-COMPATIBILITY OBJECTIVES

In the Appendix we prove that the equations for the three stereo performance objectives are satisfied if and only if the encoder matrix has the form

$$E = \begin{bmatrix} 1 & 0 & A & B \\ 0 & 1 & C & D \end{bmatrix} \qquad (10)$$

where

$$|A| = |D|, \qquad |B| = |C|, \qquad |A|^2 + |B|^2 = 1. \qquad (11)$$

There is no constraint on the phase angles of the complex numbers A, B, C, and D.

PERFORMANCE OBJECTIVE FOR MONO COMPATIBILITY

Next we state the mono performance objective in mathematical form.

Conservation of Power. For the class of uncorrelated inputs, the power in the mono output must equal the power in the input, i.e., $\overline{|y_1+y_2|^2} \equiv \overline{x^+x}$ for all inputs x such that K is diagonal.

CONSEQUENCES OF THE COMPATIBILITY OBJECTIVES

In the Appendix we prove that the equations for all four mono and stereo compatibility objectives are satisfied if and only if Eqs. (10) and (11) are satisfied for stereo compatibility along with the following additional equations needed for mono compatibility:

$$C = pc_1A \qquad (12)$$

$$D = p'B/c_1 \qquad (13)$$

where $p \triangleq \pm j$, $p' \triangleq \pm j$, and $c_1 = |B|/|A|$.

The synthesis up to this point has identified the family of compatible encoders that satisfy our mono and stereo performance objectives. The next step is to identify the decoders that satisfy our definition of optimality.

PERFORMANCE OBJECTIVES FOR QUADRAPHONIC OUTPUTS

To begin we must state the prescribed performance objectives for optimum decoders in mathematical form. For this purpose the error e in the quadraphonic output matrix x' is defined as the difference between it and the input matrix x:

$$e \triangleq x' - x. \qquad (14)$$

The mean power P_e in this error (the error power) is the sum of the mean error powers in each channel:

$$P_e \triangleq \overline{|\mathrm{LF}'-\mathrm{LF}|^2} + \overline{|\mathrm{RF}'-\mathrm{RF}|^2} + \overline{|\mathrm{RB}'-\mathrm{RB}|^2} + \overline{|\mathrm{LB}'-\mathrm{LB}|^2} \qquad (15)$$

which may be written compactly as

$$P_e \triangleq \overline{e^+e}. \qquad (16)$$

The first performance objective for the decoder is that this quadraphonic error power must be as small as possible. Such decoders yield minimum error in the mean-square sense. We are therefore led to the following synthesis problem for minimum-error decoders: Given a compatible encoder matrix E, and given a class of important inputs described by the correlation matrix K, find the set of all decoder matrices D such that the error power P_e is minimized.

This problem is analyzed in the Appendix where it is shown that the decoder D yields minimum error if $D = D_0$, where

$$D_0 = KE^+(EKE^+)^\#. \qquad (17)$$

$(EKE^+)^\#$ denotes the pseudo inverse [4]–[6] of EKE^+.

This decoder is a function of the input correlation matrix K, so that as the input correlations change, so does the decoder structure. A decoder that changes its structure to suit the input correlations is said to have logic. Eq. (17) describes the minimum-error logic decoder that has the smallest possible norm. The significance of "smallest possible norm" is discussed in the next paragraph.

In the Appendix it is proven that if the correlation matrix K has a special mathematical property, then there exist additional decoders that yield exactly the same decoding errors as D_0. These additional decoders all have larger norms than D_0, which means that their gains are larger than absolutely necessary to minimize the error power.

We now impose the condition that the decoder be time invariant (no logic). We must therefore choose a specific K for which our decoder is to yield minimum error. For this purpose we use the class of uncorrelated unit-power inputs; K is then the identity matrix. This class appears to include the set of "worst case" inputs and therefore yields the decoder that minimizes the maximum error power. For this class of inputs Eq. (17) simplifies to

$$D_0 = E^+(EE^+)^\#. \qquad (18)$$

and the resulting minimized error power is shown in the Appendix to be

$$P_e = 4 - \mathrm{rank}\,(E). \qquad (19)$$

Eq. (19) is interesting; it states that the total error power produced with the minimum-error decoder depends *only* on the rank of the encoder matrix [rank (E) = number of linearly independent rows in E]. This means that the minimized quadraphonic error power is independent of the detailed structure of the encoder for the class of uncorrelated unit-power inputs; the larger the rank of E, the smaller the minimized error power.

The maximum possible rank of E is two in 4–2–4 matrix systems. All of the compatible encoders identified earlier in this paper have this maximum rank, and they therefore all yield the same error power when decoder D_0 is used. (It can be proven that D_0 is the *only* decoder that minimizes the error power when $K = I_4$ and rank $(E) = 2$.) For these encoders EE^+ is nonsingular. It then follows that $(EE^+)^\# = (EE^+)^{-1}$, and so Eq. (18) may be written:

$$D_0 = E^+(EE^+)^{-1}. \qquad (20)$$

Any decoder (without logic) that satisfies Eq. (20) is said to be *matched* to the encoder for uncorrelated unit-power inputs. It can be shown that another way of writing Eq. (20) is $D_0 = E^{\#}$, which states that matched decoders (for uncorrelated unit-power inputs) are pseudo inverses of their encoders.

PERFORMANCE OBJECTIVE FOR ROBUST DECODERS

The only remaining performance objective for optimal decoding is the second one regarding robustness. A decoder is said to be *robust* if its performance is relatively insensitive to small errors in its element values. The matched decoder in Eq. (20) cannot be robust if EE^+ is ill-conditioned [7], [12], i.e., if the eigenvalues of EE^+ are greatly different from each other. The most robust decoder is the one for which the eigenvalues of EE^+ are equal to each other, which happens if and only if EE^+ is proportional to the identity matrix. We therefore want to choose our encoder so that

$$EE^+ = cI_2 \quad (21)$$

where I_2 is the 2×2 identity matrix, and c is a positive number. Any encoder that satisfies Eq. (21) is termed an *orthonormal* encoder, because the rows of E are then orthogonal to each other and normalized. (In the Appendix we prove that exactly half of all compatible encoders are orthonormal.) Putting Eq. (21) in Eq. (20) yields the following prescription for the optimal (robust and matched) decoder:

$$D_0 = E^+/c. \quad (22)$$

This states that the optimal decoder is proportional to the adjoint of the (orthonormal) encoder.

SOME PROPERTIES OF OPTIMAL DECODING

1) From Eqs. (21) and (22) it follows that $D_0^+D_0 = I_2/c$, which shows that the optimal decoder is orthonormal, i.e., the columns of D_0 are orthogonal to each other and normalized.

2) If the decoder is orthonormal, then the total quadraphonic output power P_Q is proportional to the total stereo output power P_S, i.e., for *any* input x

$$P_Q \equiv P_S/c \quad (23)$$

where

$$P_Q \triangleq \overline{|\mathrm{LF}'|^2} + \overline{|\mathrm{RF}'|^2} + \overline{|\mathrm{RB}'|^2} + \overline{|\mathrm{LB}'|^2} \equiv \overline{(x')^+x'} \quad (24)$$

$$P_S \triangleq \overline{|\mathrm{LT}|^2} + \overline{|\mathrm{RT}|^2} \equiv \overline{y^+y}. \quad (25)$$

This is proved in the Appendix.

3) Matched decoders yield *projective* recording systems. To explain this we let Q denote the 4×4 transfer-function matrix that maps the input x into the quadraphonic output x'. Then

$$x' = Qx \quad (26)$$

$$Q = DE. \quad (27)$$

For matched decoding we must set $D = D_0$ [see Eq. (20)], and then

$$Q = E^+(EE^+)^{-1}E. \quad (28)$$

This Q has the remarkable property of being an *orthogonal projector*, which means that it satisfies the following two requirements:

Q is a projector,

$$QQ = Q. \quad (29)$$

Q is self-adjoint,

$$Q = Q^+. \quad (30)$$

The fact that Q is a projector has a dramatic engineering significance. Eq. (29) implies that two matrix systems connected in cascade, with the quadraphonic output of the first driving the input of the second, behave the same as a single system. In other words, the quadraphonic outputs are all the same in a cascade of matched systems; this is shown in Fig. 2. Note that the stereo output y is also invariant under cascading.

The fact that Q is self-adjoint is useful in checking the validity of calculations. In combination with Eq. (29) it also guarantees that for any input x the following identity holds between the total input power $P_x \triangleq \overline{x^+x}$, the total quadraphonic error power P_e, and the total quadraphonic output power P_Q:

$$P_x \equiv P_Q + P_e. \quad (31)$$

This is proved in the Appendix.

OPTIMAL COMPATIBLE MATRICES

In the first part of this paper we determined a set of necessary and sufficient conditions for meeting stereo and mono compatibility objectives; this identified the set of compatible encoders. Then another set of necessary and sufficient conditions was derived for optimum time-invariant decoding. These two sets of conditions will now be combined to establish the family of matrices that meets the compatibility and optimality objectives simultaneously. To this end we first impose the optimality condition in Eq. (21) on the compatibility conditions dictated by Eqs. (9)–(13). In the Appendix we prove that the only encoder

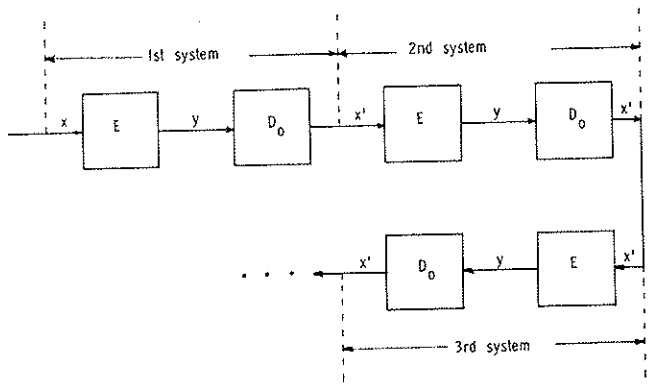

Fig. 2. Identical quadraphonic systems with matched decoders, $D_0 = E^+(EE^+)^{-1}$, can be cascaded without changing either the quadraphonic output x' or the stereo output y.

matrices that meet all these conditions simultaneously are of the form

$$E = \begin{bmatrix} 1 & 0 & A & B \\ 0 & 1 & pc_1A & -pB/c_1 \end{bmatrix} \quad (32)$$

where A and B are any two complex numbers such that $|A|^2 + |B|^2 = 1, p = \pm j$, and $c_1 = |B|/|A|$. It turns out that the orthonormalization constant $c = 2$. Eq. (22) then uniquely specifies the following optimal decoder corresponding to any E that satisfies Eq. (32):

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ A^* & -pc_1A^* \\ B^* & pB^*/c_1 \end{bmatrix} 0.5. \quad (33)$$

IMPLEMENTATION

Eqs. (32) and (33) are the necessary and sufficient conditions for compatibility and time-invariant optimality; these conditions define a family of matrix recording systems termed the SQ family. Detailed descriptions of two well-known members may be found in the literature: these are basic SQ [9]–[11] (which corresponds to $p = j, A = 0.707$, and $B = -0.707j$), and forward-oriented SQ [10] (which corresponds to $p = j,\ A = -0.707j$, and $B = 0.707$).

SUMMARY

The key features of this paper are summarized below.

1) Linear time-invariant 4–2–4 matrix recording systems are defined mathematically.

2) Compatible encoding is defined in terms of four mono and stereo performance objectives.

3) Optimal time-invariant decoding is defined in terms of two quadraphonic performance objectives.

4) All compatible encoders and their optimal time-invariant decoders (no logic) are identified; they form the SQ family of matrix systems.

5) Optimal time-invariant decoding is shown to possess three interesting properties:

a) The total quadraphonic output power is always proportional to the total stereo output power.

b) The total input power is always equal to the sum of the quadraphonic output power and the quadraphonic error power.

c) The overall transformation from quadraphonic input to quadraphonic output is an orthogonal projection.

6) The most accurate possible decoder (in the mean-square sense of minimizing the power in the decoding errors) is derived and shown to be a logic decoder with a specific structure that depends on the correlation matrix of the input signals.

APPENDIX

MATHEMATICAL DETAILS

Analytic Signals

Let $s(t)$ denote a real zero-mean function of time with Hilbert transform $\hat{s}(t)$, which is defined as

$$\hat{s}(t) \triangleq \int_{-\infty}^{\infty} (t - t')^{-1} s(t')\, dt'/\pi \quad (34)$$

where the Cauchy principle value is intended. The analytic signal representation for $s(t)$ is defined here as

$$z(t) \triangleq 2^{-0.5}[s(t) + j\hat{s}(t)]. \quad (35)$$

This definition differs from the usual one for analytic signals [1], [8] because of the scale factor $2^{-0.5}$, which is included here so that $z(t)$ and $s(t)$ have identical average powers ($\overline{|z|^2} = \overline{s^2}$, where the overbar denotes time average).

An equivalent definition for the analytic signal is stated below in terms of the Fourier transforms of $s(t)$:

$$z(t) \triangleq 2^{0.5} \int_{0}^{\infty} S(f) \exp(j2\pi ft)\, df \quad (36)$$

where the transform

$$S(f) \triangleq \int_{-\infty}^{\infty} s(t) \exp(-j2\pi ft)\, dt. \quad (37)$$

Eq. (36) states that the analytic signal is the inverse Fourier transform of $S(f)$ restricted to positive frequencies and scaled by the square root of two.

For an example, suppose that $s(t) = \cos \omega t$, then $\hat{s}(t) = \sin \omega t$ and $z(t) = 2^{-0.5} \exp(j\omega t)$. Note that $z(t)$ is the phasor representation of $s(t)$ with the $\exp(j\omega t)$ factor included; also, the average power of $z(t)$ and $s(t)$ are identical ($\overline{|z|^2} = \overline{s^2} = ½$).

STEREO COMPATIBILITY

Let the encoder E and input x be partitioned:

$$E \triangleq [W \quad T] \quad (38)$$

$$x \triangleq \begin{bmatrix} a \\ b \end{bmatrix} \quad (39)$$

where W and T are 2×2 matrices and a and b are 2×1 matrices of complex numbers. It then follows that the stereo output y is

$$y = Ex = Wa + Tb. \quad (40)$$

The first stereo compatibility objective requires that if $b = 0$, then $y = a$ for any a. This objective is satisfied if and only if $W = I_2$, where I_2 is the 2×2 identity matrix. Thus E must have the form

$$E = \begin{bmatrix} 1 & 0 & A & B \\ 0 & 1 & C & D \end{bmatrix} \quad (41)$$

where A, B, C, and D are complex numbers.

The second stereo compatibility objective states that the stereo output power ($P_S \triangleq \overline{y^+y}$) and the input power ($P_x \triangleq \overline{x^+x}$) must be equal for any input x such that the correlation matrix ($K \triangleq \overline{xx^+}$) is diagonal ($\overline{x_ix_j^*} = 0,\ i \triangleq j$). Direct calculation shows that $P_S = P_x$ in this case if and only if

$$|A| + |C|^2 = 1 \quad (42)$$

$$|B|^2 + |D|^2 = 1. \quad (43)$$

The third stereo compatibility objective states that $\overline{|y_1|^2} = \overline{|y_2|^2}$ for any x such that $K = I_4$. Direct calculation shows that this objective is satisfied if and only if

$$|A|^2 + |B|^2 = |C|^2 + |D|^2. \tag{44}$$

The general solution of Eqs. (42)–(44) is any complex A, B, C, D such that $|A| = |D|$, $|B| = |C|$, and $|A|^2 + |B|^2 = 1$.

MONO COMPATIBILITY

The mono compatibility objective states that $\overline{|y_1 + y_2|^2} = \overline{x^+x}$ for any x such that K is diagonal with nonnegative elements. Direct calculation shows that this objective is satisfied if and only if

$$|A + C| = 1 \tag{45}$$

$$|B + D| = 1. \tag{46}$$

The previous analysis implies that necessary and sufficient conditions for stereo compatibility are that A, B, C, D have the form

$$A = c \exp(j\alpha) \tag{47}$$

$$B = c' \exp(j\beta) \tag{48}$$

$$C = c' \exp(j\gamma) \tag{49}$$

$$d = c \exp(j\delta) \tag{50}$$

where $c \triangleq |A| \leqslant 1$, $c' \triangleq (1 - c^2)^{0.5}$, and $\alpha, \beta, \gamma, \delta$ = real numbers. Eqs. (45), (47), and (49) imply that $1 = |c \exp(j\alpha) + c' \exp(j\gamma)| \equiv |c \exp(j(\alpha\text{-}\gamma)) + c'|$, which implies that

$$|\alpha - \gamma| = \pi/2 \tag{51}$$

because $c^2 + (c')^2 = 1$.

Similar arguments based on Eqs. (46), (48), and (50) imply that

$$|\beta - \delta| = \pi/2. \tag{52}$$

Conversely, Eqs. (47)–(50) together with Eqs. (51) and (52) imply Eqs. (45) and (46). Therefore the encoder E satisfies the four compatibility objectives for both stereo and mono if and only if Eqs. (41) and (47)–(52) are satisfied. These conditions may be written $c = pc_1 A$, $D = p'B/c_1$, where $p \triangleq \pm j$, $p' \triangleq \pm j$, $c_1 \triangleq c_1 = |B|/|A|$, and $|A|^2 + |B|^2 = 1$.

DECODER SYNTHESIS

Statement of the problem of minimizing the decoding errors: Given any 2×4 encoder matrix E and 4×4 input correlation matrix $K = \overline{xx^+}$, find the decoder matrix D such that the error power $P_e = \overline{e^+e}$ is minimized, where the error matrix $e \triangleq x' - x = (DE - I_4)x$.

Solution: P_e is a quadratic smooth function of D. It follows from variational calculus that P_e is minimized by setting $D = D_0$ if and only if the first variation of P_e with respect to D_0 vanishes ($\delta P_e \equiv 0$), and the second variation $\delta^2 P_e \geqslant 0$.

The following notation and facts are used in deriving the minimum-error decoder. The trace[3] of the square matrix M, denoted by $\mathrm{tr}(M)$, and the adjoint, denoted by M^+, have the following useful properties, where M and N are any suitably dimensioned matrices over the field of complex numbers:

$$(MN)^+ \equiv N^+M^+ \tag{53}$$

$$\mathrm{tr}(M + N) \equiv \mathrm{tr}(M) + \mathrm{tr}(N). \tag{54}$$

For any number c,

$$\mathrm{tr}(cM) \equiv c\,\mathrm{tr}(M) \tag{55}$$

$$\mathrm{tr}(MN) \equiv \mathrm{tr}(NM) \tag{56}$$

$$\mathrm{tr}(M^+) \equiv (\mathrm{tr}(M))^*. \tag{57}$$

For any M,

$$\mathrm{tr}(M^+M) \geqslant 0. \tag{58}$$

If $\mathrm{Re}[\mathrm{tr}(MN)] = 0$ for every suitably dimensioned N, then

$$M = 0. \tag{59}$$

We are now prepared to begin calculating the minimum-error decoder D_0. This will be done by first calculating an expression for P_e in terms of D and then taking the variations of this expression.

In view of Eq. (55), we may write

$$P_e \equiv \mathrm{tr}(P_e). \tag{60}$$

By definition $P_e \triangleq \overline{e^+e}$, thus

$$P_e \equiv \mathrm{tr}\,(\overline{e^+e}). \tag{61}$$

By Eq. (57) this may be written

$$P_e \equiv \mathrm{tr}\,(\overline{ee^+}). \tag{62}$$

By definition $e \triangleq (DE - I_4)\,x$, thus

$$P_e \equiv \mathrm{tr}[\,(DE - I_4)\,\overline{xx^+}\,(DE - I_4)^+]. \tag{63}$$

By definition $K \triangleq \overline{xx^+}$, so in view of Eqs. (53) and (54), we may multiply out the terms and write

$$P_e \equiv \mathrm{tr}(DEKE^+D^+) - \mathrm{tr}(DEK) - \mathrm{tr}(KE^+D^+) + \mathrm{tr}(K). \tag{64}$$

We now calculate the variations of P_e with respect to D_0. These variations are defined as follows. In Eq. (64) set $D = D_0 + \delta D$, where δD is an arbitrary 4×2 matrix called the variation of D_0. Set $P_e \triangleq P_{e0} + \delta P_e + \delta^2 P_e$, where δP_e is called the first variation of P_e and $\delta^2 P_e$ is called the second variation of P_e. The first variation δP_e is defined as that part of P_e that depends linearly on δD, while the second variation $\delta^2 P_e$ is defined as that part of P_e that depends quadratically on δD.

The first variation of P_e with respect to D_0 is therefore (we use the fact that $K = K^+$)

$$\delta P_e = \mathrm{tr}(\delta DEKE^+D_0^+)^+ + \mathrm{tr}(D_0EKE^+\delta D^+) - \mathrm{tr}(\delta D\,EK) - \mathrm{tr}(KE^+\delta D^+). \tag{65}$$

By using Eqs. (53) and (57), Eq. (65) may be simplified:

$$\delta P_e = 2\,\mathrm{Re}[\mathrm{tr}(D_0EKE^+\delta D^+) - \mathrm{tr}(KE^+\delta D^+)]. \tag{66}$$

[3] Trace = sum of diagonal elements.

Finally, Eqs. (54) and (55) may be used to write

$$\delta P_e = 2\,\mathrm{Re}\{\mathrm{tr}[(D_0EKE^+ - KE^+)\delta D^+]\}. \quad (67)$$

In view of Eqs. (55) and (59), it follows from Eq. (67) that $\delta P_e = 0$ for any complex δD if and only if

$$D_0EKE^+ - KE^+ = 0. \quad (68)$$

The general solution to this equation [4], [6] is

$$D_0 = KE^+(EKE^+)^\# + F \quad (69)$$

$$F = Z(I_2 - EKE^+(EKE^+)^\#) \quad (70)$$

where Z is *any* 4×2 matrix. It is known that [4], [6] that D_0 has its minimum norm when $Z = 0$, and that $F = 0$ whenever rank $(EKE^+) = 2$.

To prove that this prescription for D_0 minimizes P_e we must show that the second variation $\delta^2 P_e \geqq 0$. To calculate this second variation, we use Eq. (67) and take the first variation of δP_e, which yields $\delta^2 P_e = 2\mathrm{tr}(\delta DEKE^+\delta D^+)$. Now $K = K^+$; therefore there exists a matrix S such that $K = SS^+$. Thus

$$\delta^2 P_e = 2\mathrm{tr}[(\delta DES)(\delta DES)^+] \quad (71)$$

$$\delta^2 P_e = 2\mathrm{tr}[(\delta DES)^+(\delta DES)]. \quad (72)$$

In view of Eq. (58), it follows from Eq. (72) that $\delta^2 Pe \geqq 0$ for any δD. The necessary and sufficient conditions for minimum error are now established.

We have just proven that D minimizes P_e if and only if Eqs. (69) and (70) are satisfied for some 4×2 matrix Z. The minimum P_e [calculated by using Eq. (69) for D] is found, after some algebra, to be

$$P_e = \mathrm{tr}(K) - \mathrm{tr}(EK^2E^+(EKE^+)^\#). \quad (73)$$

Note that the minimum P_e does not depend on the arbitrary matrix Z. In fact, the F term in Eq. (69) makes no contribution to the quadraphonic output $x' \triangleq D_0Ex$ because $FEx = 0$. This assertion is equivalent to asserting that the power in FEx is zero, i.e., that $\overline{(FEx)^+FEx} \equiv \mathrm{tr}(FEKE^+F^+) = 0$ for any correlation K. This last equation is easily verified by direct calculation because of the following two identities for pseudo inverses: 1) $MM^\#M \equiv M$; 2) $(M^+)^\# \equiv (M^\#)^+$.

For the special case of uncorrelated unit-power inputs, which are used to synthesize optimal time-invariant decoders, the correlation matrix $K = I_4$, and Eq. (73) reduces to $P_e = 4 - \mathrm{tr}(EE^+(EE^+)^\#)$. The pseudo inverse has the property [4], [6] that for any E, $\mathrm{tr}(EE^+(EE^+)^\#)$ = the rank of E. Thus for uncorrelated unit-power inputs,

$$P_e = 4 - \mathrm{rank}(E). \quad (74)$$

PROPERTIES OF MINIMUM-ERROR AND OPTIMAL DECODING

1) Let $Q \triangleq DE$ denote the transfer-function matrix for the quadraphonic output $(x' \equiv Qx)$. Let $D = D_0$, where D_0 is the minimum-error decoder in Eq. (69). Then the transfer-function matrix $Q = D_0E$ is a projector [6] because $QQ = Q$. This is easily proven by direct calculation with the fact that $FE = 0$. If in addition $K = I_4$, then $Q = Q^+$, and Q is therefore an orthogonal projector [6].

2) If Q is an orthogonal projector, then the input power is identical to the sum of the quadraphonic error power and the quadraphonic output power:

$$P_x \equiv P_e + P_Q. \quad (75)$$

To prove this we note that the input power $P_x \triangleq \overline{x^+x} \equiv \mathrm{tr}\,(K)$. The error power $P_e \triangleq \overline{e^+e} \equiv \mathrm{tr}\,(Q - I_4)\,\overline{xx^+}\,(Q - I_4)^+ \equiv \mathrm{tr}(QKQ^+ - KQ^+ - QK + K) \equiv \mathrm{tr}\,(KQ^+Q - KQ^+ - KQ + K) \equiv \mathrm{tr}\,(KQQ - 2KQ + K) \equiv \mathrm{tr}\,(K) - \mathrm{tr}\,(KQ) \equiv P_x - \mathrm{tr}\,(KQ)$. Finally the quadraphonic power $P_Q \triangleq \overline{(x')^+x'} \equiv \mathrm{tr}\,(Q\overline{xx^+}Q^+) \equiv \mathrm{tr}\,(QKQ^+) \equiv \mathrm{tr}\,(KQ^+Q) \equiv \mathrm{tr}\,(KQQ) \equiv \mathrm{tr}\,(KQ) \equiv P_x - P_e$.

3) If the decoder is orthonormal $(D_0^+D_0 = I_2/c,\ c > 0)$, then the stereo output power is identically proportional to the quadraphonic power:

$$P_S \equiv cP_Q. \quad (76)$$

To prove this, we note that $P_S \triangleq \overline{y^+y} \equiv \mathrm{tr}\,\overline{(Exx^+E^+)} \equiv \mathrm{tr}\,(EKE^+)$. The quadraphonic power $P_Q \triangleq \overline{(x')^+x'} \equiv \mathrm{tr}\,\overline{((Qx)^+(Qx))} \equiv \mathrm{tr}\,(Q\overline{xx^+}Q^+) \equiv \mathrm{tr}\,(D_0\,EKE^+D_0\,) \equiv \mathrm{tr}\,(EKE^+D_0^+D_0) \equiv (1/c)\,\mathrm{tr}\,(EKE^+) \equiv (1/c)\,P_S$.

REFERENCES

[1] D. Gabor, *J. Inst. Elec. Engr.*, vol. 93, pt. III, p. 429 (1946).

[2] D. H. Cooper and T. Shiga, "Discrete-Matrix Multichannel Stereo," *J. Audio Eng. Soc.*, vol. 20, pp. 346-360 (June 1972).

[3] M. A. Gerzon, "Periphony: With-Height Sound Reproduction," *J. Audio Eng. Soc.*, vol. 21, pp. 2-10 (Jan./Feb. 1973).

[4] R. Penrose, "A Generalized Inverse for Matrices," *Proc. Cambridge Phil. Soc.*, vol. 51, pt. 3, pp. 406-413 (1955).

[5] T. N. E. Greville, "Some Applications of the Pseudoinverse of a Matrix," *SIAM Rev.*, vol. 2, pp. 15-22 (Jan. 1960).

[6] L. A. Zadeh and A. Desoer, *Linear System Theory*. (McGraw-Hill, New York, 1963), pp. 577-582.

[7] B. Noble, *Applied Linear Algebra*. (Prentice-Hall, Englewood Cliffs, NJ, 1969), p. 433 and example 13.26 on p. 454.

[8] A. Papoulis, *Systems and Transforms with Applications in Optics*. (McGraw-Hill, New York, 1968), pp. 83-89, 274-276.

[9] B. B. Bauer, D. W. Gravereaux, and A. J. Gust, "A Compatible Stereo-Quadraphonic (SQ) Record System," *J. Audio Eng. Soc.*, vol. 19, pp. 638-646 (Sept. 1971).

[10] B. B. Bauer, G. A. Budelman, and D. W. Gravereaux, "Recording Techniques for SQ Matrix Quadraphonic Discs," *J. Audio Eng. Soc.*, vol. 21, pp. 19-26 (Jan./Feb. 1973).

[11] B. B. Bauer, R. G. Allen, G. A. Budelman, and D. W. Gravereaux, "Quadraphonic Matrix Perspective—Advances in SQ Encoding and Decoding Technology," *J. Audio Eng. Soc.*, vol. 21, pp. 342-350 (June 1973).

[12] M. P. Ekstrom, "A Spectral Characterization of the Ill-Conditioning in Numerical Deconvolution," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 344-348 (Aug. 1973).

Dr. White's biography appeared in the April 1975 issue.