Codes for Multiplex Spectrometry

N. J. A. Sloane, T. Fine, P. G. Phillips, and M. Harwit

A number of binary cyclic coding schemes for multiplex spectrometry are discussed and evaluated in terms of a linear, least mean square, unbiased estimate. The optical realization of such codes in dispersion instruments is briefly discussed. We show that there are many advantages both in the construction of the instrument and in its operation which accrue from cyclic codes.

I. Introduction

In two recent articles, Ibbett et al. and Decker and Harwit have independently suggested the use of sequentially stepped multiplex spectrometers. In both systems radiation enters a dispersion instrument through a single slit and is analyzed at a number of exit slits. In this process, combinations of $N$ spectral elements are observed sequentially, in $M$ successive time intervals, by placing $M$ different encoding masks in the exit plane of the spectrometer. In the arrangement of Ibbett et al. each mask has a prescribed arrangement of $2^m - 1$ slots, where $m$ is an integer. These authors have worked out the case $m = 3$, but the method can be generalized to any $N = M = 2^m - 1$.

Decker and Harwit suggested an arrangement of $N/2$ transparent and $N/2$ opaque slots. They have worked out a specific encoding technique for any number $N = 4m + 2$, where $m$ is an integer.

A recently published paper by Gottlieb brings a related approach to bear on the problem of ir television imaging.

Let $W$ be an $N \times M$ matrix whose columns represent the $M$ encoding masks successively placed in the exit plane of the spectrometer. A very desirable characteristic of $W$ is that it be cyclic; that is, the $(i + 1)^{st}$ column of $W$ is obtained by shifting the $i^{th}$ column cyclically one place downwards. This means that a mask like the one drawn in Fig. 1 can be used with the spectrometer. This corresponds to the matrix of mask weights

$$
W = \begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 & 1 & 0 & 1 \\
1 & 1 & 1 & 0 & 0 & 1 & 0
\end{bmatrix}
$$

which is the matrix $S^T$ of Theorem 1, in the case when $N = M = 8$ and $g$ is given by Eq. (3).

Such a mask has two advantages. First, it can be self-supporting and therefore permits the construction of a spectrometer which requires no transmission materials. In operation the mask is stepped one slit-width—along the length of the mask—for each successive encoding position. The required accuracy of this motion normally can be kept within micrometer tolerances and is less stringent than the tolerances associated with interferometric spectrometers. The code actually described in Ref. 1 is not cyclic, but it can be made cyclic through suitable row and column permutations [cf. Fig. 1 and Eq. (3)]. The code of Ref. 2 falls into the category of $2 \times 2$ block circulants. This code, however, is not as desirable as a purely cyclic one.

Cyclic codes do exist and their use enables one to avoid the tedious construction of $N$ masks with $N$ slots for a total of $N^2$ slots. Instead, one constructs only one mask with $2N - 1$ slots. This represents the second advantage of the proposed scheme. The cost of mask construction is reduced by $\sim N/2$ and the design of the advance mechanism is considerably simplified since the total weight of the masks also decreases as $\sim N/2$.

II. Estimation of Spectral Shape

The optical spectrum whose shape is to be estimated is spatially dispersed and the band of interest partitioned into $N$ channels. The average energy in the $i^{th}$ channel, after a selected integration time, is denoted by

Received 17 February 1969.
Fig. 1. Mask with $2N - 1$ ($N = 7$) slots for encoding the spectral distribution. Shaded areas are opaque and represent 0's.

The clear areas represent 1's.

$E_j$. The measurement process consists of observing the spectrum through $M$ filters, or masks, the energy through the $i$th mask being $\sum_{j=1}^{N} \omega_{ij} E_j$, where $W_i = (\omega_{1i}, \ldots, \omega_{Ni})$ is the $i$th vector of weights. The mask weights are assumed such that $|\omega_{ij}| \leq 1$ and correspond to attenuation. The photodetector adds a random noise $n_i$ to the signal $\sum_{j=1}^{N} \omega_{ij} E_j$ and yields a measurement

$$x_i = n_i + \sum_{j=1}^{N} \omega_{ij} E_j, \quad i = 1, 2, \ldots, M.$$ 

With the notation $\langle \rangle$ for ensemble averages, the noise $n_i$ has the following properties: $\langle n_i \rangle = 0$; $n_i$ is independent of the signal; $\langle n_i^2 \rangle = \sigma^2$; successive measurement noises are assumed to be uncorrelated ($\langle n_i n_j \rangle = 0$ if $i \neq j$).

In order to estimate the set of true energies $\{E_j\}$ by estimates (functions of the observations) $\{\hat{E}_j\}$, one needs at least as many measurements ($M$) as there are unknowns ($N$). Furthermore, at least $N$ distinct masks $\{W_i\}$ are needed if one hopes to estimate the spectral shape. Hence, assume a vector of observations $x = (x_1, \ldots, x_M)$, channel energies $E = (E_1, \ldots, E_N)$, measurement noises $n = (n_1, \ldots, n_M)$, and a matrix of masks $W = (W_1^T, \ldots, W_M^T) = (\omega_{ij})$. (The $T$ stands for transpose.) With this notation

$$x = EW + n.$$ 

An estimate $\hat{E}$ of $E$ is a function of the observations, $\hat{E}(x)$, hopefully lying close to $E$. As a measure of the accuracy of the estimate, we adopt the mean square error criterion: minimize $\epsilon = \langle (\hat{E} - E)(\hat{E} - E)^T \rangle$.

For purposes of computational convenience we agree to restrict $\hat{E}$ to be a linear function of the observations,

$$\hat{E} = xA,$$

for some matrix $A$; in the absence of more detailed statistical knowledge concerning the anticipated spectral shape or the photodetector noise characteristics, there is little alternative to the assumption of linearity. One more assumption is needed before an essentially uniquely best experimental design can be derived—the assumption that the estimator is unbiased; i.e., $\langle \hat{E} \rangle = E$. This assumption can be defended on the following grounds: (1) we desire an estimator which, on the average for a large number of applications, yields the true value; (2) the unbiased estimator can be shown to be desirable when there is a large uncertainty in $E$ relative to our prior knowledge of the spectrum and the measurement noise power. Further discussions of the approach outlined above are available in many texts on statistics.

We must now select those matrices $A$, $W$ that minimize $\epsilon$ subject to the constraint that $|\omega_{ij}| \leq 1$ and $\langle \hat{E} \rangle = E$. First observe that

$$\langle \hat{E} \rangle = \langle xA \rangle = \langle x \rangle A.$$

Furthermore,

$$\langle x \rangle = EW + \langle n \rangle = EW.$$

Hence, unbiasedness requires that $WA = I$, the identity matrix. If $M = N$, $W$ is square, and $A = W^{-1}$. If $W$ is not square but is $N \times M$ ($M > N$), one uses the generalized inverse

$$A = W^T(WW^T)^{-1}.$$

Having solved for $A$ in terms of $W$, one now selects $W$ to minimize $\epsilon = \langle (\hat{E} - E)(\hat{E} - E)^T \rangle$. Note that $\hat{E} - E = EWA + nA - E = nWW^T(WW^T)^{-1}$. By assumption, $\langle n^Tn \rangle = \sigma^2$. Thus, $\epsilon / \sigma^2 = \text{Trace} [WW^T(WW^T)^{-1}]$. If $M = N$, this simplifies to $\epsilon / \sigma^2 = \text{Trace} [W^{-1}W^{-1}]$. The optimum experimental design will be completed if one can find that $W$, subject to $|\omega_{ij}| \leq 1$, has the minimum $\epsilon / \sigma^2$.

Several convenient choices for $W$ yielding small values for $\epsilon$ are discussed in the Sec. III. The experimental design to be derived under the above hypotheses can be shown to be equivalent to that dictated by the linear, least squares method; this further supports the value of the proposed design.

III. Choice of Mask Weights

In the previous section it was shown that in the case $M = N$, when there are as many measurements as unknowns, the matrix $W = (\omega_{ij})$ of mask weights should be chosen so that $|\omega_{ij}| \leq 1$ and $Tr[W^{-1}(W^{-1})^T]$ is as small as possible. Three possible choices for the matrix $W$ are given here. They are summarized in Theorem 1.

A Hadamard matrix $H$ of order $N$ (Ref. 6) is an $N \times N$ matrix of +1's and -1's such that

$$HH^T = N I.$$

(1)

$H$ may always be normalized so that the first row and column consist entirely of +1's. Let $G$ denote the remaining $(N - 1) \times (N - 1)$ matrix, thus:

$$H = [1 \quad 1 \quad \ldots \quad 1 \quad 1]$$

Let row $i$ and column $j$ be any two distinct rows of $H$ other than the first row. Then it is easily shown that, assuming $N \geq 4$,
Table I. First Row of the G Matrix for Low Values of $N - 1$

<table>
<thead>
<tr>
<th>$N - 1$</th>
<th>First row of cyclic $G$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>++--</td>
</tr>
<tr>
<td>7</td>
<td>+---++</td>
</tr>
<tr>
<td>11</td>
<td>+---+++</td>
</tr>
<tr>
<td>15</td>
<td>++++--</td>
</tr>
<tr>
<td>19</td>
<td>++++</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N - 1$</th>
<th>First row of cyclic $G$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>++--</td>
</tr>
<tr>
<td>7</td>
<td>+---++</td>
</tr>
<tr>
<td>11</td>
<td>+---+++</td>
</tr>
<tr>
<td>15</td>
<td>++++--</td>
</tr>
<tr>
<td>19</td>
<td>++++</td>
</tr>
</tbody>
</table>

row $i$ has $+1$ and row $j$ has $+1$ in $N/4$ places

row $i$ has $+1$ and row $j$ has $+1$ in $N/4$ places

row $i$ has $+1$ and row $j$ has $+1$ in $N/4$ places

row $i$ has $+1$ and row $j$ has $+1$ in $N/4$ places

row $i$ has $+1$ and row $j$ has $+1$ in $N/4$ places

Thus $N$ must be a multiple of four.  (It is conjectured that Hadamard matrices exist whenever $N$ is a multiple of four.) Further, if $N$ is a multiple of four, and at least one of the following conditions are also satisfied:

(i) $N = p + 1$, $p$ prime
(ii) $N = p(p + 2) + 1$, $p$ and $p + 2$ prime
(iii) $N = 2^w$

then $G$ can be chosen to be a cyclic matrix, that is, a matrix in which the $(i + 1)^{st}$ row is obtained by shifting the $i^{th}$ row cyclically one place to the right. For example, when $N = 8$, $G$ may be taken to be

$$
G = \begin{pmatrix}
0 & + & - & - & + & + & + & - \\
+ & + & - & - & + & + & - & - \\
+ & - & + & + & + & - & - & + \\
- & + & + & - & - & - & + & + \\
- & - & - & - & - & - & - & + \\
- & - & - & - & - & - & - & + \\
- & - & - & - & - & - & - & + \\
- & - & - & - & - & - & - & + \\
\end{pmatrix}
$$

where $+$ stands for $+1$ and $-$ for $-1$.

Table I gives the first row of $G$ for the first few values of $N$ for which $G$ can be made cyclic. A more complete table can be found in Appendix 2 of Ref. 7, and a list of those values of $N$ for which it is presently known that a cyclic $G$ exists is given in Refs. 8 and 9.

From Eq. (2) we calculate the dot products of any two rows of $H$ or of $G$:

**In $H$:** row $i \cdot row j = \begin{cases}
0 & i \neq j \\
N & i = j
\end{cases}$

**In $G$:** row $i \cdot row j = \begin{cases}
1 & i \neq j \\
N - 1 & i = j
\end{cases}$

Also each row of $G$ contains $[(N/2) - 1] + 1$’s and $(N/2) - 1$’s.

The first choice for $W$ is the matrix $H^T$. From Eq. (1) one obtains $H^{-1} = N^{-2}HH^T$, and $Tr((H^{-1})^2H^{-1}) = Tr(N^{-2}HH^T) = Tr(N^{-1}I) = 1$.

The second choice for $W$ is the matrix $G^T$. From Eq. (4), $GG^T = NI - J$, where $J$ is an $(N - 1) \times (N - 1)$ matrix of $+1$’s, and

$$
GG^T = \left[ \begin{pmatrix} N & 1 \\ 2 & -1 \end{pmatrix} + \frac{N}{2} (-1) \right] J = -J
$$

so $G^{-1} = \left( 1/N \right) (G^T - J)$. Then

$$
Tr[(G^{-1})^2G^{-1}] = Tr \left\{ \frac{1}{N^2} [(G - J)(G^T - J)] \right\}
$$

$$
= Tr \left\{ \frac{1}{N} (I + J) - 2 - \frac{2}{N} \right\}
$$

The last choice for $W$ is $S^T$, where $S$ is the matrix obtained from $G$ by replacing $+1$’s by 0’s and $-1$’s by +1’s. Clearly each row of $S$ contains $[(N/2) - 1]$ 0’s and $N/2$ +1’s, and from Eq. (2):

In $S$: row $i \cdot row j = \begin{cases}
N/4 & i \neq j \\
N/2 & i = j
\end{cases}$

It is straightforward to calculate

$$
SS^T = (N/4)(I + J),
$$

$$
SJ = JS^T = (N/2)J,
$$

$$
S^{-1} = (2/N)(2S^T - J).
$$

These results are summarized in Theorem 1: Let $H$ be an $N \times N$ normalized Hadamard matrix, $N \geq 4$; let $G$ be the $(N - 1) \times (N - 1)$ matrix obtained by deleting the first row and column of $H$; and let $S$ be obtained from $G$ by replacing $+1$’s by 0’s and $-1$’s by +1’s. Then, three possible choices for the matrix $W$ of mask weights, together with the corresponding values of $Tr[W^{-1}(W^{-1})^T]$, are as follows:

<table>
<thead>
<tr>
<th>Matrix $W$</th>
<th>Trace</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H^T$</td>
<td>1</td>
<td>$+1, -1$</td>
</tr>
<tr>
<td>$G^T$</td>
<td>$2 - \frac{2}{N}$</td>
<td>$+1, -1$</td>
</tr>
<tr>
<td>$S^T$</td>
<td>$4 - \frac{8}{N} + \frac{4}{N^2}$</td>
<td>$+1, 0$</td>
</tr>
</tbody>
</table>

Remarks

(a) As mentioned above, for many values of $N$ the matrices can be chosen so that $G$ and $S$ are cyclic, and $H$ is cyclic with an extra row and column of $+1$’s added. This considerably simplifies the design of the masks, as is shown in the first section.

(b) For the special case $N = 2^w$, matrices with the same properties as $S^T$ were obtained by a different method in Ref. 1.

(c) Another choice for $W$ is $R^T$, where $R$ is the matrix obtained from $G$ by replacing $-1$’s by 0’s. This gives a trace, however, which is uniformly worse than that from $S^T$.

(d) All the matrices given in Theorem 1 are superior to the single slit case (where $W = I$) and to the block circulant suggested by Decker and Harwit.3

October 1969 / Vol. 8, No. 10 / APPLIED OPTICS 2105
Unfortunately we have no systematic scheme for generating successful encoding patterns. We can only evaluate a given scheme once it is presented. We therefore have no knowledge about the best code that can be generated on an absolute basis.

IV. Optical Arrangement

A multiplex spectrometer using the $S^T$ code works as follows: Radiation which passes through the entrance aperture is rendered parallel and directed toward the dispersive element. The dispersed radiation is collimated and focused upon the multislit mask at the exit plane of the instrument. The radiation transmitted by the mask passes through the postoptics of the system and impinges upon the detector. One then obtains a spectrum by sequentially stepping $M$ masks at the exit plane and recording the output for each mask. The inversion procedure described in Sec. II enables one to recover the spectrum

$$E = x(S^T)^{-1}.$$  

In some instances, depending upon instrumental design and the choice of code, it is advantageous to measure the reflected and transmitted radiation of a mask placed at an angle to the direction of the exiting radiation. This scheme is particularly useful for masks which utilize the $H^T$ and $G^T$ matrices. Here the $+1$'s represent reflecting slots and the $-1$'s represent transmitting slots. For the $H^T$ matrix, all elements of the first rows are $+1$'s and the masks corresponding to each column of $H^T$ will reflect the first spectral element at all times. The remaining $N - 1$ slots are stepped in the usual fashion (Fig. 2).

Infrared spectral measurements normally require that the radiation be chopped. One can then realize codes with 0's and 1's by chopping between the transmitted radiation and a standard source. The codes with $+1$'s and $-1$'s are realized by chopping between transmitted and reflected radiation.

There is one note of caution when using the $H^T$ matrix. Since all elements of the first column are $+1$'s, all radiation incident upon the first mask will be reflected to the detector. The remaining $N - 1$ masks will reflect only about half the radiation. As a result, a strain is placed upon the dynamic range of the instrument and nonlinear effects may arise in the detector. One solution to the problem (which would require a modification of the analysis of Sec. III) is to make the first mask only half as reflecting as the remaining $N - 1$ masks. This scheme, however, does not allow one to generate the rest of the masks by the simple stepping procedure mentioned previously. Other schemes also suggest themselves.

Partial support for this work came from a NASA contract and from a NSF grant.

References