# AN ALGORITHM FOR SIMULTANEOUS ORTHOGONAL TRANSFORMATION OF SEVERAL POSITIVE DEFINITE MATRICES TO NEARLY DIAGONAL FORM

by

Bernhard Flury*
Department of Statistics
Purdue University

Walter Gautschi
Department of Computer Sciences
Purdue University

Technical Report #84-1

Department of Statistics
Purdue University

January 1984

*The first author's work was done under contract #82.008.0.82 of the Swiss National Science Foundation.

*ABSTRACT*

For $k \geq 1$ positive definite symmetric matrices $A_1, \ldots, A_k$ of dimension $p \times p$ we define the function $\Phi(A_1, \ldots, A_k; n_1, \ldots, n_k) = \prod_{i=1}^{k} [\det(\operatorname{diag} A_i)]^{n_i} / [\det(A_i)]^{n_i}$, where $n_i$ are positive constants, as a measure of simultaneous deviation of $A_1, \ldots, A_k$ from diagonality. We give an iterative algorithm, called FG-algorithm, to find an orthogonal $p \times p$-matrix $B$ such that $\Phi(B^T A_1 B, \ldots, B^T A_k B; n_1, \ldots, n_k)$ is minimum. The matrix $B$ is said to transform $A_1, \ldots, A_k$ simultaneously to nearly diagonal form. Conditions for the uniqueness of the solution are given.

The FG-algorithm can be used to find the maximum likelihood estimates of common principal components in $k$ groups (Flury 1983b). For $k=1$, the FG-algorithm computes the characteristic vectors of the single positive definite symmetric matrix $A_1$.

**Keywords**: diagonalization, principal components, eigenvectors

## 1. THE PROBLEM

It is well known (see, e.g., Basilevsky (1983), Section 5.3) that if $A$ is a positive definite symmetric (p.d.s.) matrix of dimension $p \times p$, then there exists a real orthogonal matrix $B$ such that

$$B^T A B = \Lambda = \text{diag}\,(\lambda_1, \ldots, \lambda_p)\,, \tag{1.1}$$

where the $\lambda_i$ are all positive. For $k > 1$ p.d.s. matrices $A_1, \ldots, A_k$ the associated orthogonal matrices are in general different. We call $A_1, \ldots, A_k$ simultaneously diagonalizable if there exists an orthogonal matrix $B$ such that

$$B^T A_i B = \Lambda_i \text{ (diagonal) for } i = 1, \ldots, k\,. \tag{1.2}$$

Conditions equivalent to (1.2) have been given by Flury (1983a).

Now suppose that $A_1, \ldots, A_k$ are not simultaneously diagonalizable, but we wish to find an orthogonal matrix $B$ which makes them simultaneously "as diagonal as possible" in a sense to be defined. As a simple measure of "deviation from diagonality" of a p.d.s. matrix $F$ we can take

$$\varphi(F) = |\text{diag } F| \,/\, |F|\,, \tag{1.3}$$

where $|\cdot|$ is the determinant and diag $F$ is the diagonal matrix having the same diagonal elements as $F$. The fact that $\varphi$ is a reasonable measure of deviation from diagonality can be seen from Hadamard's inequality (Noble and Daniel (1977), exercise 11.51):

$$|F| \leq |\text{diag}\, F| \tag{1.4}$$

with equality exactly if $F$ is diagonal. Therefore, $\varphi(F) \geq 1$ holds, with equality exactly when $F$ is diagonal. Actually, $\varphi(G)$ increases monotonically as $G$ is continuously "inflated" from diag$F$ to $F$. This can be seen from the following lemma.

**Lemma 1**: *If $F = (f_{ij})$ is a p.d.s. matrix of dimension $p \times p$, then*

$$d(\alpha) := \det \begin{pmatrix} f_{11} & \alpha f_{12} & \cdots & \alpha f_{1p} \\ \alpha f_{21} & f_{22} & \cdots & \alpha f_{2p} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \alpha f_{p1} & \alpha f_{p2} & \cdots & f_{pp} \end{pmatrix} \tag{1.5}$$

***is a decreasing function of $\alpha$ for $\alpha \in [0,1]$. If $F$ is not diagonal, $d(\alpha)$ is strictly decreasing.***

**Proof**: The case $F = \operatorname{diag}(f_{11}, \ldots, f_{pp})$ is trivial; assume therefore that $F$ is not diagonal. Write

$$\begin{aligned} d(\alpha) &= |\alpha F + (1-\alpha)\operatorname{diag} F| \\ &= |\operatorname{diag} F| \cdot |\alpha(\operatorname{diag} F)^{-\frac{1}{2}} F (\operatorname{diag} F)^{-\frac{1}{2}} + (1-\alpha) I_p| \end{aligned} \tag{1.6}$$

and note that $d(\alpha)>0$ for all $\alpha \in [0,1]$, since both $F$ and $\operatorname{diag} F$ are p.d.s. Let $R = (\operatorname{diag} F)^{-\frac{1}{2}} F (\operatorname{diag} F)^{-\frac{1}{2}}$. $R$ is p.d.s. with 1's on the main diagonal. Let $d_1(\alpha) = |\alpha R + (1-\alpha) I_p|$. Then $d_1(0) = 1$ and $d_1(1) < 1$ by Hadamard's inequality. It remains to show that $d_1(\alpha)$ is strictly decreasing in $(0,1)$. Let $\rho_1 \geq \rho_2 \geq \cdots \geq \rho_p > 0$ denote the eigenvalues of $R$. The eigenvalues of $\alpha R + (1-\alpha) I_p$ are

$$\gamma_j = \alpha \rho_j + 1 - \alpha \quad (j = 1, \ldots, p), \tag{1.7}$$

and therefore

$$d_1(\alpha) = \prod_{j=1}^{p} \gamma_j = \prod_{j=1}^{p} (1 + \alpha(\rho_j - 1)). \tag{1.8}$$

Taking the first derivative gives

$$\begin{aligned} \frac{\partial d_1}{\partial \alpha} &= \sum_{h=1}^{p} (\rho_h - 1) \prod_{\substack{j=1 \\ j \neq h}}^{p} [1 + \alpha(\rho_j - 1)] \\ &= d_1(\alpha) \cdot \sum_{j=1}^{p} \frac{\rho_j - 1}{1 + \alpha(\rho_j - 1)}, \end{aligned} \tag{1.9}$$

where all denominators are positive because of $\rho_j > 0$ and $\alpha \leq 1$. Letting

$$d_2(\alpha) = \sum_{j=1}^{p} \frac{\rho_j - 1}{1+\alpha(\rho_j - 1)}, \tag{1.10}$$

we note that $d_2(0) = \sum_{j=1}^{p} (\rho_j - 1) = \text{tr } R - p = 0$ and

$$\frac{\partial d_2}{\partial \alpha} = - \sum_{j=1}^{p} \frac{(\rho_j - 1)^2}{(1+\alpha(\rho_j - 1))^2} < 0. \tag{1.11}$$

Therefore, $d_2(\alpha)<0$ on $(0,1]$, implying that $\frac{\partial d_1}{\partial \alpha} < 0$ for $0 < \alpha \leq 1$. This proves the lemma.

The reader may notice a similarity to ridge regression: Hoerl and Kennard (1970, Theorems 4.1 and 4.2) have given monotonicity properties of some functions related to the trace of the matrix $(F + \alpha I_p)^{-1}$ for $\alpha>0$.

Let us now consider $k$ p.d.s. matrices $F_1, \ldots, F_k$ and positive weights $n_1, \ldots, n_k$. Then we define the simultaneous deviation from diagonality of the matrices $F_1, \ldots, F_k$ with given weights $n_1, \ldots, n_k$ as

$$\Phi(F_1, \ldots, F_k; n_1, \ldots, n_k) = \prod_{i=1}^{k} [\varphi(F_i)]^{n_i}. \tag{1.12}$$

Let now $F_i = B^T A_i B$ $(i=1, \ldots, k)$ for a given orthogonal matrix $B$. Then we can take

$$\Phi_0(A_1, \ldots, A_k; n_1, \ldots, n_k) = \min_{B \in O(p)} \Phi(B^T A_1 B, \ldots, B^T A_k B; n_1, \ldots, n_k) \tag{1.13}$$

where $O(p)$ is the group of orthogonal $p \times p$-matrices, as a measure of simultaneous diagonalizability of $A_1, \ldots, A_k$. Clearly, $\Phi_0 \geq 1$ holds, with equality if and only if (1.2) is satisfied.

It can be shown (Flury 1983b) that if the minimum $\Phi_0$ is attained for a matrix $B_0 = (b_1, \ldots, b_p) \in O(p)$, then the following system of equations holds:

$$b_l^T \left( \sum_{i=1}^{k} n_i \frac{\lambda_{il} - \lambda_{ij}}{\lambda_{il}\lambda_{ij}} A_i \right) b_j = 0 \qquad (l,j=1, \ldots, p;\ l \neq j) \tag{1.14}$$

where

$$\lambda_{ih} = b_h^T A_i b_h \qquad (i=1, \ldots, k;\ h=1, \ldots, p). \tag{1.15}$$

In this paper we give an algorithm for finding $B_0$.

## 2. THE FG-ALGORITHM

The FG-algorithm consists of two algorithms, called $F$ and $G$ respectively, which minimize $\Phi$ by iteration on two levels:

- On the outer level (F-level), every pair $(b_l, b_j)$ of column vectors of the current approximation $B$ to the solution $B_0$ is rotated such that the corresponding equation in (1.14) is satisfied. One iteration step of the F-algorithm consists of rotations of all $p(p-1)/2$ pairs of vectors of $B$. The F-algorithm is similar to algorithms used in factor analysis to perform varimax and other rotations (see, e.g., Weber (1974)).
- On the inner level (G-level), an orthogonal 2×2-matrix $Q$ which solves a two-dimensional analog of (1.14) is found by iteration. This matrix defines the rotation of a pair of vectors currently being used on the F-level.

### The F-algorithm

Let

$$\Phi(B) = \Phi(B^T A_1 B, \ldots, B^T A_k B; n_1, \ldots, n_k) \tag{2.1}$$

denote the simultaneous deviation of $B^T A_1 B, \ldots, B^T A_k B$ from diagonality as a function of $B$, the $A_i$ and $n_i$ being considered as fixed. The F-algorithm yields a converging sequence of orthogonal matrices $B^{(0)}, B^{(1)}, \ldots$, such that $\Phi(B^{(f+1)}) \leq \Phi(B^{(f)})$.

The algorithm proceeds as follows:

*step* $F_0$: Define $B = (b_1, \ldots, b_p) \in O(p)$ as an initial approximation to the orthogonal matrix minimizing $\Phi$, e.g. $B \leftarrow I_p$. Put $f \leftarrow 0$.

*step* $F_1$: Put $B^{(f)} \leftarrow B$ and $f \leftarrow f+1$

*step* $F_2$: Repeat steps $F_{21}$ to $F_{24}$ for all pairs $(l,j)$, $1 \leq l < j \leq p$:

*step* $F_{21}$: Put $H(p \times 2) \leftarrow (b_l, b_j)$ and

$$T_i(2\times 2) \leftarrow \begin{pmatrix} b_l^T A_i b_l & b_l^T A_i b_j \\ b_j^T A_i b_l & b_j^T A_i b_j \end{pmatrix} \quad (i = 1, \ldots, k).$$

The $T_i$ are p.d.s.

*step* $F_{22}$: Perform the G-algorithm on $(T_1, \ldots, T_k)$ to get an orthogonal

$$2\times2\text{–matrix } Q = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

*step* $F_{23}$: Put $H^*(p\times2) = (b^*_l, b^*_j) \leftarrow HQ$. (This corresponds to an orthogonal rotation of the two columns of $H$ by an angle $\alpha$).

*step* $F_{24}$: In the matrix $B$, replace columns $b_l$ and $b_j$ by $b^*_l$ and $b^*_j$, respectively, and call the new matrix again $B$.

*step* $F_3$: If, for some small $\varepsilon_F>0$, $\Phi(B^{(f-1)}) - \Phi(B) < \varepsilon_F$ holds, stop. Otherwise, start the next iteration step at $F_1$.

**The G-algorithm**

This algorithm solves the equation

$$q_1^T \left[ n_1 \frac{\delta_{11}-\delta_{12}}{\delta_{11}\delta_{12}} T_1 + \cdots + n_k \frac{\delta_{k1}-\delta_{k2}}{\delta_{k1}\delta_{k2}} T_k \right] q_2 = 0 , \qquad (2.2)$$

where $T_1, \ldots, T_k$ are fixed p.d.s. 2×2-matrices, $n_i>0$ are fixed constants,

$$\delta_{ij} = q_j^T T_i q_j \quad (i=1,\ldots,k;\ j=1,2) , \qquad (2.3)$$

and $Q = (q_1, q_2)$ is an orthogonal 2×2-matrix. The iteration of the algorithm yields a sequence of orthogonal matrices $Q^{(0)}, Q^{(1)}, \ldots$, converging to a solution of (2.2).

The algorithm proceeds as follows:

*step* $G_0$: Define $Q$ (2×2) as an initial approximation to the solution of (2.2), e.g. $Q \leftarrow I_2$. Put $g \leftarrow 0$.

*step* $G_1$: Put $Q^{(g)} \leftarrow Q$ and $g \leftarrow g+1$.

*step* $G_2$: Compute the $\delta_{ij}$ (2.3), using the current $Q$.
Put $T$ (2×2) $\leftarrow n_1 \frac{\delta_{11}-\delta_{12}}{\delta_{11}\delta_{12}} T_1 + \cdots + n_k \frac{\delta_{k1}-\delta_{k2}}{\delta_{k1}\delta_{k2}} T_k$ .

*step* $G_3$: Compute the (normalized) eigenvectors of $T$. In $Q = (q_1, q_2)$, put $q_1 \leftarrow$ first eigenvector of $T$, $q_2 \leftarrow$ second eigenvector of $T$.

*step* $G_4$: If $||Q^{(g-1)} - Q|| < \varepsilon_G$ (where $||\cdot||$ denotes a matrix norm and $\varepsilon_G>0$ is a small positive constant), stop. Otherwise, start the next iteration step at $G_1$. Note that, since the order of eigenvectors is arbitrary, as well as their signs, it may be necessary to exchange $q_1$ and $q_2$ and/or to multiply one or both columns of $Q$ by $-1$ before comparing $Q$ with $Q^{(g-1)}$.

The motivation for the two algorithms and their connection with the basic system of equations (1.14) is as follows. Suppose that the $(l,j)$-th equation of (1.14) is to be solved. With $H = (b_l : b_j)$ denoting the current $l$-th and $j$-th columns of $B$, and $\lambda_{ih}$ being defined as in (1.15), $b_l$ and $b_j$ are the desired solution if and only if the 2×2-matrix

$$\sum_{i=1}^{k} n_i \frac{\lambda_{il} - \lambda_{ij}}{\lambda_{il}\lambda_{ij}} T_i \tag{2.4}$$

is diagonal, where

$$T_i = H^T A_i H \quad (i=1, \ldots, k). \tag{2.5}$$

Assume now that $b_l$ and $b_j$ do not solve the $(l,j)$-th equation, but $b^*_l = Hq_1$ and $b^*_j = Hq_2$ do, where $Q = (q_1 : q_2)$ is an orthogonal 2×2 matrix. Then

$$b^{*T}_l \Big[ \sum_{i=1}^{k} n_i \frac{\lambda^*_{il} - \lambda^*_{ij}}{\lambda^*_{il}\lambda^*_{ij}} A_i \Big] b^*_j = 0, \tag{2.6}$$

where

$$\lambda^*_{ih} = b^{*T}_h A_i b^*_h \quad (i=1, \ldots, k;\ h=l,j). \tag{2.7}$$

Putting $H^* = (b^*_l : b^*_j) = HQ$, (2.6) holds precisely if

$$\sum_{i=1}^{k} n_i \frac{\lambda^*_{il} - \lambda^*_{ij}}{\lambda^*_{il}\lambda^*_{ij}} H^{*T} A_i H^* \tag{2.8}$$

is diagonal. Now we note that

$$\begin{aligned} H^{*T} A_i H^* &= (HQ)^T A_i (HQ) \\ &= Q^T T_i Q, \end{aligned} \tag{2.9}$$

$$\begin{aligned} \lambda^*_{il} &= (Hq_1)^T A_i (Hq_1) \\ &= q_1^T T_i q_1 \quad (i=1,\ldots,k) \end{aligned} \tag{2.10}$$

and

$$\lambda^*_{ij} = q_2^T T_i q_2 \quad (i=1,\ldots,k). \tag{2.11}$$

Thus the problem of rotating the $l$-th and $j$-th columns of $B$ so as to satisfy

(1.14) can be reduced completely to the problem of finding an orthogonal 2×2-matrix $Q = (q_1:q_2)$ such that

$$q_1^T \left[ \sum_{i=1}^{k} n_i \frac{\delta_{i1}-\delta_{i2}}{\delta_{i1}\delta_{i2}} T_i \right] q_2 = 0, \tag{2.12}$$

where $\delta_{i1}$ and $\delta_{i2}$ have been written in place of $\lambda^*_{il}$ and $\lambda^*_{ij}$, respectively.

Since (2.12) is a 2-dimensional analog of (1.14), and since the group of orthogonal 2×2-matrices is compact, it follows that a solution of (2.12) always exists.

The problem of solving (2.12) is itself nontrivial. Although (2.12) can be written in terms of a rotation angle $\alpha$, solving for $\alpha$ would involve solving a polynomial equation of degree $4k$ in $\cos\alpha$ and $\sin\alpha$, which seems rather tedious. A more elegant solution is provided by the G-algorithm, which is based on the observation that the vectors $q_1$, $q_2$ satisfying (2.12) are eigenvectors of the matrix in brackets. Since the latter, however, depends also on $q_1$, $q_2$, an iterative procedure is required.

## 3. CONVERGENCE OF THE FG-ALGORITHM

### 3.1 Convergence of the F-algorithm

We show that the F-algorithm, in theory (i.e. if $\varepsilon_F=\varepsilon_G=0$), does not stop unless the equations (1.14) are satisfied for the current $B$, and that, if $B$ does not satisfy (1.14), an iteration step of the F-algorithm will decrease $\Phi$.

Suppose that the current orthogonal matrix $B = (b_1, \ldots, b_p)$ does not satisfy the $(l,j)$-th equation of (1.14). For notational simplicity, we can take $l=1$ and $j=2$ without loss of generality. Let us write $B = (B^{(1)}:B^{(2)})$, where $B^{(1)} = (b_1,b_2)$. In step $F_{21}$, the matrices $T_i = B^{(1)T}A_iB^{(1)}$ are passed to the G-algorithm. The G-algorithm gives back an orthogonal 2×2 matrix $Q$ (step $F_{22}$). (Note that $Q$ is not necessarily unique, depending upon the conventions used in the G-algorithm. We will consider every matrix $\bar{Q}$ obtained from $Q$ by interchanging the columns of $Q$ and/or multiplying one or both columns by $-1$ as *equivalent* to $Q$). Steps $F_{23}$ and $F_{24}$ correspond to the transformation

$$B^* = B \begin{pmatrix} Q & 0 \\ 0 & I_{p-2} \end{pmatrix} = (B^{(1)}Q:B^{(2)}) . \tag{3.1}$$

$B^*$ is orthogonal, since it is the product of two orthogonal matrices. Now we have

$$\Phi(B^*) = \prod_{i=1}^{k} [\,|\text{diag } B^{*T}A_iB^*| \,/\, |B^{*T}A_iB^*|\,]^{n_i}$$

$$= \prod_{i=1}^{k} [\,|\text{diag } Q^TB^{(1)^T}A_iB^{(1)}Q| \cdot |\text{diag } B^{(2)^T}A_iB^{(2)}| \,/\, |A_i|\,]^{n_i} . \qquad (3.2)$$

It will be shown in section 3.2 that if, as assumed, (1.14) is not satisfied for $l=1$ and $j=2$, then

$$\prod_{i=1}^{k} |\text{diag } Q^TB^{(1)^T}A_iB^{(1)}Q|^{n_i} < \prod_{i=1}^{k} |\text{diag } B^{(1)^T}A_iB^{(1)}|^{n_i} . \qquad (3.3)$$

If (1.14) is satisfied, $Q$ will be equivalent to $I_2$, and hence (3.3) holds with equality. Therefore, each iteration step of the F-algorithm decreases $\Phi$, and the algorithm will stop only if (1.14) is satisfied.

### 3.2 Convergence of the G-algorithm

The G-algorithm starts with $k$ p.d.s. 2×2-matrices $T_1, \ldots, T_k$ and $k$ weights $n_1, \ldots, n_k > 0$. Let $Q^{(g)}$ denote the orthogonal 2×2 matrix after the $g$-th iteration. We will show that

$$\prod_{i=1}^{k} |\text{diag } Q^{(g+1)^T}T_iQ^{(g+1)}|^{n_i} \le \prod_{i=1}^{k} |\text{diag } Q^{(g)^T}T_iQ^{(g)}|^{n_i} , \qquad (3.4)$$

and that the sequence $Q^{(g)}$ converges to an orthogonal matrix which solves (2.2).

Suppose now that the $(g+1)$-st iteration of the G-algorithm is being performed. It is somewhat simpler to prove the convergence if we introduce the following notation: Let $Q = (q_1, q_2)$ contain the current approximation to the solution of (2.2) and $\delta_{ij}$ be defined as in (2.3). Then we put

$$a_i = \frac{\delta_{i1}-\delta_{i2}}{\delta_{i1}\delta_{i2}} \quad (i=1, \ldots, k) , \qquad (3.5)$$

$$T = \sum_{i=1}^{k} n_i a_i T_i , \qquad (3.6)$$

and

$$U_i = Q^TT_iQ = \begin{bmatrix} u_{11}^{(i)} & u_{12}^{(i)} \\ u_{21}^{(i)} & u_{22}^{(i)} \end{bmatrix} . \qquad (3.7)$$

The $U_i$ are p.d.s., and clearly

$$T = \sum_{i=1}^{k} n_i a_i Q U_i Q^T \tag{3.8}$$

and

$$\delta_{i1} = u_{11}^{(i)}, \ \delta_{i2} = u_{22}^{(i)} . \tag{3.9}$$

In step $G_3$ the characteristic vectors of $T$ are computed. Denote the solution by $Q^*$, so that

$$Q^{*T} T Q^* = \Lambda \tag{3.10}$$

is diagonal. From (3.8) it follows that

$$\sum_{i=1}^{k} n_i a_i Q^{*T} Q U_i Q^T Q^* = \Lambda . \tag{3.11}$$

The characteristic vectors of the symmetric matrix

$$U = \sum_{i=1}^{k} n_i a_i U_i \tag{3.12}$$

are therefore given by the orthogonal matrix

$$P = Q^T Q^* , \tag{3.13}$$

and $Q^*$ can therefore be obtained by

$$Q^* = QP . \tag{3.14}$$

Note that $U = Q^T T Q$ is diagonal if and only if $Q = (q_1, q_2)$ is a solution of (2.2). From (3.9) and (3.12) it follows that

$$U = \sum_{i=1}^{k} n_i \frac{u_{11}^{(i)} - u_{22}^{(i)}}{u_{11}^{(i)} u_{22}^{(i)}} \begin{bmatrix} u_{11}^{(i)} & u_{12}^{(i)} \\ u_{21}^{(i)} & u_{22}^{(i)} \end{bmatrix} . \tag{3.15}$$

Let

$$\vartheta_i = \begin{cases} 1 & \text{if } u_{11}^{(i)} > u_{22}^{(i)} \\ -1 & \text{if } u_{11}^{(i)} < u_{22}^{(i)} \\ 0 & \text{if } u_{11}^{(i)} = u_{22}^{(i)} , \end{cases} \tag{3.16}$$

$$a'_i = \vartheta_i a_i , \tag{3.17}$$

and

$$S_i = \begin{pmatrix} s_{11}^{(i)} & s_{12}^{(i)} \\ s_{21}^{(i)} & s_{22}^{(i)} \end{pmatrix} = a'_i U_i \quad (i=1, \ldots, k) . \tag{3.18}$$

$S_i$ is p.d.s., unless $\vartheta_i = 0$. With this notation, we have

$$U = \sum_{i=1}^{k} n_i \vartheta_i S_i . \tag{3.19}$$

Now let $k' \leq k$ denote the number of $\vartheta_i$'s which are not zero. Without loss of generality assume that the $k-k'$ matrices $S_i$ which are zero have the indices $k'+1, \ldots, k$. Therefore the sum (3.19) extends only up to $k'$, and we are going to show that

$$\prod_{i=1}^{k'} |\text{diag } P^T S_i P|^{n_i} \leq \prod_{i=1}^{k'} (s_{11}^{(i)} s_{22}^{(i)})^{n_i} , \tag{3.20}$$

with equality if and only if $U$ is diagonal. Assuming for the moment that (3.20) holds true, the proof of (3.4) can be completed by noting that (3.20) implies

$$\prod_{i=1}^{k'} (a_i^2 \vartheta_i^2 |\text{diag } P^T U_i P|)^{n_i} \leq \prod_{i=1}^{k'} (a_i^2 \vartheta_i^2 u_{11}^{(i)} u_{22}^{(i)})^{n_i} \tag{3.21}$$

and therefore

$$\prod_{i=1}^{k'} |\text{diag } Q^{*T} T_i Q^*|^{n_i} \leq \prod_{i=1}^{k'} |\text{diag } Q^T T_i Q|^{n_i} . \tag{3.22}$$

For the remaining $k-k'$ matrices $U_i$ $(i=k'+1, \ldots, k)$ we have $u_{11}^{(i)} = u_{22}^{(i)}$, and therefore, as is easily verified,

$$|\text{diag } B^T U_i B| \leq |\text{diag } U_i| \tag{3.23}$$

for any $B \in O(2)$, with equality exactly if $B$ is equivalent to $I_2$ or $u_{12}^{(i)} = 0$. This holds, in particular, for $B=P$. Putting (3.22) and (3.23) together gives now the

desired result (3.4). It remains to show (3.20).

Let $P=(p_1,p_2)$ denote the eigenvectors of $U = \sum_{i=1}^{k'} \vartheta_i n_i S_i$, with $p_1$ being associated with the algebraically larger root. Since $U$ is symmetric, $P$ is orthogonal (or can be so chosen if the two roots are identical), and both characteristic roots are real. Assume that $U$ is not diagonal, and let $\varepsilon_i$ $(i=1, \ldots, k')$ be defined by

$$P^T S_i P = \begin{bmatrix} s_{11}^{(i)} + \vartheta_i \varepsilon_i & \cdot \\ \cdot & s_{22}^{(i)} - \vartheta_i \varepsilon_i \end{bmatrix}. \tag{3.24}$$

From (3.5), (3.9) and (3.16) to (3.18) we have

$$s_{11}^{(i)} s_{22}^{(i)} = \vartheta_i (s_{11}^{(i)} - s_{22}^{(i)}), \quad \text{or} \quad 1 = \vartheta_i \left( \frac{1}{s_{22}^{(i)}} - \frac{1}{s_{11}^{(i)}} \right) \quad (i=1, \ldots, k'), \tag{3.25}$$

which implies that either $s_{11}^{(i)}$ or $s_{22}^{(i)}$ is smaller than 1. It then follows that $\varepsilon_i < 1$ $(i=1, \ldots, k')$. Indeed, if $\vartheta_i = 1$, then $s_{22}^{(i)} < 1$ by (3.25), and the positivity of $s_{22}^{(i)} - \varepsilon_i$ implies $\varepsilon_i < 1$. If $\vartheta_i = -1$, then $s_{11}^{(i)} < 1$, and $s_{11}^{(i)} - \varepsilon_i > 0$ implies again $\varepsilon_i < 1$.

The product of the diagonal elements of $P^T S_i P$ is

$$\begin{aligned} |\text{diag } P^T S_i P| &= (s_{11}^{(i)} + \vartheta_i \varepsilon_i)(s_{22}^{(i)} - \vartheta_i \varepsilon_i) \\ &= s_{11}^{(i)} s_{22}^{(i)} - \varepsilon_i \vartheta_i (s_{11}^{(i)} - s_{22}^{(i)}) - \varepsilon_i^2 \\ &= (1-\varepsilon_i) s_{11}^{(i)} s_{22}^{(i)} - \varepsilon_i^2 \\ &\leq (1-\varepsilon_i) s_{11}^{(i)} s_{22}^{(i)} \quad (i=1, \ldots, k'). \end{aligned} \tag{3.26}$$

Thus,

$$\prod_{i=1}^{k'} |\text{diag } P^T S_i P|^{n_i} \leq \left[ \prod_{i=1}^{k'} (1-\varepsilon_i)^{n_i} \right] \left[ \prod_{i=1}^{k'} |\text{diag } S_i|^{n_i} \right], \tag{3.27}$$

and (3.20) holds if we can prove that

$$\prod_{i=1}^{k'} (1-\varepsilon_i)^{n_i} < 1. \tag{3.28}$$

To demonstrate this, we note that, since $U$ is assumed not diagonal,

$$p_1^T U p_1 > u_{11}, \tag{3.29}$$

or equivalently,

$$\begin{aligned} &\sum_{i=1}^{k'} \vartheta_i n_i p_1^T S_i p_1 > \sum_{i=1}^{k'} \vartheta_i n_i s_{11}^{(i)} , \\ &\sum_{i=1}^{k'} \vartheta_i n_i (p_1^T S_i p_1 - s_{11}^{(i)}) > 0 . \end{aligned} \tag{3.30}$$

Since $p_1^T S_i p_1 - s_{11}^{(i)} = \vartheta_i \varepsilon_i$ $(i=1, \ldots, k')$, this implies

$$\sum_{i=1}^{k'} n_i \varepsilon_i > 0 , \tag{3.31}$$

so that not all $\varepsilon_i$ can be zero. On the other hand, if $U$ is diagonal, then $P$ is equivalent to $I_2$, and all $\varepsilon_i$ are zero. Therefore the $\varepsilon_i$ vanish simultaneously if and only if $U$ is diagonal. Now we need the following lemma.

**Lemma 2**: *If* $x_i>0$, $n_i>0$ $(i=1, \ldots, k')$ *and* $\sum_{i=1}^{k'} n_i x_i \le \sum_{i=1}^{k'} n_i$, *then* $\prod_{i=1}^{k'} x_i^{n_i} \le 1$.

**Proof**: Maximize the function $\prod_{i=1}^{k'} x_i^{n_i}$ under the restriction $\sum_{i=1}^{k'} n_i x_i = g$ $(>0)$, using a Lagrange multiplier. The maximum has the value $(g/n)^n$ and is attained for $x_1 = \cdots = x_{k'} = g/n$, where $n = \sum_{i=1}^{k'} n_i$. Noting that $g \le n$ completes the proof.

Since $\varepsilon_i<1$ $(i=1, \ldots, k')$ and $\sum_{i=1}^{k'} n_i \varepsilon_i > 0$, we can use Lemma 2 with $x_i = 1-\varepsilon_i$ and get (3.28). Note that equality in (3.28) holds exactly if all $\varepsilon_i$ are zero. This completes the proof of convergence of the G-algorithm.

## 4. CONDITIONS FOR UNIQUENESS OF THE SOLUTION

In Section 3 we have shown that the FG-algorithm converges to a minimum of (2.1), unless the initial approximation of the orthogonal matrix $B$ is (badly) chosen as a stationary point of $\Phi$. However, we do not know whether $\Phi$ has a unique minimum. We are now going to show that in some "extreme" cases there exist more than one local minimum, and we give approximate conditions when this will happen. Throughout this section (unless otherwise stated) we will only consider the case $k=2$ and $p=2$.

Let the p.d.s. matrix $S_1$ have the characteristic roots $l_1>l_2$ (the case $l_1=l_2$ being trivial), and assume, for simplicity, that

$$S_1 = \begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix}. \tag{4.1}$$

From (3.5) it can be seen that the solutions of (2.2) are unaffected by proportionality, i.e., we can assume (see also (3.25))

$$l_1 - l_2 = l_1 l_2 \tag{4.2}$$

without loss of generality. Consider now an orthogonal matrix

$$B = B(\varphi) = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}. \tag{4.3}$$

The product of the diagonal elements of $B^T S_1 B$ is

$$\begin{aligned} |\mathrm{diag}\,(B^T S_1 B)| &= [l_2+(l_1-l_2)\cos^2\varphi][l_1-(l_1-l_2)\cos^2\varphi] \\ &= l_1 l_2 + (l_1-l_2)^2\cos^2\varphi\,\sin^2\varphi \\ &= l_1 l_2[1 + l_1 l_2\cos^2\varphi\,\sin^2\varphi] \\ &= r_1[1 + r_1\cos^2\varphi\,\sin^2\varphi]\,, \end{aligned} \tag{4.4}$$

where

$$r_1 = l_1 l_2 \tag{4.5}$$

denotes the product of the characteristic roots of $S_1$. Let the excentricity $d_1$ of $S_1$ be defined as the ratio of the larger to the smaller root of $S_1$,

$$d_1 = l_1 / l_2\,, \tag{4.6}$$

which is also the Euclidean condition number of $S_1$. From (4.2) it follows that $l_2 = l_1/(l_1+1)$, and therefore $d_1 = l_1+1$. Similarly, $d_1 = 1/(1-l_2)$, and therefore

$$\begin{aligned} l_1 &= d_1-1\,, \\ l_2 &= (d_1-1)/d_1\,. \end{aligned} \tag{4.7}$$

Multiplying these two equations gives

$$r_1 = (d_1-1)^2/d_1\,. \tag{4.8}$$

Note that $d_1$ does not depend on the absolute size of $S_1$ (every matrix proportional to $S_1$ has the same excentricity), and so the same is true for $r_1$.

For a second p.d.s. matrix $S_2$, let $d_2$ denote its excentricity, and

$$r_2 = (d_2-1)^2/\, d_2\,. \tag{4.9}$$

Let $B_0 = \begin{pmatrix} \cos\varphi_0 & -\sin\varphi_0 \\ \sin\varphi_0 & \cos\varphi_0 \end{pmatrix}$ denote the orthogonal matrix which diagonalizes $S_2$. Then, in analogy to (4.4), we get

$$|\operatorname{diag}(B^T S_2 B)| = r_2[1+r_2\cos^2(\varphi-\varphi_0)\,\sin^2(\varphi-\varphi_0)]\,. \tag{4.10}$$

The function $\Phi$ to be minimized is

$$\Phi(\varphi) = [1+r_1\cos^2\varphi\,\sin^2\varphi]^{n_1}\,[1+r_2\cos^2(\varphi-\varphi_0)\,\sin^2(\varphi-\varphi_0)]^{n_2}\,. \tag{4.11}$$

Let us now assume that $n_1=n_2$, so that it remains to minimize

$$G(\varphi) = [1+\tfrac{1}{4}r_1\sin^2(2\varphi)][1+\tfrac{1}{4}r_2\sin^2(2(\varphi-\varphi_0))]\,. \tag{4.12}$$

$G(\varphi)$ is $\pi/2$-periodic, and from (4.12) it becomes clear that for $\varphi_0\neq 0$, $G(\varphi)$ may have more than one local minimum in one period, depending on $r_1$, $r_2$ and $\varphi_0$ (and, in the general situation, on $n_1$ and $n_2$). Note that $\varphi_0$ is the minimum angle between two characteristic vectors of $S_1$ and $S_2$.

Let us first look at the extreme situation $\varphi_0=\pi/4$. From a Taylor expansion it can be seen that in a neighborhood of 0,

$$G(\varphi) = 1 + \tfrac{1}{4}r_2 + (r_1-r_2+\tfrac{1}{4}r_1r_2)\,\varphi^2 + O(\varphi^4)\,. \tag{4.13}$$

The function $G(\varphi)$ has therefore a stationary point at $\varphi=0$, which is a

$$\begin{aligned} &\text{minimum, if } r_1-r_2+\tfrac{1}{4}r_1r_2 > 0\,, \\ &\text{maximum, if } r_1-r_2+\tfrac{1}{4}r_1r_2 < 0\,. \end{aligned} \tag{4.14}$$

Note that for $r_1\geq 4$ or $r_1=r_2$ this is always a minimum.

Similarly, at $\varphi = \pi/4$, we get a

$$\begin{aligned} &\text{minimum, if } r_2 - r_1 + \frac{1}{4} r_1 r_2 > 0 , \\ &\text{maximum, if } r_2 - r_1 + \frac{1}{4} r_1 r_2 < 0 . \end{aligned} \tag{4.15}$$

For $r_2 \geq 4$ or $r_1 = r_2$ this is always a minimum.

Since $r_1$ and $r_2$ are both positive, there cannot be a maximum at 0 and $\pi/4$ simultaneously. Local minima at both points, however, are obtained e.g. if both $r_1$ and $r_2$ are larger than 4, or if $r_1 = r_2$ (even if $r_1 = r_2$ is very close to zero!). Thus the case of equal excentricity of both matrices seems most "dangerous" in terms of multiple local minima.

Using the relation $r_i = (d_i - 1)^2 / d_i$ (4.8, 4.9), the conditions (4.14) and (4.15) can be transformed to conditions on the excentricities $d_i$ ($i=1,2$). Figure 1 shows a partition of $[1, \infty) \times [1, \infty)$ into three areas in which a minimum is attained at 0 only, at $\pi/4$ only, or at both points, depending on the values of $d_1$ and $d_2$. Note that for $d_1 > 5.828427$ ($d_2 > 5.828427$) there is always a minimum at $\varphi = 0$ ($\varphi = \pi/4$), and if $d_1 = d_2$, there are always two minima.

Figure 1. *Conditions for minima and maxima if* $n_1 = n_2$, $\varphi_0 = \pi/4$.

The case $\varphi_0 = \pi/4$ treated so far is of course the "worst possible" case, since the minimum angle between two characteristic vectors of $S_1$ and $S_2$ cannot exceed $\pi/4$. For the application in common principal component analysis (Flury 1983b), however, we expect $\varphi_0$ rather close to zero, if the null hypothesis of identical principal components in the populations holds. Therefore we look now at the situation where $\varphi_0$ is close to zero. Without loss of generality we can assume $\varphi_0 > 0$. Again, for simplicity, we take $n_1 = n_2 = 1$.

Approximating the trigonometric functions in the two factors of $G$ by Taylor series at $\varphi=0$ (first factor) and at $\varphi=\varphi_0$ (second factor), and taking the first derivative of $G$ with respect to $\varphi$ yields

$$\begin{aligned} G'(\varphi) = {} & 2r_1[\varphi+O(\varphi^3)][1+r_2((\varphi-\varphi_0)^2 + O(\varphi-\varphi_0)^4)] \\ & + 2r_2[(\varphi-\varphi_0) + O(\varphi-\varphi_0)^3][1+r_1(\varphi^2+O(\varphi^4))] . \end{aligned} \tag{4.16}$$

If $\varphi_0$ is close to zero, sufficient accuracy can be had for $0 \le \varphi \le \varphi_0$ if we ignore all terms of order higher than 2. An approximation to the solution(s) of $G'(\varphi) = 0$ within $[0, \varphi_0]$ is therefore given by the solution(s) of

$$r_1\varphi[1+r_2(\varphi-\varphi_0)^2] + r_2(\varphi-\varphi_0)[1+r_1\varphi^2] = 0 . \tag{4.17}$$

This equation has either one or three real roots, depending on $r_1$, $r_2$ and $\varphi_0$. For $r_1 = r_2 = r$, (4.17) can be written as

$$\begin{aligned} & \varphi(1+r(\varphi-\varphi_0)^2) + (\varphi-\varphi_0)(1+r\varphi^2) \\ & = 2(\varphi-\frac{\varphi_0}{2})\,(r\varphi^2 - r\varphi_0\,\varphi+1) = 0 . \end{aligned} \tag{4.18}$$

Thus $\varphi=\varphi_0/2$ is a solution of (4.18) (and also of (4.16) if $r_1=r_2$). If, approximately,

$$\frac{4}{r} > \varphi_0^2 , \tag{4.19}$$

$G(\varphi)$ takes a minimum at $\varphi_0/2$. Under the same condition (4.19), the polynomial

$$r\,\varphi^2 - r\,\varphi_0\varphi+1 \tag{4.20}$$

has no real root, and the minimum at $\varphi_0/2$ is unique.

If, always approximately for small $\varphi_0$, $4/r<\varphi_0^2$, we get a maximum at $\varphi_0/2$, and two minima at

$$\frac{1}{2}(\varphi_0 \pm \sqrt{\varphi_0^2-4/r}) . \tag{4.21}$$

In terms of the excentricity parameters $d_1=d_2=d$, condition (4.19) becomes

$$\frac{4d}{(d-1)^2} > \varphi_0^2 , \tag{4.22}$$

which shows that two minima are to be expected only if the excentricity is high. For large $d$, (4.22) is approximately the same as

$$d < (\frac{2}{\varphi_0})^2 . \tag{4.23}$$

For example, if $\varphi_0=.2$ ($\approx 11.5$ degrees), a single minimum can be expected approximately if $d<100$.

Figure 2 shows the typical behavior of the function $G(\varphi)$ for $\varphi_0=.2$ and $r=160$ ($d=161.99$). The two minima are approximately at .039 and .161. If different values are chosen for $r_1$ and $r_2$, the two minima are in general not identical, but the shape of the graph is similar, with one "valley" being less deep than the other.

Figure 2. *Graph of* $G(\varphi)$ *for* $\varphi_0=.2$, $n_1=n_2=1$ *and* $d_1=d_2=162$.

Although these results are only approximate, they give a general idea about the conditions for uniqueness of the minimum. For $k>2$ matrices, the relations are of course more complicated, but still we can expect a unique minimum

unless some of the matrices are highly excentric.

For dimension $p>2$, the minimum is certainly unique if all the $p(p-1)/2$ equations (1.14) have a unique minimizing solution. (By a minimizing solution we mean a solution which corresponds to a local minimum of $G$, or, in the $p$-dimensional case, of the function $\Phi$.) On the other hand, if some of the equations have more than one minimizing solution, this does not necessarily imply that the whole system (1.14) has more than one minimizing solution.

A solution given by the FG-algorithm does of course not prove its uniqueness. However, Figure 2 suggests the following: If we start the FG-algorithm with $B(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ as an initial approximation, it will converge to the left minimum, while $B(\varphi_0) = \begin{pmatrix} \cos\varphi_0 & -\sin\varphi_0 \\ \sin\varphi_0 & \cos\varphi_0 \end{pmatrix}$ as an initial approximation leads to convergence to the right minimum. (This is indicated by the arrows in Figure 2). Minimizing solutions can always be expected to be somehow "close" to the characteristic vectors of one of the matrices. Therefore, if there is doubt about the uniqueness of the solution, it is recommended that one run the FG-algorithm $k$ times, using the $k$ sets of characteristic vectors of $S_1, \ldots, S_k$ as initial approximations. If all $k$ solutions found are equal, it is reasonable to assume that there is a unique global minimum.

As a numerical example, let $S_1 = \begin{pmatrix} 100 & 0 \\ 0 & 1 \end{pmatrix}$ and $S_2 = \begin{pmatrix} 96.0143 & 19.4603 \\ 19.4603 & 4.9857 \end{pmatrix}$, so that $d_1=d_2=100$ and $\varphi_0=.202$ ($\approx 11.57$ degrees), which is a borderline case according to approximation (4.22). $G(\varphi)$ assumes two minima at .08 and .12, approximately. If we reduce the excentricity to 90 (leaving $\varphi_0$ unchanged), we get the matrices $S_1 = \begin{pmatrix} 90 & 0 \\ 0 & 1 \end{pmatrix}$ and $S_2 = \begin{pmatrix} 86.4168 & 17.4946 \\ 17.4946 & 4.5831 \end{pmatrix}$. For these two matrices, there is a unique minimum at $\varphi_0/2$. The bound (4.22) for $d$ is in general too high, but the approximation becomes better when $\varphi_0$ gets smaller.

## 5. REMARKS

1. The proof of convergence of the G-algorithm makes strong use of the assumption that the matrices $T_i$ are positive definite. If one or several of the matrices $A_i$ are close to singularity, this could cause numerical problems, because the $a_i$ (3.5) might become very large.
2. Since the stopping rule given in step $F_3$ depends on the absolute size of the matrices $A_i$, it may be better to replace it by a criterion similar to the one used in the G-algorithm:

   $F_3$ : If $||B^{(j-1)}-B|| < \varepsilon_F$ for some small $\varepsilon_F>0$, stop. Otherwise,

start the next iteration step at $F_1$.

3. If the current version of $B$ in the F-algorithm is a stationary point of $\Phi$, and $I_2$ is taken as an initial approximation of $Q$ in the G-algorithm, FG will not change $B$, since (1.14) is satisfied. This occurs, e.g., if the diagonal elements of the $A_i$-matrices are identical for each $A_i$, that is, diag $A_i$=diag $(c_i, \ldots, c_i)$ for some $c_i>0$ $(i=1, \ldots, k)$, and $I_p$ is taken as an initial approximation of $B$. An important special case of this are correlation matrices, where the diagonal elements are all 1. If the first iteration of the F-algorithm does not change $B$, it might therefore be helpful to try FG with another initial approximation.

4. On the F-level, a better initial approximation than $I_p$ might be to take the eigenvectors of one of the $A_i$ (e.g. the one with the largest $n_i$), or the eigenvectors of $\sum_{i=1}^{k} n_i A_i$. On the G-level, $I_2$ is a good initial approximation for $Q$, when the current $B$ on the F-level is already close to the correct solution.

5. In step $F_{24}$, the $l$-th and $j$-th column of $B$ are adjusted using the matrix $Q$ given by the G-algorithm. Since these two columns will undergo changes in subsequent executions of steps $F_{21}$ to $F_{24}$, it is not necessary to iterate on the G-level until full convergence is reached. In most cases the first iteration steps of the G-algorithm will decrease $\Phi(B)$ much more than the later iterations. If $k=1$, only one iteration step is required in each execution of the G-algorithm.

6. In order to avoid permutations of the columns of $B$ and multiplications by $-1$, it is convenient to order the columns of $Q$ such that

$$Q = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \tag{5.1}$$

where $-\pi/2<\alpha<\pi/2$.

7. If $k=1$, the FG-algorithm computes the characteristic vectors of the single p.d.s. matrix $A=A_1$. Nothing is known, however, about the efficiency of the FG-algorithm compared with the usual algorithms used to compute characteristic vectors of a p.d.s. matrix.

8. The listing of a FORTRAN-program performing the FG-algorithm can be obtained from the first author upon request.

**REFERENCES**

BASILEVSKY, A. (1983), *Applied Matrix Algebra in the Statistical Sciences*, New York: North Holland.

FLURY, B.N. (1983a), "Some relations between the comparison of covariance matrices and principal component analysis", *Computational Statistics and Data Analysis*, 1, 97-109.

FLURY, B.N. (1983b), "A generalization of principal component analysis to $k$ groups", Technical Report No. 83-14, Dept. of Statistics, Purdue University.

HOERL, A.E. and KENNARD, R.W. (1970), "Ridge regression: Biased estimation for nonorthogonal problems", *Technometrics*, 12, 55-67.

NOBLE, B. and DANIEL, J.W. (1977), *Applied Linear Algebra*, Englewood Cliffs, N.J.: Prentice Hall.

WEBER, E. (1974), *Einführung in die Faktorenanalyse*, Stuttgart & New York: Verlag Gustav Fischer.

FIGURE 2

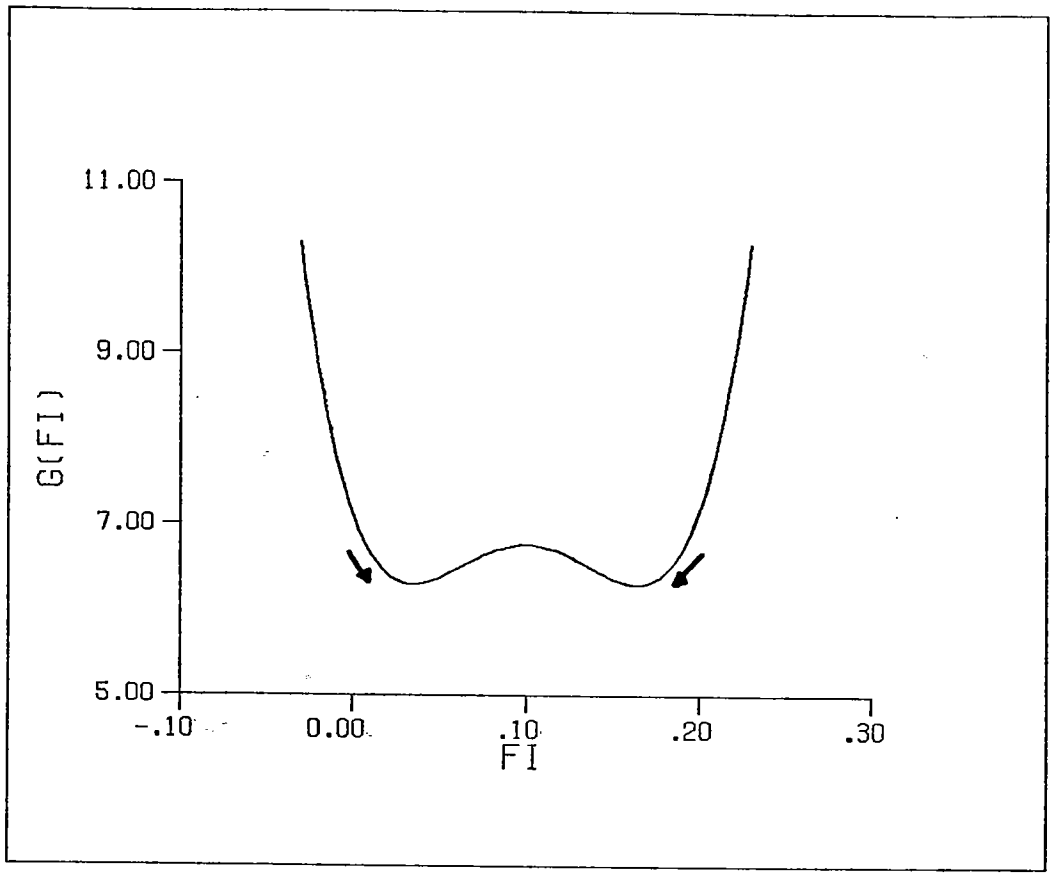


FIGURE 1