NOAA Technical Memorandum ERL PMEL-40

DATA INTERCOMPARISON THEORY

III. S-PHASE AND T-PHASE TESTS FOR SPATIAL PATTERN AND TEMPORAL EVOLUTION

Rudolph W. Preisendorfer
Curtis D. Mobley

Pacific Marine Environmental Laboratory
Seattle, Washington
December 1982

UNITED STATES
DEPARTMENT OF COMMERCE
Malcolm Baldrige,
Secretary

NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION
John V. Byrne,
Administrator

Environmental Research
Laboratories
George H. Ludwig
Director

NOTICE

Mention of a commercial company or product does not constitute an endorsement by NOAA Environmental Research Laboratories. Use for publicity or advertising purposes of information from this publication concerning proprietary products or the tests of such products is not authorized.

Contribution No. 602 from NOAA's Pacific Marine Environmental Laboratory

TABLE OF CONTENTS

Abstract

This is the third of a series of five reports on some new techniques in data intercomparison theory, devised particularly with observed and model climate data sets in mind. In this study we derive a set of statistics which describe differences in spatial patterns and temporal evolutions of data sets. These are derived from the SHAPE statistic of the second study. It is shown that SHAPE and its derivates are all of a correlational nature, thereby permitting immediate statistical significance tests using existing tables. However, it is also shown that correlation-type statistics, especially in multivariate settings, are limited in their ability to discern differences in spatial and temporal patterns of data sets. Two multiparameter tests (the S-Phase and T-Phase) are devised to remedy this limitation and their properties are given a preliminary examination. The S-Phase test is based on the canonic rotation angles between the spatial (eigenvector) frames of two data sets that have been given their singular value decomposition. The T-Phase test is based on the canonic correlation angles between the temporal (principal component) frames of two data sets under the same type of decomposition. These multiparameter tests, unfortunately, are difficult to interpret, even though they appear to perform well under realistic data conditions. Further multiparameter tests are devised and outlined with research programs suggested for future study. The five general procedures of the second study of the series (IOP, EOP, APP, PPP, CIP) are potential sources of reference distributions for the new multiparameter statistics.

Data Intercomparison Theory

III. S-Phase and T-Phase Tests for Spatial Pattern and Temporal Evolution

Rudolph W. Preisendorfer

Curtis D. Mobley

I. Introduction

We continue our development of data intercomparison methods begun in* DIT(II). In that study we began with a systematic analysis of the squared distance (DIST2) between two data sets. We saw how that squared distance could be additively split into three parts: one (namely SITES) describing the separation of the centroids or average location of the data sets, another (SPRED) describing the difference in radial spread of the data sets (thought of as swarms of points in a euclidean p-space), and finally (SHAPE), a part describing the difference in spatial and temporal configurations of the data sets in p-space. It is this third statistic, SHAPE, that we shall analyze into its elemental parts in the present study.

The statistic SHAPE is by far the richest of the three that we carved out of DIST2, for it contains within it the information about the difference in spatial patterns of the two data sets, and also differences in their temporal evolution. A preliminary examination in DIT(II) of the power of SHAPE, using random sampling from gaussian populations to simulate different pairs of data sets to be intercompared, showed that the power was relatively low compared to that of SITES and SPRED, under similar testing conditions. This is not surprising in view of the fact that SHAPE is the result of the collapse of the multivariate

---

* A list of titles in the present series is given in the reference section (§8), below. It will be assumed that the reader has access to DIT(II) in the developments that follow.

internal temporal and spatial structure of a data set down into a single number. We illustrate this information crunch by some examples in §4, and use that discussion to motivate some attempts at multiparameter tests (in §§5, 6) for spatial pattern and temporal evolution differences between data sets.

Besides attempting the construction of new multiparameter tests, we also devote some effort to analyzing SHAPE. By splitting apart the various pieces of information within a data set we can produce correspondingly simpler SHAPE-type statistics (such as ORIEN, COREL, DIAGS, e.g., in §2) that concentrate on one attribute at a time, namely space properties, time properties, or variance properties of a data set. This splitting apart of information is effected through the singular value decomposition of a data set. The resultant simpler statistics have somewhat higher power.

All of these SHAPE-derived statistics (including SHAPE itself) share a common statistical structure, namely that of a correlation. This fact endows these statistics with the advantage that they could be relatively simply applied using standard statistical correlation tests. But there are also disadvantages. First of all, the classical correlation test is designed for the usual gaussian-population sampling procedure. Under real-life sampling conditions, the ORIEN, COREL, and DIAGS statistics (for example), along with many other correlation-type statistics presently used in practice, can in principal have distinctly *non-classical* distributions (unless special transformations are devised), and so there could be abuses of the classical test. But this abuse seems to be common practice (so common that many practitioners go about the ritual of declaring a correlation significant without a moment's thought as to the validity of the underlying test). To some extent we shall go along with such practice, especially in large-sample settings where the data-derived correlations will very likely

obey a reasonable facsimile of the classical theoretical distribution. However, in §7 we will outline some research that can be done to provide some valid procedures for the new SHAPE-derived correlation-type statistics.

Even with this improved testing procedure of the SHAPE-derived correlation-type statistics, there remains one more disadvantage, beyond that discussed in the preceding paragraph, and one that in our opinion is the more serious of the two. This is the disadvantage in a correlation statistic of having much information (spatial pattern, temporal evolution, distribution of variance) compressed into a single number. It is this information-loss property of correlation that should be the focus of new research on multiparameter statistics for intercomparing multivariate phenomena in the geosciences. In §§5,6 we introduce two new procedures (the S-Phase, T-Phase tests) for possible use in data intercomparison. In §7 we make some simple suggestions for possible multiparameter statistics (beyond those of §§5 and 6) that are compromises between the simplicities of a correlation and the complexities of a full point-by-point description of the phenomenon of interest. In this way we set the groundwork for a family of statistical procedures that would allow intercomparisons of such objects as principal components, principal vectors (empirical orthogonal functions), and other multivariate objects used in data intercomparisons.

Acknowledgments

We are grateful to Ryan Whitney for expertly typing and editing the present manuscript, the most demanding of the present series. Gini May drew the figures. Curt Mobley's work was supported in part by a grant from the EPOCS (Equatorial Pacific Ocean Climate Study) Council. We also acknowledge Jean Chatfield's cooperation in overseeing the production of the typed manuscript.

2. SHAPE and Its Descendants

We shall now analyze the statistic SHAPE into various simpler pieces using the concept of a singular value decomposition of data sets. First of all, SHAPE was derived in DIT(II), and according to (2.2) of that study we have,

$$\text{SHAPE} = 2\left\{1 - \sum_{t=1}^{n}\left[\frac{\underline{d}'(t)-\underline{d}_o}{\sigma_D}\right]^T\left[\frac{\underline{m}'(t)-\underline{m}_o}{\sigma_M}\right]\right\} \tag{2.1}$$

The notation in (2.1) is also defined in Appendix B below where we exhibit the decomposition of a standardized n×p data set $\underline{\tilde{D}} = \underline{A}'\underline{\tilde{K}}^{\frac{1}{2}}\underline{E}^T$ into its singular-value form.* Thus the first of the two factors in (2.1) has the alternate representation (cf. (B3.21)):

$$\frac{\underline{d}'(t)-\underline{d}_o}{\sigma_D} = \sum_{j=1}^{p} \kappa_j \alpha_j(t)\underline{e}_j \quad , \quad t = 1,\ldots,n \tag{2.2}$$

If the standardized n×p data set $\underline{\tilde{M}}$ is given an SVD similar to $\underline{\tilde{D}}$ (using the theory of Appendix B as a pattern):

$$\underline{\tilde{M}} = \underline{B}'\underline{\tilde{L}}^{\frac{1}{2}}\underline{F}^T \tag{2.3}$$

where

$$\underline{B}' \equiv \begin{bmatrix} \beta_1(1) & \ldots & \beta_p(1) \\ \vdots & & \vdots \\ \beta_1(n) & \ldots & \beta_p(n) \end{bmatrix} \tag{2.4}$$

---

* The name "singular value decomposition" of a matrix comes from the fact that the data matrix may be rectangular and its associated covariance matrix may have zero (singular) eigenvalues. Another name that would be natural is the *principal decomposition* of a matrix, since its factors are closely related to the principal-components and principal-eigenvectors of a matrix.

and where

$$\tilde{\underline{L}} \equiv \text{diag}[\lambda_1^2,\ldots,\lambda_p^2] \tag{2.5}$$

$$\underline{F} \equiv [\underline{f}_1 \ \ldots \ \underline{f}_p] \tag{2.6}$$

$$\underline{f}_j \equiv [f_j(1),\ldots,f_j(p)]^T \tag{2.7}$$

then the second factor in (2.1) can be written as

$$\frac{\underline{m}'(t)-\underline{m}_o}{\sigma_M} = \sum_{j=1}^{p} \lambda_j \beta_j(t) \underline{f}_j \quad , \quad t = 1,\ldots,n. \tag{2.8}$$

Introducing these representations into (2.1) we find

$$\text{SHAPE}(\tilde{\underline{D}},\tilde{\underline{M}}) = 2\{1 - \sum_{j=1}^{p} \sum_{k=1}^{p} \kappa_j \lambda_k (\underline{\alpha}_j^T \underline{\beta}_k)(\underline{e}_j^T \underline{f}_k)\} \tag{2.9}$$

Here there is revealed the rich inner structure of $\text{SHAPE}(\tilde{\underline{D}},\tilde{\underline{M}})$ in terms of the temporal evolution $(\underline{\alpha}_j,\underline{\beta}_k)$ of $\tilde{\underline{D}}$ and $\tilde{\underline{M}}$, and their principal spatial patterns $(\underline{e}_j,\underline{f}_k)$, along with the distribution in space of the dimensionless variances (κ_j,λ_k) of $\tilde{\underline{D}}$ and $\tilde{\underline{M}}$. It should be plausible from this view of its innards that the power of SHAPE is relatively low. We shall next systematically reduce $\text{SHAPE}(\tilde{\underline{D}},\tilde{\underline{M}})$ to its elemental parts which have simpler structure and presumably somewhat higher power. The results are summarized in Table 2.1, in the left two columns.

The statistics of simplest structure are ORIEN_k, COREL_k, and DIAGS, as defined in (2.10)-(2.12). For example, ORIEN_k ("orientation") gives the norm between the corresponding kth principal vectors (EOFs) of $\underline{D}'$ and $\underline{M}'$, respectively. COREL_k ("correlation") gives the norm between the kth normalized principal components of the data sets, while DIAGS ("diagonal spread") compares the way all the eigenvalues κ_j^2,λ_j^2 change with j. These three types of statistic,

TABLE 2.1

NAME	SHAPE DESCENDANT	How obtained from SHAPE	Inner Product DEFINITION
S-SHAPE	$2[1 - \sum_{j=1}^{p} \kappa_j \lambda_j (\underline{e}_j^T \underline{f}_j)]$	set $\underline{A}' = \underline{B}'$	$2\{1 - (\tilde{\underline{K}}^{\frac{1}{2}} \underline{E}^T, \tilde{\underline{L}}^{\frac{1}{2}} \underline{F}^T)\}$
T-SHAPE	$2[1 - \sum_{j=1}^{p} \kappa_j \lambda_j (\underline{\alpha}_j^T \underline{\beta}_j)]$	set $\underline{E} = \underline{F}$	$2\{1 - (\underline{A}' \tilde{\underline{K}}^{\frac{1}{2}}, \underline{B}' \tilde{\underline{L}}^{\frac{1}{2}})\}$
ST-SHAPE	$2[1 - \frac{1}{p} \sum_{j=1}^{p} (\underline{\alpha}_j^T \underline{\beta}_j)(\underline{e}_j^T \underline{f}_j)]$	set $\tilde{\underline{K}} = \tilde{\underline{L}} = \underline{I}_p$	$2\{1 - (\underline{A}' \underline{E}^T, \underline{B}' \underline{F}^T)\}$
ORIEN	$2[1 - \frac{1}{p} \sum_{j=1}^{p} \underline{e}_j^T \underline{f}_j]$	set $\underline{A}' = \underline{B}'$, $\tilde{\underline{K}} = \tilde{\underline{L}} = \underline{I}_p$	$2\{1 - \frac{1}{p} (\underline{E}, \underline{F})\}$
COREL	$2[1 - \frac{1}{p} \sum_{j=1}^{p} \underline{\alpha}_j^T \underline{\beta}_j]$	set $\underline{E} = \underline{F}$, $\tilde{\underline{K}} = \tilde{\underline{L}} = \underline{I}_p$	$2\{1 - \frac{1}{p} (\underline{A}', \underline{B}')\}$
DIAGS	$2[1 - \sum_{j=1}^{p} \kappa_j \lambda_j]$	set $\underline{A}' = \underline{B}'$, $\underline{E} = \underline{F}$	$2\{1 - (\tilde{\underline{K}}^{\frac{1}{2}}, \tilde{\underline{L}}^{\frac{1}{2}})\}$

Special fragments of ORIEN and COREL, and also the vector form of DIAGS, are of interest:

$$\text{ORIEN}_k = 2(1 - \underline{e}_k^T \underline{f}_k) \quad , \quad k = 1, \ldots, p \tag{2.10}$$

$$\text{COREL}_k = 2(1 - \underline{\alpha}_k^T \underline{\beta}_k) \quad , \quad k = 1, \ldots, p \tag{2.11}$$

$$\text{DIAGS} = 2(1 - \underline{\kappa}^T \underline{\lambda}) \quad , \tag{2.12}$$

where $\underline{\kappa} = [\kappa_1, \ldots, \kappa_p]^T$, $\underline{\lambda} = [\lambda_1, \ldots, \lambda_p]^T$

HIERARCHY OF MODEL/DATA INTERCOMPARISON STATISTICS DERIVED FROM THE (EUCLIDEAN DISTANCE)^2 BETWEEN TWO DATA SETS

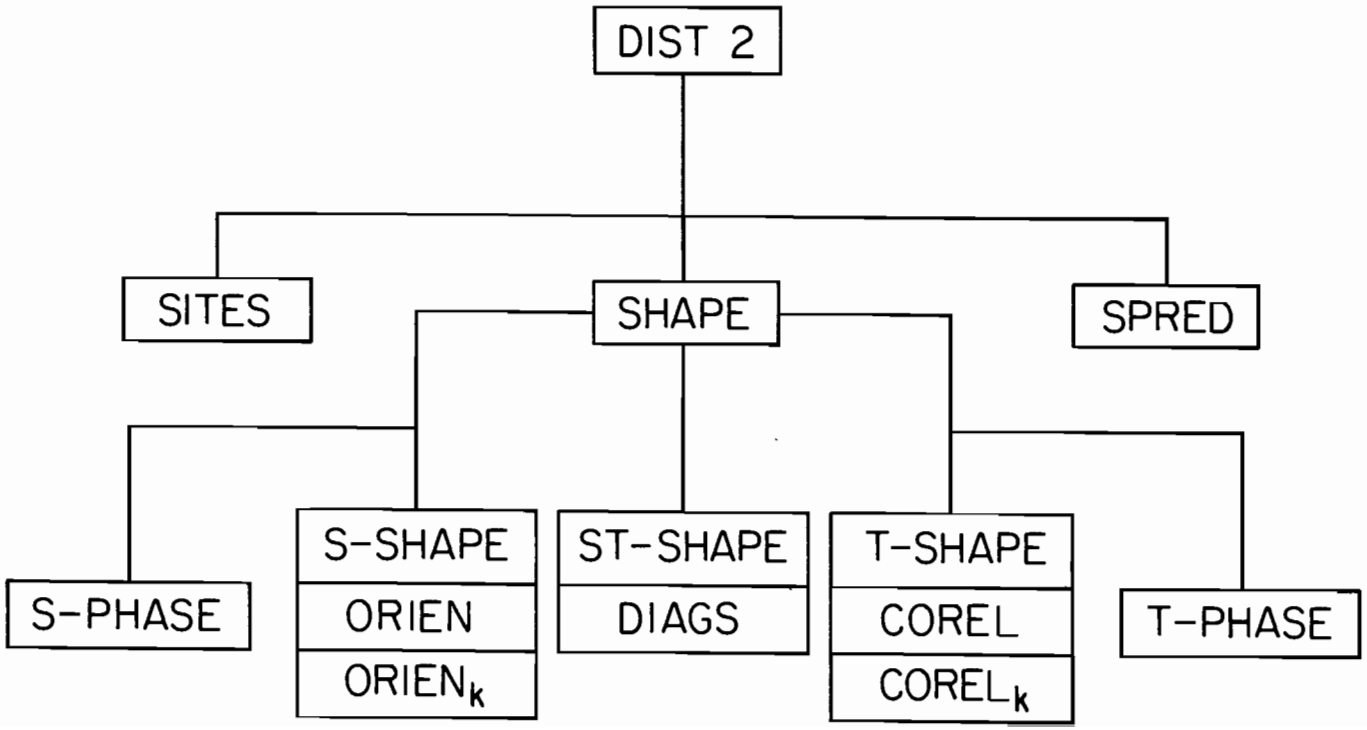


Fig. 2.1

respectively, intercompare in great detail the spatial, temporal, and variance structures of the data sets beyond those structures described by SITES, SPRED and SHAPE.

Going up the scale of complexity, we come to ORIEN and COREL, in Table 2.1. These intercompare entire eigenframes and principal component frames: $\underline{E}$ with $\underline{F}$ and $\underline{A}'$ with $\underline{B}'$, respectively, as can be seen from their forms involving inner products of $\underline{e}_j, \underline{f}_j$ and $\underline{\alpha}_j, \underline{\beta}_j$, respectively. Since there is now more structure crammed into one number by these statistics, we would expect lower power than their counterparts in (2.10), (2.11).

The S-SHAPE statistic in Table 2.1 weights the vectors $\underline{e}_j, \underline{f}_j$ with their respective eigenvalues as shown. In this way ORIEN is somewhat sharpened by helping it to emphasize those vectors in the sets $\{\underline{e}_1, \ldots, \underline{e}_p\}$, $\{\underline{f}_1, \ldots, \underline{f}_p\}$ that are the more important with regard to variance. A similar set of observations holds for T-SHAPE in Table 2.1. Finally, ST-SHAPE eliminates only the statistical effects of the eigenvalues, again (as SHAPE itself) leaving a considerable amount of structure to be represented by one number. The ST-SHAPE statistic is included mainly for completeness. It is not anticipated that ST-SHAPE will be of much use in practice, since it mixes together spatial and temporal information much in the way SHAPE does. An overview of these statistics, and others that are derived from DIST2, is given in Fig. 2.1.

3. Correlational Structure of SHAPE and its Descendants

When we view the $n \times p$ data sets $\underline{\tilde{D}}, \underline{\tilde{M}}$ as points in E_{np}, then it becomes clear that the essential structures of SHAPE, S-SHAPE, T-SHAPE, and all the other descendants of SHAPE in Table 2.1, are manifestations of the classical correlation coefficient. The basis for this assertion is developed in Appendix A.

Using the notion of the inner product $(\tilde{\underline{D}},\tilde{\underline{M}})$ of $\tilde{\underline{D}}$ and $\tilde{\underline{M}}$, we see that

$$\text{SHAPE}(\tilde{\underline{D}},\tilde{\underline{M}}) = 2\ \{1 - (\tilde{\underline{D}},\tilde{\underline{M}})\} \tag{3.1}$$

It is clear from this (cf. (A2.19)) that $\text{SHAPE}(\tilde{\underline{D}},\tilde{\underline{M}})$ is the distance between two unit vectors $\tilde{\underline{D}},\tilde{\underline{M}}$ in E_{np}, where $\tilde{\underline{D}},\tilde{\underline{M}}$ in another view (sec. 2 of Appendix B) are standardized n×p data sets. Using this broad algebraic perspective, we can then view all the descendants of SHAPE in similar manners, and the results are displayed in the right column of Table 2.1. To deduce the indicated inner product forms listed in the table, use the SVD forms of $\tilde{\underline{D}},\tilde{\underline{M}}$ (cf. (B3.18), (C1.2)) and the appropriate one of the formulas (A2.22), (A2.23).

4. Limitations of Correlational Statistics

A. Information Loss by Compression

We alluded several times in the Introduction to the fact that much information compression takes place in the statistic SHAPE. This behavior is a general characteristic of correlational statistics, i.e., of that class of statistics whose algebraic form involves inner products or algebraic combinations of inner products. This interrelation of correlations and inner products is summarized in (A2.19) (on the matrix level) and in (A2.20) (on the vector level). We shall now give a simple example that illustrates this compression property of correlations.

Let $\underline{D} = \{d(t,x):\ t = 1,\ldots,n;\ x = 1,\ldots,p\}$ and $\underline{M} = \{m(t,x):\ t = 1,\ldots,n;\ x = 1,\ldots,p\}$ be two space-centered data sets. Specifically, we set $p = 24$, and these 24 points are arranged as a 4×6 rectangle on the ocean's surface, say. One may picture the set as one of numbered points as sketched below:

1	2	3	4	5	6	N
7	8	9	10	11	12	
13	14	15	16	17	18	N′
19	20	21	22	23	24	

At each one of these points x, $m(\cdot,x)$ and $d(\cdot,x)$ are functions that take on values for $t = 1,\ldots,n$. At each x and t, let $m(t,x)$ be related to $d(t,x)$ by the rule

$$m(t,x) = g(x)\, d(t,x) \tag{4.1}$$

where for the moment we set

$$g(x) = \begin{cases} g & \text{for } x = 1,\ldots,12 \quad \text{(the domain N)} \\ 1 & \text{for } x = 13,\ldots,24 \quad \text{(the domain N′)} \end{cases} \tag{4.2}$$

Hence, depending on the value of g, the $m(t,x)$ values are greater or smaller than the $d(t,x)$ in the "northern half" N of the common data domain sketched above. On the southern half N′ of the whole domain, $m(t,x)$ and $d(t,x)$ agree. Computing the inner product (Appendix A) of this $\underline{M}$ and $\underline{D}$, we find

$$\begin{aligned} (\underline{D},\underline{M}) &= \sum_{x=1}^{p} \sum_{t=1}^{n} d(t,x)\, m(t,x) \\ &= g \sum_{x\in N} \sum_{t=1}^{n} d^2(t,x) + \sum_{x\in N'} \sum_{t=1}^{n} d^2(t,x) \\ &\equiv gV + V' \end{aligned} \tag{4.3}$$

Where V and V′ are defined in the context of (4.3). The norms of $\underline{D}$ and $\underline{M}$ are

$$\|\underline{D}\|^2 = (\underline{D},\underline{D}) = V+V' \tag{4.4}$$

$$\|\underline{M}\|^2 = (\underline{M},\underline{M}) = g^2V+V' \tag{4.5}$$

Then we find the inner product, i.e., the correlation, of the normalized data sets to be:

$$r = \left(\frac{\underline{D}}{\|\underline{D}\|}, \frac{\underline{M}}{\|\underline{M}\|}\right) = \frac{gV+V'}{[(g^2V+V')(V+V')]^{\frac{1}{2}}} \tag{4.6}$$

To see the point of the present example very simply, let us set $n = 2m$, and $d(t,x) = 1$ for all x and $t = 1,\ldots,m$; and $d(t,x) = -1$ for all x and $t = m+1,\ldots,n$. Then $V = V' = 12n$, and so (4.6) reduces to

$$r = \frac{g+1}{[2(g^2+1)]^{\frac{1}{2}}} \tag{4.7}$$

Now r is a simple function of the magnification factor g over the domain N. If $g = 1$, then $r = 1$, as expected (perfect correlation). If $g = 0$, then for each t we have a marked contrast in $m(t,x)$ and $d(t,x)$ values between the northern and southern halves of the domain. For this, $r = 0.707$. If g were very large (say infinite), then $r = 1/2^{\frac{1}{2}} = 0.707$ once again. Therefore the correlation coefficient doesn't distinguish between these two different spatial patterns for $g = 0$ or $g = \infty$. Indeed, it may be checked that the same value of r in (4.7) holds not only for an equal north-south partition of the 24 point field but also for an equal east-west partition for the two choices of g (i.e., $g = 0$ or $g = \infty$); and for that matter, $r = 0.707$ for any of the 2,704,156 partitions of the 24-point domain into two 12-point subfields. Therefore, as the spatial configurations vary over all possible partitions of the 24-point field into two equal subfields, r stays fixed at $r = 0.707$ for either of these two choices of g. Under the present sampling conditions, the sample size is 24. With this sample size the r value of 0.707 is significant on the 1% level. If we set $g = -1$, then the visual contrast between the two fields (North-South, East-West, etc.) is just as great as the preceding $g = 0$ or

$g = \infty$ cases. But now $r = 0$, and the correlation is not significant for any of the 2,704,156 possible partitions.

To summarize, as long as g was non negative, the correlation coefficient did a creditable job of spotting significant departures of the overall patterns from the usual random types. When g went negative, the potentially significant patterns for $g = -1$ were arithmetically wiped out as the correlation summation extended over N and N′. Moreover, even in the statistically significant cases ($g = 1, \infty$), we observed an insensitivity of the correlation coefficient to geographical redistributions of the two contrasting regions N and N′.

Of the two limitations in the correlation coefficient just summarized, that dealing with insensitivity to geographical distributions of variance is essentially removable. One simply performs a new correlation significance test over the smaller regions of interest. For example if an arithmetic cancellation is suspected over N+N′ when $g = -1$ and N,N′ are the north, south partition elements, then, while $r = 0$ over N+N′, we would in turn find $r = 1$ over N and over N′ separately for $d(t,x)$, $m(t,x)$ restricted to these smaller domains. However, the limitation of arithmetic cancellation, the first and more essential limitation of the correlation calculation, is still around to do its nullifying work over these smaller regions.

B. Power Loss by Dimension

The second example of correlation limitations concerns the angular separation of two unit vectors such as $\underline{e}_j, \underline{f}_j$ or $\underline{\alpha}_j, \underline{\beta}_j$ occurring in the SVD of the data sets $\underline{M}$ and $\underline{D}$. This separation is measured by $\underline{e}_j^T \underline{f}_j$ or $\underline{\alpha}_j^T \underline{\beta}_j$, respectively. Thus, specifically, in E_p we would have $\underline{e}_j^T \underline{f}_j = \cos\theta$, where θ is the angle between $\underline{e}_j$ and $\underline{f}_j$. We will now describe how the probability distribution of $\cos\theta$ changes with increasing dimension p of the space containing $\underline{e}_j$ and $\underline{f}_j$.

The change is such that it induces power losses in the ORIEN test as p increases. To aid in visualizing the distribution, *we will hold $\underline{f}_j$ fixed*, and project $\underline{e}_j$ onto the *fixed* $\underline{f}_j$. Then, if we draw $\underline{e}_j$ randomly from a population of unit vectors that are distributed uniformly* over the (p-1)-dimensional surface of a unit sphere in E_p, we find that the distribution (for fixed $\underline{f}_j$) of $\underline{e}_j^T\underline{f}_j = \cos\theta$ ($\equiv x$) is

$$P(x) = D(1-x^2)^{\frac{1}{2}(p-3)} \quad , \quad p \geq 2 \ , \tag{4.8}$$

where

$$D = \Gamma(\tfrac{1}{2}p)/\pi^{\frac{1}{2}}\Gamma(\tfrac{1}{2}(p-1)) \ , \ -1 \leq x \leq 1.$$

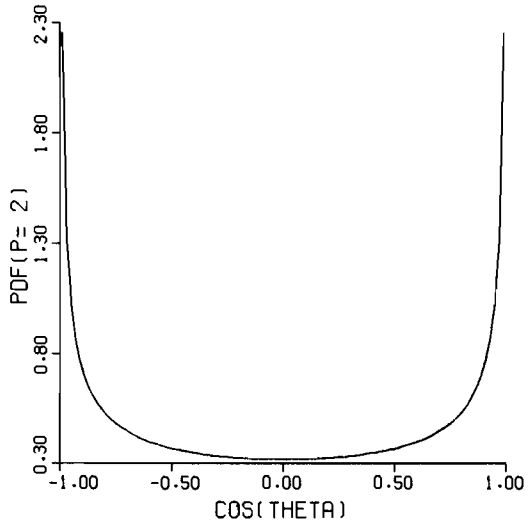
This formula is derived in Appendix E. Some plots of P(x) vs x are shown in Figs. 4.1-4.8 for selected values of p. Starting with p = 2, we have the pdf of the component of a unit vector in the plane: Picture the vector on a unit circle in the xy plane, base at the origin, arrow on the circle. The tips of the vectors in the population are uniformly distributed, by construction, over the circle. The projections of the tips on the x axis (along which $\underline{f}_j$ lies) produce cosθ values ranging from -1 to +1. Notice how the cosθ's dwell mostly around +1 and -1. This is intuitively clear. In Figure 4.2 we have p = 3 and we can still visualize the geometry: the tips of the vectors lie uniformly distributed over the two dimensional unit sphere in E_3. The projections of these vectors on the x axis are distributed uniformly† on the axis, as shown.

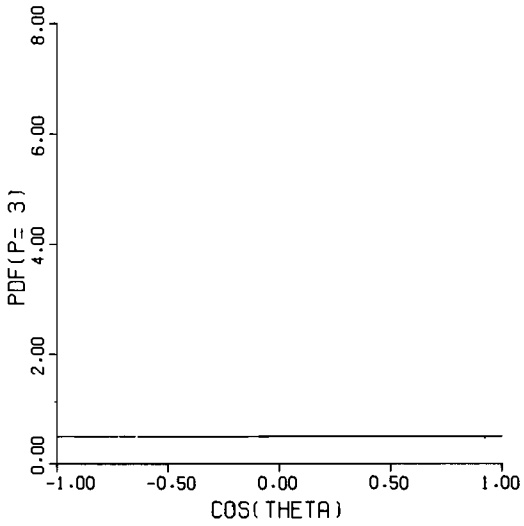
---

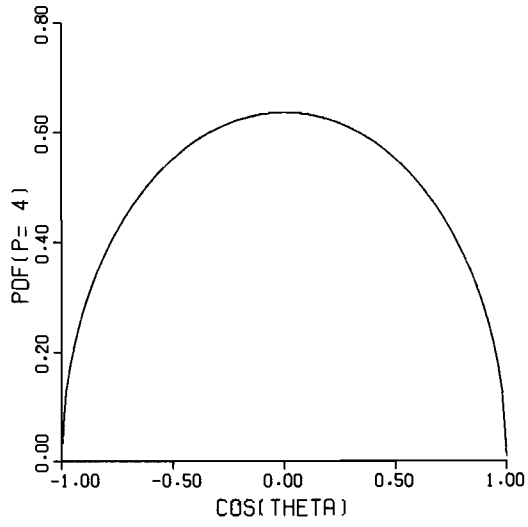
* If we draw p random, independent samples from N(0,1) and form a unit vector in E_p, then this type of vector will have a uniform distribution over the (p-1)-sphere in E_p.

† This is related to the euclidean-geometry theorem that says the area of a spherical cap or zone is proportional to its height.

PROBABILITY DENSITIES FOR RANDOM UNIT VECTOR COMPONENTS







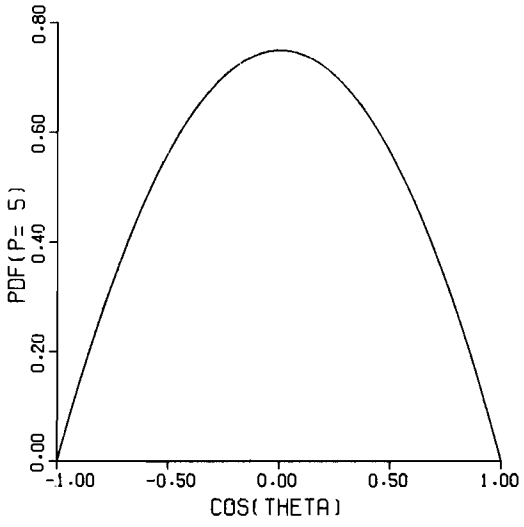
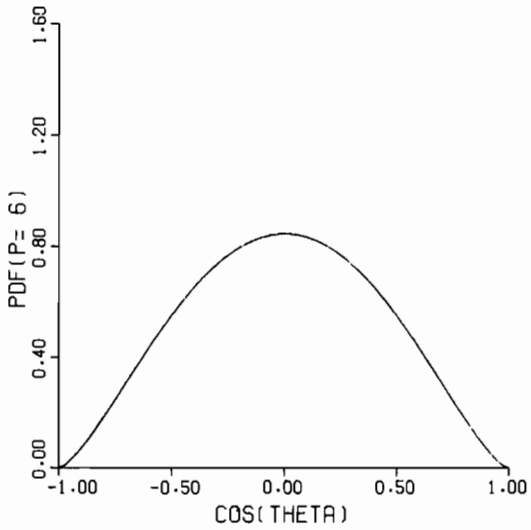
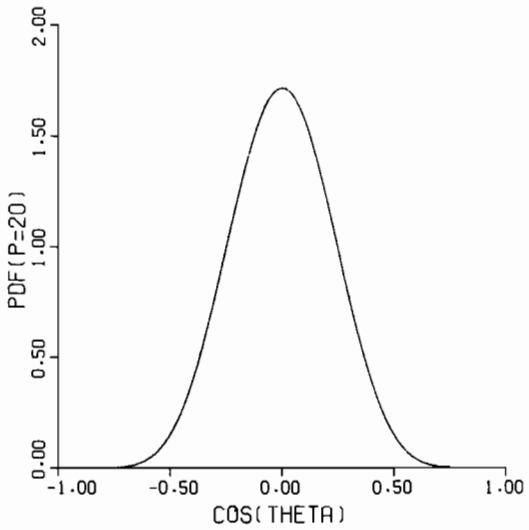


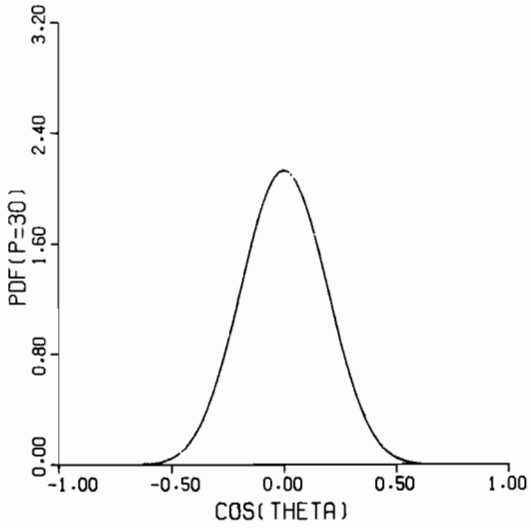
Fig. 4.1 **Fig. 4.2**

Fig. 4.3 **Fig. 4.4**

PROBABILITY DENSITIES FOR RANDOM UNIT VECTOR COMPONENTS







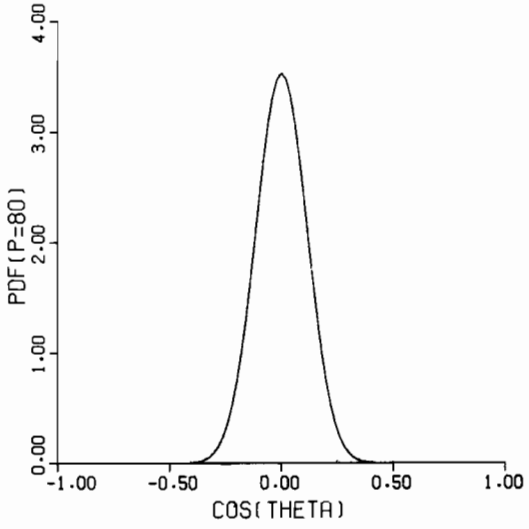


Fig. 4.5 **Fig. 4.6**

Fig. 4.7 **Fig. 4.8**

The ordinate of the straight line is ½. As we go on to p = 4, visualization ceases, and we stay close to (4.8) for guidance. The curve of the pdf is now a circle (do not be misled by the choice of ordinate scale in Fig. 4.3). The projections of the four dimensional unit vector on the x axis now tend to gather more toward $\cos\theta = 0$. For p = 5, the pdf curve in Fig. 4.4 is a parabola; the tendency for $\cos\theta$ to cluster around 0 is beginning to be more pronounced. At p = 6, in Fig. 4.5, a definite change of character of the pdf curve takes place: it is now bell-shaped. From this point on, i.e., from p = 6 in Fig. 4.5, to p = 80 in Fig. 4.8, the trend of the population's $\cos\theta$ values toward the origin is inexorable; the height increases linearly with p, the width goes down as $1/p^{½}$, and the pdf approaches Dirac's delta singularity, as $p \to \infty$.

The practical consequence of this crowding of the components of a random unit vector (and hence the values of random correlations) toward zero is that the power of the ORIEN statistic decreases with increasing p: in higher dimensional settings, two randomly selected unit vectors are very likely to be found nearly orthogonal. This is in contrast to the everyday case of p = 2 where the opposite is the case. As a result, as we rotate one unit vector away from another in E_p, and $\cos\theta$ drops from 1 down toward 0, the larger the p is, the longer the drop is before $\cos\theta$ is out of the (say) 5% critical region of the right tail of the pdf. Translated into angles, this means that, starting from 0°, in an E_{10} setting, say, we reach 75° or so before a statistic such as ORIEN can detect with confidence 95% that the null hypothesis no longer holds. We shall illustrate this phenomenon with some examples.

In Fig. 4.9 we have displayed some power curves of ORIEN. In the upper panel we have the case of n = 50 and one of the curves labeled "p = 2." This is how that curve was made: Two 50×2 (= n×p) data sets $\underline{D}', \underline{M}'$ were drawn from $N_2(0, \underline{I}_2)$; i.e., the 50 rows of these matrices were drawn randomly, one by one,

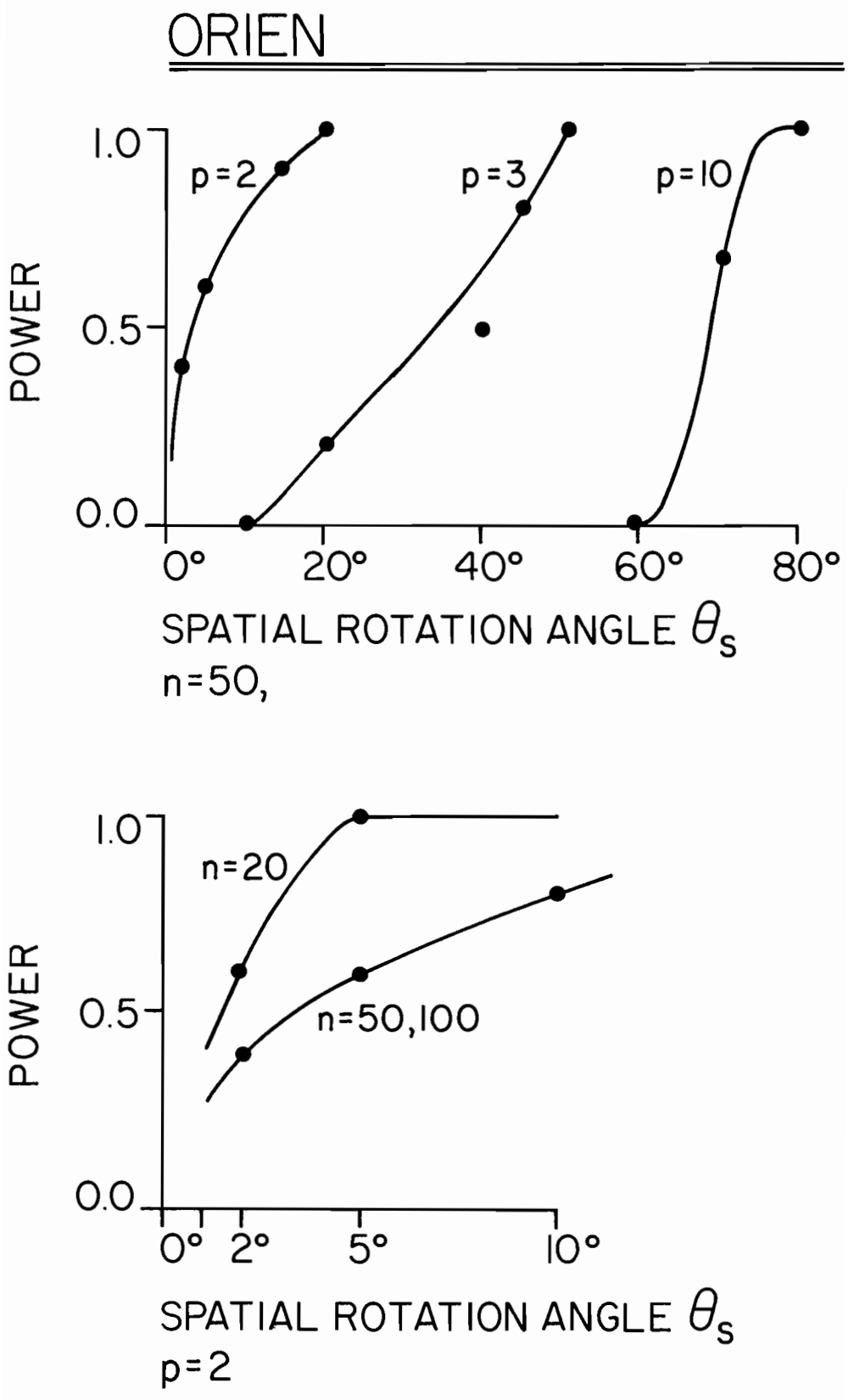

ORIEN
1.0
0.5
0.0
POWER
p=2
p=3
p=10
0°
20°
40°
60°
80°
SPATIAL ROTATION ANGLE θ_s
n=50,
1.0
0.5
0.0
POWER
n=20
n=50,100
0°
2°
5°
10°
SPATIAL ROTATION ANGLE θ_s
p=2

Fig. 4.9

from a spherical two dimensional normal distribution. The $\underline{e}_j$ and $\underline{f}_j$ eigenvectors, j = 1,2, were found using the theory of Appendix B, and ORIEN was then formed, as defined in Table 2.1. One hundred such draws of $\underline{D}'$ and $\underline{M}'$ were made and from the 100 resultant realizations of ORIEN, a cumulative distribution function (cdf) was constructed. This constituted the cdf of the null hypothesis--i.e., that $\underline{D}'$,$\underline{M}'$ were drawn from the same population. Next, 10 fresh realizations of $\underline{D}'$ were drawn from the population. Each of the 10 $\underline{E}$ frames of the realizations was homogeneously spatially rotated* by 2½° (to simulate draws of $\underline{D}'$ and $\underline{M}'$ from populations of different orientation), and it was noted whether the ORIEN between the rotated $\underline{E}$ frame and the original $\underline{E}$ frame of $\underline{D}'$ lay in the 5% right critical region of the constructed ORIEN cdf. If it did, the trial was called a "success." (The null hypothesis was detected to be false.) Out of 10 such trials 4 were successful, and this is recorded by the leftmost point of the "p = 2" curve (with ordinate 0.4) in Fig. 4.9. The remaining three points were made in a similar manner from 10 trials each. From the resultant curve for p = 2 we see that the power of ORIEN increases rapidly, and by 15° the power is 0.9. The remaining two curves in the upper panel of Fig. 4.9 were constructed similarly. The lower panel in Fig. 4.9 gives more detail on the power of ORIEN for p = 2 when different sample sizes are taken.

The curve for p = 10 in Fig. 4.9 is of special interest here. This illustrates our comments above about the loss of power of ORIEN in higher dimensional spaces. Thus we must rotate an $\underline{E}$ frame in E_{10} on the order of 75°

---

* Refer to (C2.19) in Appendix C, for the theory of spatial rotations. There, p = 2, so ℓ = 1 for the present example. Hence there was only one canonic rotation angle by which to rotate in this example. In the other curves of Fig. 4.9, say for p = 10, then ℓ = 5, and a *homogeneous* rotation means simply that all five θ_j, j = 1,...,5, were set equal to a common value θ, which was then varied.

before ORIEN can detect (with confidence 95%) 9 out of 10 such rotations. The power magnitude and confidence level here are satisfactory; it is the size of the angle that is not satisfactory, being unworkably large.

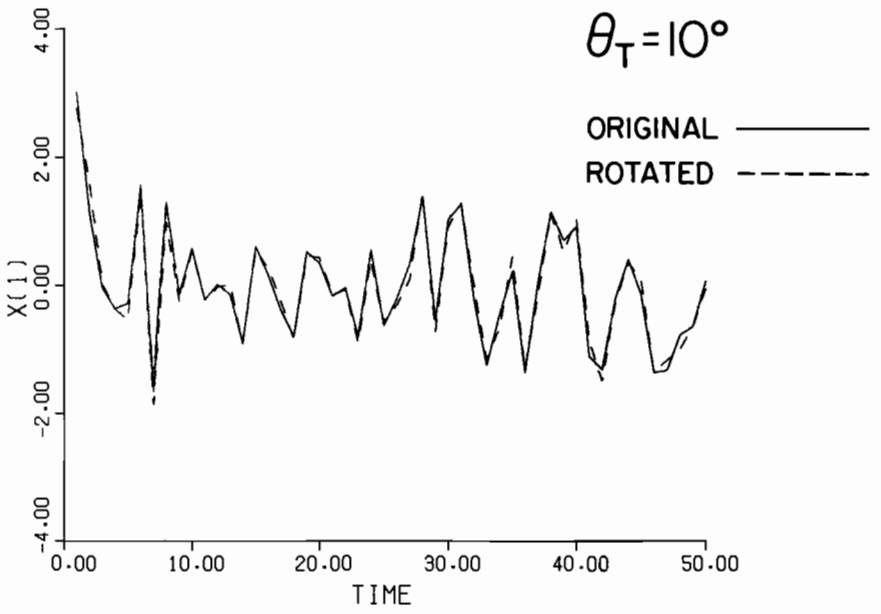
We shall next illustrate what it means physically for a unit vector in E_{50} to be rotated by various amounts. The result is quite representative of general rotations in E_p with high values of p.

One way to illustrate this rotation effect in higher-dimension spaces and still retain a semblance of visualizability is to work with a principal component time series, as follows. The rotations will then be temporal rather than spatial. We generated a 50×2 random data set $\underline{D}'$ just as we did in the context of the p = 2 curve of Fig. 4.9. We then found the first (unstandardized)* principal component time series $\{a_1(t): t = 1,\ldots,50\}$ as defined in (B3.13) of Appendix B. We next constructed a 50×50 temporal rotation matrix $\underline{L}$ as shown in (C2.19). Note that the rotation theory leading to (C2.19) is general, so that we can replace "p" everywhere by "n" in that theory. Moreover, we can choose the 50×50 matrix $\underline{W}$ to be anything we please in the present example, and we chose it to be $\underline{I}_{50}$. We further may set all θ_j in $\underline{L}$ equal to a common angle θ, thereby producing a *homogeneous* temporal rotation. We then applied the homogeneous temporal rotation, for varying amounts of θ, to the vector $\underline{a}_1 = [a_1(1),\ldots,a_1(50)]^T$. The results are shown in the sequence of Figures 4.10-4.16. In Fig. 4.10 the solid curve represents the unrotated, original 50-component principal vector $\underline{a}_1$; the dashed curve represents the 50-component $\underline{a}_1$ rotated by an homogeneous rotation of 10°. As shown in the

---

* That is, do not divide the centered data set by σ_D. The result is $a_j(t) = \sigma_D \tilde{a}_j(t)$, j = 1,...,p.

TEMPORAL ROTATION OF
FIRST PRINCIPAL COMPONENT
$a_1(t)$ FOR p=2, n=50



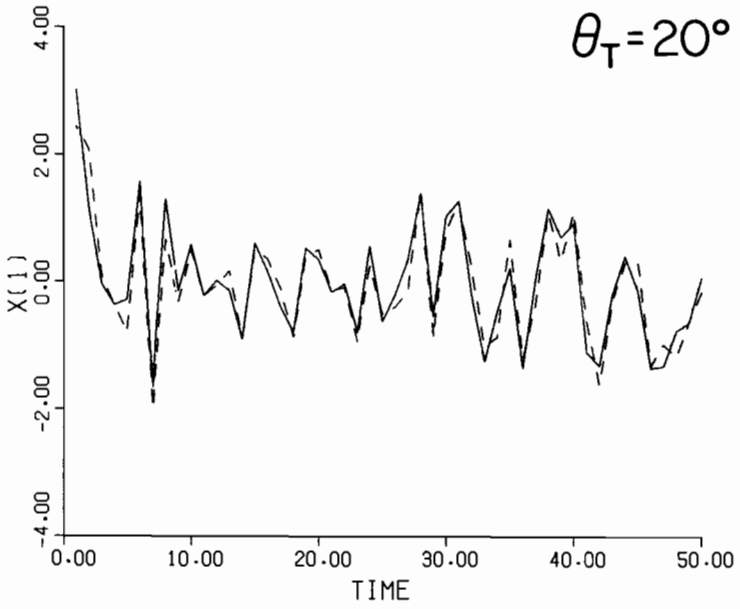


Fig. 4.11

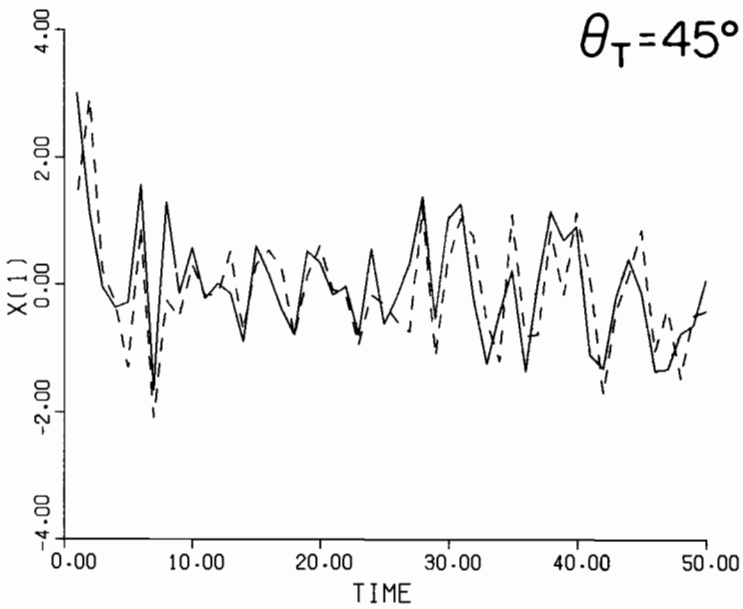


Fig. 4.10

Fig. 4.12

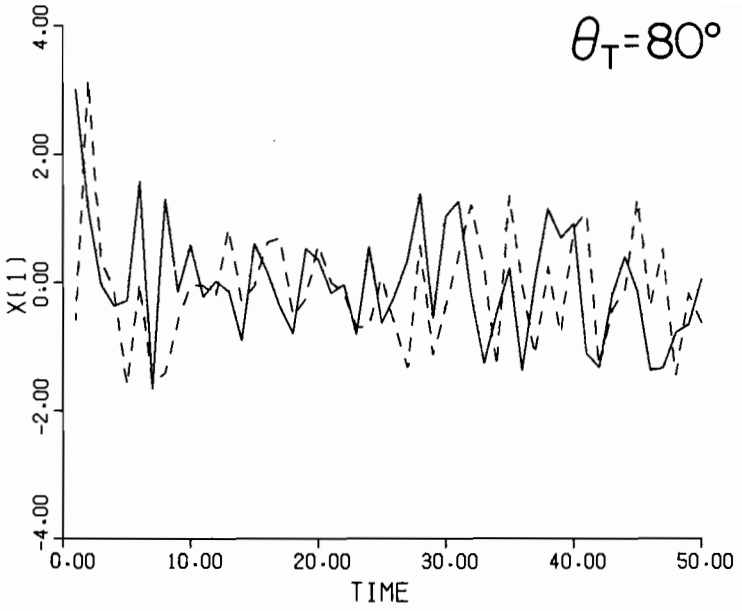


Fig. 4.13

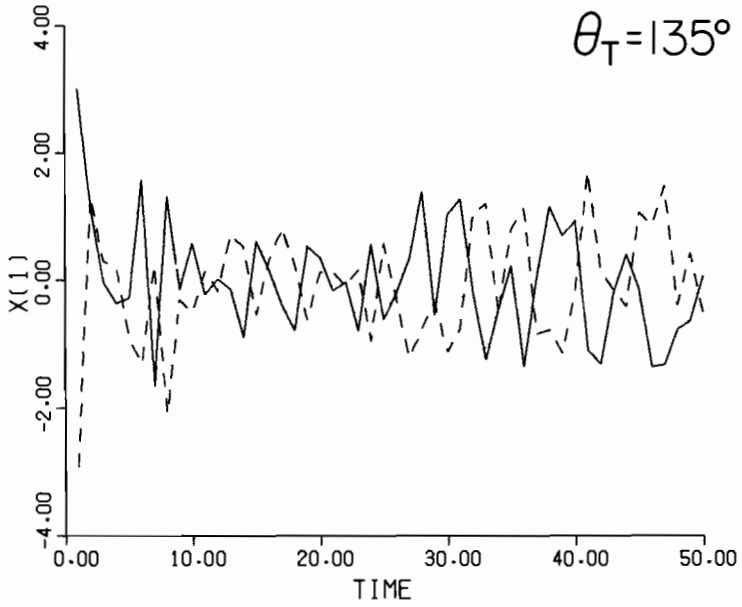


Fig. 4.15

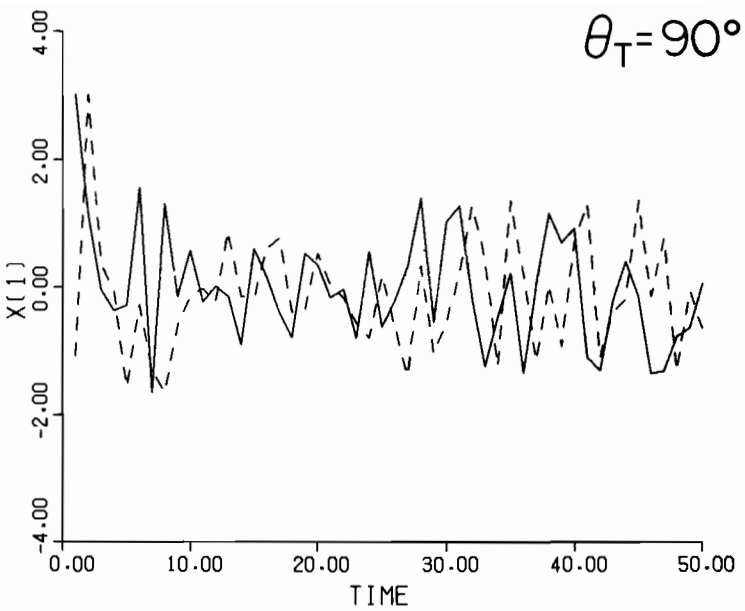


Fig. 4.14

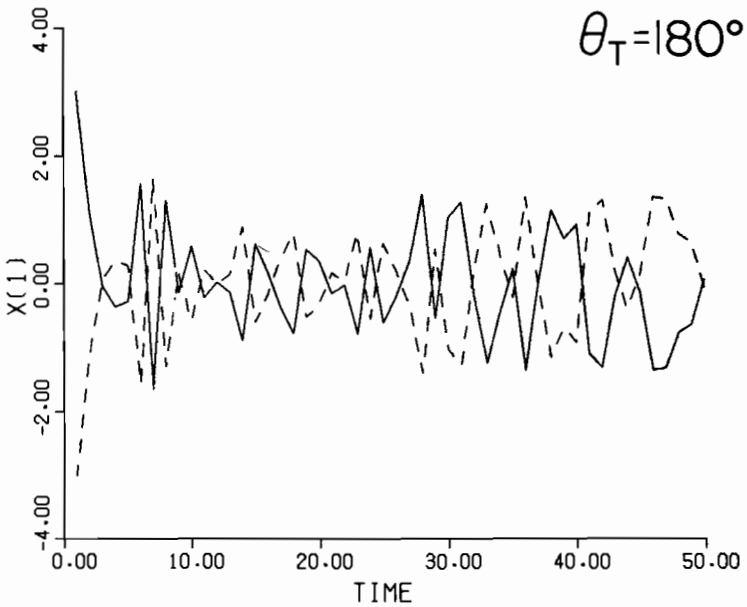


Fig. 4.16

subsequent figures, the rotation effect becomes noticeable between 20° and 45°. By 90°, the two curves are shifted considerably. Observe that it is not a rigid shift, but a less obvious one dictated by the structure of (C2.19). It is easy to obtain explicit expressions for each t-value of $a_1(t)$ as a function of θ to see what this dependence is. For this, we note parenthetically that we can write $\underline{R}$ in (C2.19) as:

$$\underline{R} = \sum_{j=1}^{\ell} [c_j(\underline{x}_j\underline{x}_j^T + \underline{y}_j\underline{y}_j^T) + s_j(\underline{x}_j\underline{y}_j^T - \underline{y}_j\underline{x}_j^T)] \qquad (4.9)$$

See Sec. 6 of Appendix C. Hence if $\underline{z}$ is a $p\times 1$ vector, then its rotated form $\underline{w}$ is given by:

$$\underline{w} = \underline{R}\,\underline{z} = \sum_{j=1}^{\ell} [(c_j\xi_j + s_j\eta_j)\underline{x}_j + (-s_j\xi_j + c_j\eta_j)\underline{y}_j] \qquad (4.10)$$

where we have written

$$'\xi_j' \text{ for } \underline{x}_j^T\underline{z}$$
$$'\eta_j' \text{ for } \underline{y}_j^T\underline{z} \quad , \quad j = 1,\ldots,\ell.$$

Continuing to study the sequence of figures, particularly the figures for θ_T = 80°, 90°, we gain an immediate impression of the effect on the time series $\{a_j(t): \; t = 1,\ldots,50\}$ of an homogeneous rotation of 80° to 90°. This is, as suggested by the trend of curves in the upper panel of Fig. 4.9, the amount of rotation needed to attain a power of 0.9 or more for the statistics ORIEN or COREL, when p or n is 50. This amount of rotation is large and easily detectable by eye. In the presence of such shifts we do not need a statistic such as ORIEN to tell us the two curves are significantly different.

What we need for this task are more sensitive statistics, ones that are not of the correlation type, when working in high-dimensional settings with smaller rotation angles. We next turn to consider this matter.

5. S-Phase Test

We now introduce the S-Phase test, which is designed to tell whether the $\underline{E}$ and $\underline{F}$ frames of a given data/model pair $\underline{D},\underline{M}$ (and hence their spatial patterns) are significantly close or distant, as the case may be. The central concept used in the test is the set of canonic rotation angles between $\underline{E}$ and $\underline{F}$. The theory of these angles is given in Appendix C, and how they enter the S-Phase test will be fully described in the discussion paragraphs below. A preliminary glance at Appendix C for notation and terminology will help the reader prepare for a first reading of the S-Phase test.

We should state at the outset that the S-Phase test, despite its somewhat finished appearance, is actually but a first attempt to systematically and objectively gauge the distance between data/model eigenframes $\underline{E}$ and $\underline{F}$ in a practical way. On the one hand, the test, as it is presently constructed, has several strong features, particularly in the way the rotation angle distributions are linked to the values of the data matrix $\underline{D}$ and hence to the physics of the process under study. On the other hand the test shows some weakness in the behavior of the canonic rotation angles as measures of distance. These positive and negative features will be discussed after the test has been defined.

The S-Phase test below has been specifically devised for the case where the number of samples n and the number p of points in space are such that $n-1 \geq p$. The theory of Appendix B then shows that the $p \times p$ $\underline{E}$ and $\underline{F}$ frames are uniquely defined, and rotations in E_p can uniquely relate $\underline{E}$ and $\underline{F}$. Moreover,

it is then possible to define a unique set of canonic rotation angles between $\underline{E}$ and $\underline{F}$. The complementary case n-1 < p, as the theory of Appendix B shows, has associated at most only n-1 non-zero eigenvalues and their uniquely defined eigenvectors. In this case there is no unique rotation between those sets of column vectors of $\underline{E}$ and of $\underline{F}$ that are uniquely determined. However, it is possible in this case to use the formalism of the T-Phase test. How this can be done will be outlined in the closing paragraph of §6.

A. S-Phase Test Stages

The S-Phase test has four main stages. Stage I finds the canonic rotation angles $\theta_1,\ldots,\theta_\ell$, $\ell = [p/2]$ between the space-centered versions of the given n×p primitive data sets $\underline{D}',\underline{M}'$. In Stage II the cumulative distribution of the θ_j is found under the null hypothesis, using a Monte Carlo procedure. Stage III connects the θ_j with the values of $\underline{D}$ in such a way that we know on the average how much of a rotation of $\underline{D}$'s $\underline{E}$-frame will produce a given percent change in $\underline{D}$'s field values. Finally, in Stage IV, the results of Stages II, III are combined with that of Stage I to decide whether $\underline{E}$ and $\underline{F}$ are significantly close or significantly distant, as the case may be.

STAGE I. (Canonic Rotation Angles θ_j)

1. Given: two n×p matrices $\underline{D}',\underline{M}'$. Center them in space to obtain $\underline{D},\underline{M}$, respectively.
2. Find the $\underline{E}$ and $\underline{F}$ frames of $\underline{D}$ and $\underline{M}$, respectively.
3. Apply the octant and chirality conditions to $\underline{E},\underline{F}$.
4. Construct $\underline{R} \equiv \underline{F}\,\underline{E}^T$ from the conditioned $\underline{E},\underline{F}$.
5. Find the canonic rotation angles $\theta_1,\ldots,\theta_\ell$, of $\underline{R}$, where $\ell = [p/2]$.

STAGE II. (Reference Distribution for Canonic Rotation Angles)

1. Select n×p matrices $\underline{D}^{(i)},\underline{M}^{(i)}$ randomly from $N_p(\underline{0},\underline{I}_p)$. Center $\underline{D}^{(i)},\underline{M}^{(i)}$ in space (average over n dimension). Here i is a realization index i = 1,...,r.
2. Find the $\underline{E}^{(i)},\underline{F}^{(i)}$ frames of $\underline{D}^{(i)},\underline{M}^{(i)}$, respectively, i = 1,...,r.
3. Apply the octant and chirality conditions to $\underline{E}^{(i)},\underline{F}^{(i)}$, i = 1,...,r.
4. Construct $\underline{R}^{(i)} \equiv \underline{F}^{(i)}\underline{E}^{(i)T}$ from the conditioned $\underline{E}^{(i)},\underline{F}^{(i)}$, i = 1,...,r.
5. Find the canonic rotation angles $\theta_1^{(i)},\ldots,\theta_\ell^{(i)}$ of $\underline{R}^{(i)}$, where $\ell = [p/2]$ and i = 1,...,r.
6. Pool all ℓr angles $\theta_j^{(i)}$, j = 1,...,ℓ; i = 1,...,r and find the cumulative distribution of this set of ℓr angles.

STAGE III. (Connecting Spatial Rotations and D-field Changes)

1. Construct rotation platform bases $\{\underline{x}_1^{(i)},\underline{y}_1^{(i)},\ldots,\underline{x}_\ell^{(i)},\underline{y}_\ell^{(i)}\}$ for the case p = 2ℓ or $\{\underline{x}_1^{(i)},\underline{y}_1^{(i)},\ldots,\underline{x}_\ell^{(i)},\underline{y}_\ell^{(i)},\underline{x}_{\ell+1}^{(i)}\}$ for the case p = 2ℓ+1, i = 1,...,r.
2. Produce $\bar{f}^{(i)}(\theta)$ vs θ, i = 1,...,r, via homogeneous spatial rotations of data set $\underline{D}$ using the ith rotation platform.
3. Average the $\bar{f}^{(i)}(\theta)$ over the i = 1,...,r to obtain $\bar{f}(\theta)$.

STAGE IV. (Deciding whether $\underline{E},\underline{F}$ are significantly close or distant)

1. Choose an $\bar{f}_a$ value, $0 \leq \bar{f}_a \leq 1$. Find that θ_a which solves $\bar{f}(\theta) = \bar{f}_a$. θ_a is the *acceptance angle*, and $\bar{f}_a$ is the *acceptance fraction*.
2. Find p_a that pairs with θ_a using the reference distribution of Step 6, STAGE II. p_a is the *acceptance probability*.
3. Choose *confidence level* 1-α of test. α is the *size* of the test.

4. Compute from p_a and α the *critical acceptance number* a(close), or a(distant), as the case may be.

5. $\underline{E}$ is significantly close to (significantly distant from) $\underline{F}$ if the number a of canonic rotation angles $\theta_1,\ldots,\theta_\ell$ (of Step 5, STAGE I) in $[0,\theta_a]$ is greater than a(close) (is less than or equal to a(distant)). In each case the decision results in a rejection of the null hypothesis H_o, i.e., that $\underline{D}$ and $\underline{M}$ of STAGE I are drawn from the same population as the $\underline{D}^{(i)}$ and $\underline{M}^{(i)}$ in STAGE II.

B. Discussion of S-Phase Test Stages

STAGE I (Discussion).

The theory of singular value decompositions of data sets is used to find $\underline{E}$ and $\underline{F}$, as explained in Appendix B. The octant and chirality conditions are defined in §5 of Appendix C. Their purpose is to uniquely fix $\underline{E}$ and $\underline{F}$ so that the canonic rotation angles are in turn uniquely defined. The computation of the canonic rotation angles is described in Appendix C. A standard IMSL eigenvalue routine can be used for the complex eigenvalues of $\underline{R}$. The theory of Appendix C provides the reader with the background that will help in the intelligent use of the subroutine. Our program, as a result of the octant condition, places all the θ's in the range $[0,\pi]$ (see also the remarks below (C4.4)).

STAGE II (Discussion).

The first five steps of this stage are exact replications of the five steps of STAGE I. The randomly selected pairs $\underline{D}^{(i)},\underline{M}^{(i)}$ are drawn from $N_p(\underline{0},\underline{I}_p)$. (We chose r = 100 for our work.) For example, $\underline{D}^{(i)}$ is obtained by np random samples from N(0,1), the normal distribution with zero mean and unit

variance. The np numbers so drawn can be used to build $\underline{D}^{(i)}$ row by row (or column by column). Ordinarily, i.e., without the octant condition, as both experiment and theory show, the $\theta_j^{(i)}$ so found, will be uniformly distributed over $[0,2\pi]$. In this case Step 6 would not be needed, as we would know the theoretical form of the reference distribution: a rising straight line over the domain $[0,2\pi]$ starting with ordinate 0 and ending with ordinate 1 at 2π. However, the octant condition prevents the θ_j of the frames from going beyond π. This effect shows up first of all in the elimination of rotation angles in $[\pi,2\pi]$, and secondly, in the slower rise of the reference distribution inside $[0,\pi]$ near π. Hence the whole point of STAGE II is to find the actual distribution of the θ_j's over the range $[0,\pi]$ as produced by the octant condition.

We put up with these nonlinear anomalies in the θ-distribution because the octant condition seems to us a sensible way to define the distance between two orthonormal frames: since the signs of the $\underline{e}_j$ and $\underline{f}_j$ in $\underline{E}$ and $\underline{F}$, respectively, do not matter to the physical representation of the fields $\underline{D}$ and $\underline{M}$, we can change these signs at will until ORIEN($\underline{E}$,$\underline{F}$) is a minimum. This minimum is what we would naturally think of as the (distance)^2 between the two frames.*

By contrast to this convention, there is another way of fixing the signs of the $\underline{e}_j$, and $\underline{f}_j$: by making the first components of the $\underline{e}_j$ and $\underline{f}_j$ nonnegative (the "hemispheric" condition). It is clear tht this generally would uniquely fix the θ's, but not at the minimum possible ORIEN. (The chirality condition would of course have to be retained.) We haven't made the experiment to see

---

* This is in analogy to the definition of distance between two arbitrary sets A,B in E_p: let d(a,b) be the distance between any point a in A and b in B. Then let a,b vary over their respective sets. The minimum of d(a,b) so found is the *distance* between A and B. This is also the sensible way to define the distance between two islands A,B on the ocean, e.g.

if the θ-distribution is uniform over $[0,\pi]$ with this "hemispheric" condition; we would conjecture that it is. If it were, then STAGE II could be eliminated, and a new version of the S-Phase test constructed. The loss of the minimum ORIEN (the octant) condition may be compensated for by a simpler and quicker S-Phase computer program.

There was ample opportunity to explore the possibilities inherent in the hemispheric condition prior to writing this report. However, we felt the matter was of secondary importance since the null hypothesis implicitly used in STAGE II's present form deals with a rather special population from which the $\underline{D}^{(i)},\underline{M}^{(i)}$ are drawn, namely $N_p(\underline{0},\underline{I}_p)$. A more realistic population would be $N_p(\underline{0},\underline{\Sigma})$ for some non-isotropic covariance matrix $\underline{\Sigma}$. The reference distribution of θ_j's for this $\underline{\Sigma}$ is most conveniently found by Monte Carlo methods. Moreover, estimating $\underline{\Sigma}$ from real data would lend a new dimension of reality to the test; also a new dimension of expense. Therefore we leave the matter of STAGE II here for the present. Its generalization to the case of $\underline{\Sigma}$ can be very simply made whenever there is a reason to do so (it will affect only Step 1).

STAGE III (Discussion)

The theory of construction of the bases involving the eigenvectors $\underline{x}_i,\underline{y}_i$ is fully covered in Appendix C. We call these bases "rotation platforms" because they are the scaffoldings from which the homogeneous spatial rotations of $\underline{D}$ are made. These rotation platforms are represented by the matrices $\underline{W}$ in (C2.22) and (C2.24). For Step 1, one may, e.g., construct these $\underline{W}$'s by repeating Steps 1-4 of STAGE II and extracting the $\underline{W}$-matrices from the $\underline{R}^{(i)}$ of Step 4. Alternatively, one may select a p×p matrix $\underline{W}'$ randomly from $N_p(\underline{0},\underline{I}_p)$. The columns of $\underline{W}'$ are linearly independent with probability 1. Then orthonormalize

these columns using the Gram-Schmidt procedure to find $\underline{W}^{(i)}$. Let $\underline{W}^{(i)}$ be the ith such random rotation platform obtained, i = 1,...,r, in one of these ways. (In our work we set r = 100.) We then rotate the given centered $\underline{D}$ matrix via the formula:

$$\underline{D}^{(i)}(\theta) \equiv \underline{D}[\underline{R}^{(i)}(\theta)]^T \quad , \quad \begin{matrix} 0 \leq \theta \leq \pi \\ i = 1,\ldots,r. \end{matrix} \tag{5.1}$$

where

$$\underline{R}^{(i)}(\theta) \equiv \underline{W}^{(i)}\underline{L}(\theta)\underline{W}^{(i)T} \quad , \quad \begin{matrix} 0 \leq \theta \leq \pi \\ i = 1,\ldots,r \end{matrix} \tag{5.2}$$

Here $\underline{L}(\theta)$ is as defined in (C2.19) or (C2.23) and *with all* θ *the same*. This is what is meant by an *homogeneous rotation*. By applying $[\underline{R}^{(i)}(\theta)]^T$ in this way, we are in effect rotating $\underline{D}^{(i)}$'s $\underline{E}$ frame:

$$\begin{aligned} \underline{D}[\underline{R}^{(i)}(\theta)]^T &= \underline{A}'\,\underline{K}^{\frac{1}{2}}\underline{E}^T\,[\underline{R}^{(i)}(\theta)]^T \\ &= \underline{A}'\,\underline{K}^{\frac{1}{2}}\,[\underline{R}^{(i)}(\theta)\underline{E}]^T \quad , \quad \begin{matrix} 0 \leq \theta \leq \pi \\ i = 1,\ldots,r. \end{matrix} \end{aligned} \tag{5.3}$$

Thus we induce changes in $\underline{D}$ via spatial rotations only; variance (via $\underline{K}^{\frac{1}{2}}$) and temporal evolution (via $\underline{A}$) remain unaffected, as well as location (via $\underline{D}_o$). Let $d^{(i)}(t,x;\theta)$ be the tx element of $\underline{D}^{(i)}(\theta)$, t = 1,...,n; x = 1,...,p, i = 1,...,r. Let d(t,x) be the tx element of the original centered matrix $\underline{D}$. Then for each t,x,θ and i, form the quantity

$$f^{(i)}(t,x;\theta) \equiv \left| \frac{d^{(i)}(t,x;\theta)-d(t,x)}{\sigma_D'} \right| \tag{5.4}$$

where

$$\sigma_D' \equiv \left[\frac{1}{np} \sum_{t=1}^{n} \sum_{x=1}^{p} [d(t,x)-d_o(x)]^2 \right]^{\frac{1}{2}} \tag{5.5}$$

Form the average

$$\bar{f}^{(i)}(\theta) \equiv \frac{1}{np} \sum_{t=1}^{n} \sum_{x=1}^{p} f^{(i)}(t,x;\theta) \tag{5.6}$$

$$i = 1,\ldots,r \quad ; \quad 0 \leqq \theta \leqq \pi,$$

Average once again to find

$$\bar{f}(\theta) \equiv (\frac{1}{r} \sum_{i=1}^{r} \bar{f}^{(i)}(\theta))/\bar{f}_{max} \tag{5.7}$$

$$0 \leqq \theta \leqq \pi$$

where $\bar{f}_{max}$ is the maximum of the numerator over the domain $[0,\pi]$. $\bar{f}_{max}$ invariably occurs at $\theta = \pi$. Hence $0 \leqq \bar{f}(\theta) \leqq 1$; $0 \leqq \theta \leqq \pi$.

From these constructions we see that $\bar{f}(\theta)$ gives the average fractional change of the original $\underline{D}$-field induced by an homogeneous rotation of θ radians. As θ increases over $[0,\pi]$, our experience is that $\bar{f}(\theta)$ increases nearly linearly with θ. In all cases the rise should be monotonic over the entire range.

STAGE IV (Discussion)

The acceptance angle θ_a of Step 1 defines a proper subinterval $[0,\theta_a]$ of $[0,\pi]$ in which we keep watch for the occurrence of canonic rotation angles θ_j, $j = 1,\ldots,\ell$. It is intuitively clear that, no matter what θ_a is or how it is obtained, the more θ_j that fall in $[0,\theta_a]$, the better is the fit between $\underline{E}$ and $\underline{F}$, i.e., the smaller is the norm $\|\underline{E}-\underline{F}\|^2$. This is the basis for the significantly-close decision in step 5. On the other hand, too few θ_j falling in $[0,\theta_a]$ raises the possibility that $\underline{E}$ and $\underline{F}$ are significantly distant from each other. The key word in each of these decisions is "significantly." This entails associating to $[0,\theta_a]$ a probability of a θ_j falling in $[0,\theta_a]$. This is done

by means of the reference distribution of Step 6, STAGE II. Moreover θ_a itself is chosen using $\bar{f}(\theta)$ in an attempt to connect up this probability with the field values of $\underline{D}$. The smaller the $\bar{f}_a$ we choose, the smaller is the resulting θ_a, and hence p_a, and so the more stringent we are intending the test to be for the significantly-close option; and the less stringent for the significantly-distant option.

What option in Step 5 is chosen is largely a function of who is doing the test. A GCM developer who is in the first stages of developing his model would likely choose the significantly-distant option. In the resultant S-Phase test for moderate choices of $\bar{f}_a$ and α, unless the model is really bad in its $\underline{F}$ frame match to the data $\underline{E}$-frame, the number a of accepted θ_j will not be less than a(distant). In the final stages of development of a GCM that appears to have promise in its spatial pattern simulations of real data, the modeler may wish to test whether $\underline{F}$ is significantly close to $\underline{E}$. Our experience is that the significantly-close test is a very demanding one (say for $\bar{f}_a = 0.1$, $\alpha = 0.1$) in contrast to the significantly-distant test which is relatively forgiving. Whether one chooses this or that option, he still has the ability to vary $\bar{f}_a$ and α.

The details for finding a(close) and a(distant), which are instrumental in reaching these decisions, are as follows. On the assumption that the θ_j are pairwise independent (which by experiment seems largely so) we can compute the probability P(j) that precisely j of the ℓ canonic rotation angles are in $[0,\theta_a]$:

$$P(j) = \binom{\ell}{j} p_a^j (1-p_a)^{\ell-j} \quad , \quad j = 1,\ldots,\ell. \tag{5.8}$$

For the significantly-close decision, find that a for which

$$\sum_{j=a+1}^{\ell} P(j) = \alpha \qquad (5.9)$$

and call the value "a(close)." The critical region is [a(close)+1,...,ℓ].

For the significantly-distant decision, find that a for which

$$\sum_{j=a+1}^{\ell} P(j) = 1-\alpha \qquad (5.10)$$

and call the value "a(distant)." The critical region is [0,...,a(distant)]. In our work we most often used $\alpha = 0.1$.

C. Some Preliminary Studies of the S-Phase Test

The feature of the S-Phase test that makes it particularly interesting from the physical point of view is STAGE III, wherein the changes in $\underline{D}$ induced by rotating the $\underline{E}$ frame of $\underline{D}$ are connected to the homogeneous rotation angles θ. The quantities $\bar{f}(\theta)$ and $\bar{f}_a$ summarize these connections, where $\bar{f}_a$ is a choosable parameter. In this paragraph we examine the effect, of changing $\bar{f}_a$, on the decision concerning the proximity of the $\underline{E}$ and $\underline{F}$ frames. Specifically, on the one hand, we are interested in seeing the S-Phase test continuing to declare that the $\underline{E}$ and $\underline{F}$ are significantly distant under a wide range of increasing-$\bar{f}_a$ conditions (starting from $\bar{f}_a = 0$) when we know that these two frames are indeed distant. On the other hand, $\underline{E}$,$\underline{F}$ frames that have been deliberately built close together should be declared by the S-Phase test to be significantly close under a wide range of decreasing-$\bar{f}_a$ conditions (starting from $\bar{f}_a = 1.0$). It turns out that these two situations are more or less symmetrical within the S-Phase formalism, and so we will illustrate only one

of them. We consider the significantly-distant case, and set up some test cases as follows.

We choose an n×p data matrix $\underline{D}$ and produce four rotated versions $\underline{M}_i$ of it, i = 1,2,3,4. We apply the S-Phase test to each pair($\underline{D}$,$\underline{M}_i$) to see if it will detect our attempts to make the $\underline{E}$ frame of $\underline{D}$ and the $\underline{F}_i$ frame of $\underline{M}_i$ close or distant. The choices of the rotation angles between $\underline{E}$ and $\underline{F}_i$ are shown in Fig. 5.1. In part (a) we display 12 rotation angles uniformly distributed over $[0,\frac{1}{2}\pi]$ to simulate a random rotation between two 24×24 matrices $\underline{E}$ and $\underline{F}_1$. In part (b) the rotation angles are crowded around 0 to simulate relatively close frames $\underline{E}$ and $\underline{F}_2$, while in (c) the reverse is true for $\underline{E}$ and $\underline{F}_3$. Case (d) is intermediate between cases (b) and (c), and it was this case that produced a surprising result, which we shall discuss presently.

To begin the intercomparisons we constructed a 36×24 (= n×p) matrix $\underline{D}$ by drawing its 864 entries randomly from the normal population N(0,1). We then used (C2.21) with $\underline{W} = \underline{I}_{24}$ to rotate $\underline{D}$ spatially (cf. (5.3)) via $\underline{R} = \underline{L}$, where the θ_j in $\underline{L}$ are specified as in (a)-(d) of Fig. 5.1. The result was $\underline{M}_i$, i = 1,...,4. The analytic forms of the θ_j, j = 1,...,12, in these cases are given below:

case (a):	$\theta_j = 90° \ (j/12)$,	uniform on [0°,90°]
case (b):	$\theta_j = 90° \ ((j-1)/12)^3$,	packed toward 0°
case (c):	$\theta_j = 90° \ (1-((j-1)/12)^3)$,	packed toward 90°
case (d):	$\theta_j = 90° \ (1+(j-6)/40)$,	packed around 90°

Once we had a data matrix pair $\underline{D}$,$\underline{M}_i$, we found the frames $\underline{E}$ and $\underline{F}_i$ and then applied the S-Phase test to this pair of frames. For all four test cases we had fixed α at 0.10, and made various choices of $\bar{f}_a$. These choices and the salient results of the S-Phase test are shown in Table 5.1. In cases (a)-(c)

FOUR DIFFERENT MODES OF E, F FRAME SEPARATION FOR CHECKING S-PHASE TEST

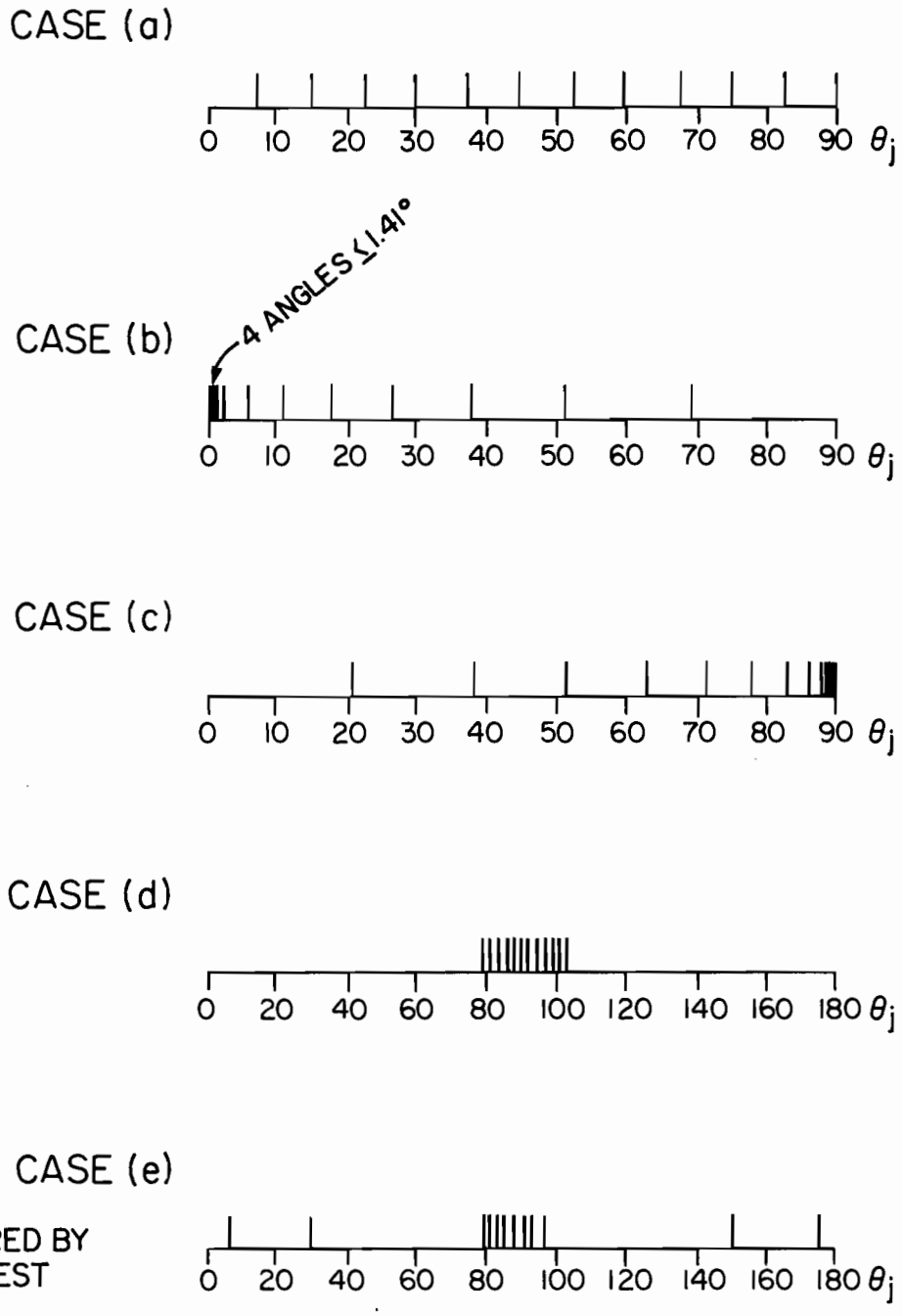


Fig. 5.1

TABLE 5.1

The results of the S-Phase Test on four different modes of E,F-frame separation. Significantly-Distant Option.

Case	$\bar{f}_a$	θ_a°	p_a	Accept or Reject H_o	{no. of θ_j in $[0,\theta_a]$}/a(distant) Reject H_o if fraction is 1 or less. Rejected H_o means signific.-distant frames.
(a)	0.10	11.6	0.095	R	1/1
	0.99	168.5	0.94	A	12/11
(b)	0.10	11.6	0.095	A	7/1
	0.99	168.5	0.94	A	12/11
(c)	0.10	11.6	0.095	R	0/1
	0.20	24.0	0.22	R	1/2
	0.55	67.5	0.48	R	4/5
	0.60	74.6	0.52	R	5/5
	0.99	168.5	0.94	A	12/11
(d)	0.10	11.6	0.095	R	1/1
	0.80	108.0	0.69	A	10/7
	0.99	168.5	0.94	R	11/11

the θ_j were recovered exactly in STAGE I of the test. In case (d), the recovered θ_j are shown in case (e) of Fig. 5.1.

The case we were initially most interested in was case (c), and we shall discuss this first. The rotation angles were deliberately selected to be large, crowding around 90° from below 90°, as can be seen from Fig. 5.1. We

set $\alpha = 0.10$ and chose the acceptance fraction $\bar{f}_a$ to be 0.10. The Monte Carlo runs in STAGE III produced an $\bar{f}(\theta)$ curve which associated to $\bar{f}_a = 0.10$ the acceptance angle θ_a of 11.6°. This angle in turn, via the reference distribution of STAGE II, yielded up the acceptance probability $p_a = 0.095$. This p_a defined, via (5.10), a critical acceptance number a(distant) = 1. When the program computed the θ_j for $\underline{D}$ and $\underline{M}_3$, as described in STAGE I, it was found that none of the 12 θ_j was in $[0,\theta_a] = [0,11.6°]$. Since we needed 0 or 1 of the θ_j's in [0,11.6°] for rejection of H_o, the null hypothesis H_o (Step 5, STAGE IV) was accordingly rejected. We conclude, for this level of $\bar{f}_a$, that $\underline{E}$ and $\underline{F}_1$ are significantly distant. This was what we designed case (c) for and what we wanted, intuitively, the S-Phase to declare. We then raised $\bar{f}_a$ to 0.20 and still obtained rejection of H_o. H_o was rejected because 1 canonic rotation angle was found in [0, 24.0°] while a(distant) in this case was 2. At $\bar{f}_a = 0.60$ we we obtained the final rejection of H_o, and at $\bar{f}_a = 0.99$, H_o was accepted. Therefore, at about $\bar{f}_a = 0.60$ we have the acceptance/rejection threshold of H_o on the 90% (= (1-α)100%) confidence level. The reader will now see that, in case (c), had we crowded the θ_j more closely toward 90°, we could have raised $\bar{f}_a$ somewhere beyond 0.60 and perhaps still have had rejection of H_o, i.e. the test would have continued to declare the $\underline{E}$ and $\underline{F}_3$ frames significantly distant.*

---

* The dual result for this case (c) would occur in the *significantly-close option* using *case (b)*: Starting with (say) $\bar{f}_a = 0{\cdot}9$, the S-Phase test would reject H_o and declare $\underline{E}$ and $\underline{F}_3$ significantly close. This rejection of H_o would continue until $\bar{f}_a$ was decreased to around 0.50 or somewhat below, where there would occur an acceptance/rejection threshold. For $\bar{f}_a$ smaller than the threshold, H_o would be accepted. For a given α, the smaller $\bar{f}_a$ is before acceptance of H_o, the closer are the $\underline{E}$ and $\underline{F}$ frames.

In practice, i.e., with real data, this sweep of the S-Phase test parameter $\bar{f}_a$ over (0,1) starting near 0 would, for given α, give a succession of acceptances and rejections of H_o. The user of the S-Phase test would then receive an impression of the proximity of an $\underline{E}$ and $\underline{F}$ frame which would be a much more detailed impression than had he selected just one $\bar{f}_a$ and arrived at a single acceptance or rejection of H_o. In the significantly-distant option, and for a given α, the larger $\bar{f}_a$ is before acceptance of H_o, the more distant are the $\underline{E}$ and $\underline{F}$ frames. This many-look situation is similar to that in the APP (auto-cross permutation procedure) of DIT(II) where many looks at a statistic, such as SITES or SPRED, were made before declaring that the data sets were significantly separated or not in the sense of SITES or SPRED, respectively.

In case (a) we obtained acceptance of H_o all across the $\bar{f}_a$ range, except at $\bar{f}_a = 0.10$, indicating that, as far as the S-Phase tests analysis was concerned, the $\underline{E}$ and $\underline{F}_1$ frames could have been obtained from two samples drawn randomly out of $N_{24}(\underline{0},\underline{I}_{24})$.

In case (b) we obtained acceptance of H_o all across the $\bar{f}_a$ range. This is a reasonable result for the present option, and since we know, by construction, that $\underline{E}$ and $\underline{F}_2$ are relatively close, this suggests that a user of the S-Phase test would do well to try *both options* in Step 5, STAGE IV on a given pair of data sets. In the present case, the result of applying the significantly-close option is anticipated in the preceding footnote. As a user's experience with the S-Phase test grows, and a history of results accumulates using both options and a small set of α values, the proximity analysis of two new $\underline{E}$ and $\underline{F}$ frames would be viewable against the accumulated results of earlier tests and more than just an objective conclusion would result.

We turn finally to case (d). This case differs from the others because we have allowed the angles θ_j to spill over the 90° threshold in (0,180°).

After we produced $\underline{F}_4$ by rotating $\underline{E}$ the indicated amount, we sent $\underline{E}$ and $\underline{F}_4$ off to S-Phase test and observed the following: In STAGE I of the test an independent determination of the canonic rotation angles between $\underline{E}$ and $\underline{F}_4$ was made. The result is shown in Figure 5.1 (e) which we call the *recovered angles*. Unlike cases (a), (b), and (c), wherein the θ_j distribution was in each case recovered exactly, we obtained the distinctly different distribution shown in (e) when we started with that in (d). When this result appeared, we were momentarily puzzled. It eventually became clear what was happening: the program was correct; the octant condition was doing what it was designed to do, namely to flip the $\underline{e}_j$ vectors of the $\underline{D}$ matrix until $\underline{E}$ was changed to some new $\underline{E}'$ such that ORIEN($\underline{E}'$,$\underline{F}_4$) is a minimum (cf. §5, Appendix C, and the discussion of STAGE II in par. B, above). It turned out that this minimum was below the ORIEN($\underline{E}$,$\underline{F}_4$) value for the *original* $\underline{E}$. The net result is that, for $\bar{f}_a = 0.10$, the test rejected H_o, on the grounds of the presence of the single recovered angle of size 6.94° in [0,11.6°]. This angle, near 7°, can be seen in Fig. 5.1(e). Clearly the 1/1 fraction in Table 5.1 for this case is a borderline decision. As we continued to raise $\bar{f}_a$ we obtained a strong acceptance at $\bar{f}_a = 0.80$. Continuing on to $\bar{f}_a = 0.99$, we obtained a final (borderline) rejection of H_o. Recall that, in the significantly distant option, a rejection of H_o means the frames are significantly distant. Hence, in our sweep through the range of $\bar{f}_a$ values, we see the $\underline{E}$,$\underline{F}_4$ frames first as distant, then close, and then distant. Closer examination of all these results shows that, while at first puzzling, they are nevertheless reasonable: by rotating $\underline{D}$ out past the 90° threshold of rotation angles, we produced an $\underline{F}_4$ frame that could be transformed, via successive sign changes on its vectors $\underline{f}_j$, to a new frame $\underline{F}_4'$, of equivalent physical validity for representing $\underline{M}_4$, that was definitely

closer* to $\underline{D}$ than $\underline{F}_4$. However, the distribution of the recovered angles, although distinctly new, was still qualitatively like the original: concentrated around 90°, sparse near 0° and 180°.†

Despite this rationalized reasonableness of the result of the S-Phase test in case (d), we are left with a sense of overall disappointment in the test. For one thing, we have discovered that, under the octant condition, a continuous rotational transformation of a data matrix $\underline{D}$ can induce a discontinuous transformation of the canonic rotation angles θ_j. This is illustrated by the θ_j distributions in Fig. 5.1(d), (e). This is a property, however, that can be tolerated until something better is devised. (A possible remedy for this weakness of the test is to drop the octant condition and explore other ways of keeping the frames $\underline{E},\underline{F}$ uniquely defined. One such possibility is the "hemispheric" condition discussed in par. B above.)

While the S-Phase test seems to have the general configuration of a potentially good research tool, it is clear that its interpretation is somewhat complex, and not unlike reading tea leaves. Thus, it probably can be handled

---

* Since ORIEN($\underline{E},\underline{F}$) is symmetric in $\underline{E},\underline{F}$, the minimum ORIEN condition can be attained by acting on either E's or F's vectors.

† Nevertheless, observe that, in the process of continuously rotating $\underline{D}$ out past 90°, we encounter an abrupt change in the recovered canonic rotation angles of STAGE I. It might be noted that we could avoid the discontinuity of the recovered θ_j's by defining an *equivalent homogeneous canonic rotation angle* θ_e such that (on the basis of (C4.4)) $4(1-\cos\theta_e) \equiv (4/p) \sum_{k=1}^{\ell} (1-\cos\theta_k)$. θ_e varies continuously as E is rotated continuously. However, such a tactic returns us to the *single number* description of distance, and this is what the S-Phase test was intended to avoid in the first place.

by researchers with a bent for statistical niceties; but it is not for the researcher wanting a rough and ready "cook-book" answer to his data intercomparison problem. Perhaps some future research will take the basic ideas here and rework them into something more practicable.

6. T-Phase Test

We consider next the T-Phase test, which is designed to tell whether the $\underline{A}'$ and $\underline{B}'$ frames of the given data/model pair $\underline{D},\underline{M}$ (and hence their temporal evolution) are significantly close or distant, as the case may be. The central concept used in the test is the set of canonic correlation angles between $\underline{A}'$ and $\underline{B}'$. The theory of these angles is given in Appendix D, and how they enter the T-Phase test will be described in the discussion following the statement of the test. The general features of the T-Phase test have been made as close to those of the S-Phase test as their inherent difference allows. The inherent difference is in the dimensionalities of $\underline{A}',\underline{B}'$ vis a vis those of $\underline{E},\underline{F}$. The latter matrices are $p\times p$, while the former are $n\times p$. $\underline{E}$ and $\underline{F}$ therefore can be connected *uniquely* by a rotation in E_p whenever $n-1 \geq p$. Rotations in E_n, when $n-1 \geq p$, can always connect $\underline{A}',\underline{B}'$, but the rotations are not unique. Thus, a unique set of canonic rotation angles linking $\underline{A}',\underline{B}'$ is impossible whenever $n-1 \geq p$, a case which often occurs in practice, and the case on which the S-Phase test was built. When $n-1 < p$, it is possible to make a unique rotational connection between $\underline{A}',\underline{B}'$. In this case the formalism of the S-Phase test can be adopted once any n-1 of the linearly independent p columns in $\underline{A}'$ and any n-1 of the linearly independent p columns in $\underline{B}'$ are chosen as bases for E_{n-1}, the common space for these bases.* In view of this fact we will

---

* That the columns of $\underline{A}',\underline{B}'$ are vectors in a common E_{n-1} is demonstrated in Appendix D. See (D6.4), (D6.5).

dwell exclusively in this section on the more difficult of the two cases for the T-Phase test namely where $n-1 \geq p$. As noted above, the central concept needed in this case is the canonic correlation angle. In par. C below, we shall give an overview of the S-Phase and T-Phase tests.

A. T-Phase Test Stages

STAGE I (Canonic Correlation Angles)

1. Given: two $n \times p$ matrices $\underline{D}', \underline{M}'$. Center them in space to obtain $\underline{D}, \underline{M}$, respectively.
2. Find the $\underline{A}'$ and $\underline{B}'$ frames of $\underline{D}, \underline{M}$, repectively.
3. Construct $\underline{P}_\alpha = \underline{A}'\underline{A}'^{T}$ and $\underline{P}_\beta = \underline{B}'\underline{B}'^{T}$, and then $\underline{P}_\alpha \underline{P}_\beta$.
4. Find the canonic correlation angles $\psi_1, \ldots, \psi_p$ of $\underline{P}_\alpha \underline{P}_\beta$.

STAGE II (Reference Distributions for Canonic Rotation and Correlation Angles)

1. Construct pairs $\underline{N}^{(i)}, \underline{P}^{(i)}$ of orthonormal frames in E_n, of same chirality, $\ell = 1, \ldots, r$. (See also Step 3).
2. Find canonic rotation angles $\theta_1^{(i)}, \ldots, \theta_{[\frac{1}{2}n]}^{(i)}$ between $\underline{N}^{(i)}, \underline{P}^{(i)}$, $\ell = 1, \ldots, r$. Pool all $[\frac{1}{2}n]r$ angles; form a reference distribution from them.
3. During process in Step 1, select the first p orthonormal vectors from $\underline{N}^{(i)}$ and from $\underline{P}^{(i)}$ for each $i = 1, \ldots, r$, resulting in $n \times p$ frames $\underline{A}^{(i)}$, $\underline{B}^{(i)}$, respectively. Find canonic correlation angles between $\underline{A}^{(i)}, \underline{B}^{(i)}$, namely $\psi_1^{(i)}, \ldots, \psi_t^{(i)}$, $t = p - \max[0, 2p-(n-1)]$. Pool all tr correlation angles; form a reference distribution from them.

STAGE III (Connecting temporal rotations with D-field changes)

1. Construct rotation platform bases $\{\underline{x}_1^{(i)}\underline{y}_1^{(i)},\ldots,\underline{x}_m^{(i)}\underline{y}_m^{(i)}\}$ for the case $n = 2m$ or $\{\underline{x}_1^{(i)}\underline{y}_1^{(i)},\ldots,\underline{x}_m^{(i)},\underline{y}_m^{(i)},\underline{x}_{m+1}^{(i)}\}$ for the case $n = 2m+1$, $i = 1,\ldots,r$.
2. Produce $\bar{f}^{(i)}(\theta)$ vs θ, $i = 1,\ldots,r$, via homogeneous temporal rotations of data set $\underline{D}$ using the ith rotation platform.
3. Average the $\bar{f}^{(i)}(\theta)$ over the $i = 1,\ldots,r$ to obtain $\bar{f}(\theta)$.

STAGE IV (Deciding whether $\underline{A}',\underline{B}'$ are significantly close or distant)

1. Choose an $\bar{f}_a$ value, $0 \leqq \bar{f}_a \leqq 1$. Using $\bar{f}(\theta)$ of STAGE III, find that θ_a which solves $\bar{f}(\theta) = \bar{f}_a$. θ_a is the *acceptance rotation angle*, and $\bar{f}_a$ is the *acceptance fraction*.
2. Use the reference distribution in Step 2, STAGE II to find p_a that pairs with θ_a. p_a is the *acceptance probability*.
3. Use the reference distribution in Step 3, STAGE II to determine *acceptance correlation angle* ψ_a from p_a.
4. Choose *confidence level* $1-\alpha$ of test, α is the *size* of the test.
5. Compute from p_a and α the *critical acceptance number* a(close) or a(distant), as the case may be.
6. $\underline{A}'$ is significantly close (significantly distant from) $\underline{B}'$ if the number a of canonic correlation angles $\psi_1,\ldots,\psi_p$ (of Step 4, STAGE I) in $[0,\psi_a]$ is greater than a(close) (is less than or equal to a(distant)). In each case a decision results in the rejection of the null hypothesis H_o that the $\underline{A}',\underline{B}'$ frames of $\underline{D},\underline{M}$ in STAGE I are drawn from the same population as the $\underline{A}^{(i)},\underline{B}^{(i)}$ frames in STAGE II.

B. Discussion of T-Phase Test Stages

STAGE I (Discussion)

The theory of singular value decompositions of data sets, used to find $\underline{A}'$ and $\underline{B}'$, is as described in Appendix B. The canonic correlation angle theory is given in Appendix D. A standard IMSL eigenvalue routine will find the eigenvalues of $\underline{P}_\alpha\underline{P}_\beta$, and from these a simple algorithm finds the corresponding correlation angles.

STAGE II (Discussion)

This stage differs from that in the S-Phase test by going directly to orthonormal frames in E_n, rather than mimicking $\underline{D}^{(i)},\underline{M}^{(i)}$ (as in S-Phase), to find the $n\times p$ $\underline{A}^{(i)},\underline{B}^{(i)}$ frames. This is apparently an unavoidable step, and its role in the T-Phase test will be clear once STAGE IV has been discussed. We note, in passing this point, that we could modify STAGE II of S-Phase to look like Step 1 of this stage without any loss of generality to S-Phase. In this way the two tests, in STAGE II at any rate, can be structured more closely.

STAGE III (Discussion)

This stage is similar, in all details, to STAGE III of the S-Phase test. Now, of course, we use the $n\times n$ temporal rotation $\underline{R}^{(i)}(\theta)$ to map $\underline{D}$: $\underline{R}^{(i)}(\theta)\underline{D} \equiv \underline{D}^{(i)}(\theta) = [\underline{R}^{(i)}(\theta)\underline{A}']\underline{K}^{\frac{1}{2}}\underline{E}^T$. The $\underline{D}^{(i)}(\theta)$ are now processed just as in the S-Phase test.

STAGE IV (Discussion)

The use of the three distributions in this stage can be seen in Figure 6.1. By means of the STAGE III homogeneous temporal rotations in E_n we find the

FIXING θ_a, p_a, AND ψ_a IN THE T-PHASE TEST

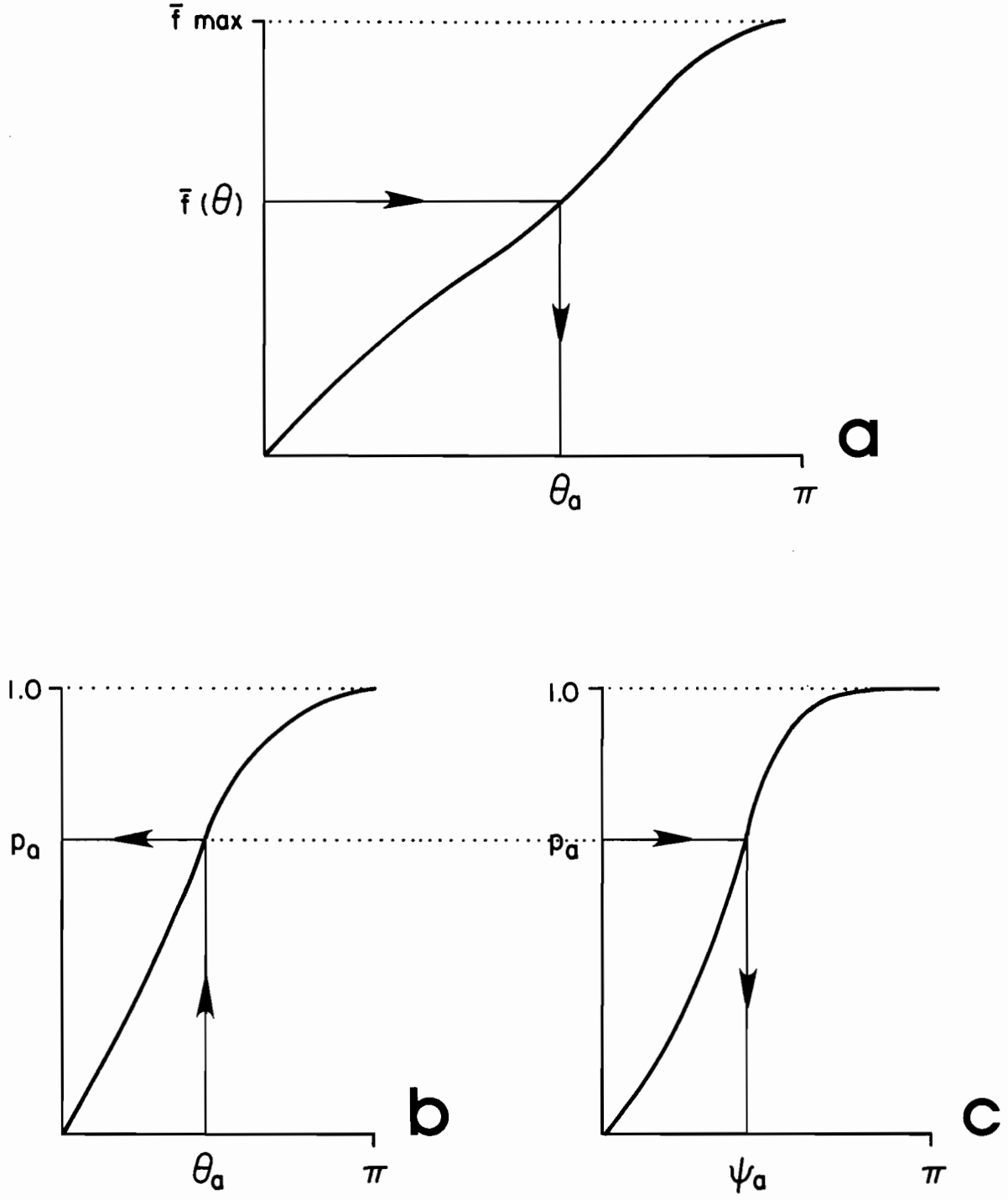


Fig. 6.1

curve in Fig. 6.1(a). Choosing $\bar{f}_a$ then fixes θ_a. This θ_a in turn fixes p_a via the curve in Fig. 6.1(b), coming from Step 2, STAGE II. This p_a is then used to find ψ_a in Fig. 6.1(c), using the curve obtained in Step 3, STAGE II. This roundabout way to link p_a with ψ_a is unavoidable because we have *rotations* in E_n that must generate $\bar{f}(\theta)$. Somehow, then, these θ have to be linked with ψ, and that is the task of the (p_a,θ_a) in Fig. 6.1(b). Of course it is in principle easy to think up ways to link p_a and ψ_a directly: imagine a data set available for use in which a goodly supply of n-samples of p-variate fields can be obtained. The resultant $\underline{A}',\underline{B}'$ sets from each sample will then collectively produce a reference distribution of canonic correlation angles. From this an acceptance interval $[0,\psi_a]$ can be produced. The ψ_j of $\underline{D},\underline{M}$ can then be tested for membership in $[0,\psi_a]$ and the number of these angles in $[0,\psi_a]$ can be gauged for significance (in either the *close* or *distant* option). What has made this attractively direct procedure possible is, of course, an *adequate* data setting (cf. DIT(II)). What we are struggling with (by choice) in the present version of T-Phase is an *inadequate* data setting (cf. DIT(II)), to show how one can do something along the line of a T-Phase test in the poorest of settings. a(close), a(distant) in the T-Phase test are reckoned as in (5.8), (5.9), (5.10).

C. Overview of S-Phase and T-Phase Tests

In the present study, we have introduced the S-Phase and T-Phase tests in specific physical contexts and perhaps have thereby given the impression that one test should always be used to tell if spatial patterns are close or distant and the other to tell if the temporal evolutions of data sets are close or distant. In this paragraph we will try to dispel any such impressions and point out the dual structure of these tests, and the general nature of each,

but with a minimum of mathematics. If at some later time these tests will have proven to be useful, then perhaps it will be worthwhile for a general mathematical formalism to be developed to free each test from its specific origins.

The first hint that each of the two tests is not restricted to just spatial or just temporal patterns occurs in the introduction to each test. The S-Phase and T-Phase tests were each stated for the case $n-1 \geq p$, and hints were made as to the treatment of the complementary case $n-1 < p$. The reader, having perused Appendix B, will then have discerned that the relative sizes of $n-1$ and p are critical to the matters of degenerate eigenvalues and to non-unique eigenvectors in the singular value decomposition of an $n \times p$ data matrix $\underline{D}$. These matters directly concern the conduct of each test. When $n-1 \geq p$ and n pertains to time samples and p to spatial points at which the samples are taken, then it is natural to call this the *nondegenerate case*, while $n-1 < p$ denotes the *degenerate case*. The S-Phase and T-Phase tests above are written for the nondegenerate case.

In the degenerate case, the S-Phase test is to be conducted along the lines of the nondegenerate-setting T-Phase test: select from the p columns of $\underline{E}$ and $\underline{F}$ $n-1$ linearly independent columns, and these will now be treated as if they were $\underline{A}'$ and $\underline{B}'$ frames. Dually, in the degenerate case, the T-Phase test is conducted along the lines of the nondegenerate-setting S-Phase test: select from the p columns of $\underline{A}'$ and $\underline{B}'$ $n-1$ linearly independent columns and these will now be treated as if they were $\underline{E}$ and $\underline{F}$ frames.

Out of these observations comes the realization that the essence of an S-Phase test is testing the closeness of two p-member *orthonormal frames* in some p-dimensional space E_p, while the essence of a T-Phase test is testing the closeness of two p-member orthonormal frames in some q-dimensional space E_q, where $p < q$. It doesn't matter to these tests whether the interpretation

of a vector in the frame at hand is that of a principal component time series or a principal spatial vector (empirical orthogonal function) of a data set.

7. Some Research Problems

Our findings in §4 indicate that comparison techniques for high-dimensional multivariate data sets require multiparameter structure; nothing less will serve as a reliable research tool. The lesson of §4 was that a single parameter test (e.g., via correlations or norms) simply lost too much of the multivariate information being fed into it for the results to be fully informative. This does not mean we advocate discarding correlation- or norm-type statistics in the multivariate setting. We shall continue to use them and draw inferences by means of them; but we intend to continue the search for multiparameter tests such as the S-Phase and T-Phase tests of §§5, 6, and beyond. In this section we outline some potentially powerful multivariate tests especially concerned with tests of significance of differences between principal component time series, eigenvalue sequences, and principal vectors (EOF's); that is, statistical significance tests for intercomparisons of the three main classes of objects arising from a singular value decomposition of data sets.

A. Intercomparing factors $\underline{A}', \underline{K}^{\frac{1}{2}}, \underline{E}$ and $\underline{B}', \underline{L}^{\frac{1}{2}}, \underline{F}$

The SVD decomposition of n×p data sets $\underline{D}$ and $\underline{M}$ results in comparable objects such as $\underline{A}', \underline{B}'$ and $\underline{K}^{\frac{1}{2}}, \underline{L}^{\frac{1}{2}}$, and $\underline{E}, \underline{F}$ (see Appendix B). Their intercomparisons by means of inner products (correlations) are summarized in Table 2.1. Besides this form, we can compare these objects in more detail, as summarized in the following list.

(i) Intercomparison of the normalized principal components in $\underline{A}' = [\underline{\alpha}_1 \ldots \underline{\alpha}_p]$ and $\underline{B}' = [\underline{\beta}_1,\ldots,\underline{\beta}_p]$. Specifically, compare the n values $\alpha_j(1)$, $\alpha_j(2),\ldots,\alpha_j(n)$ of $\underline{\alpha}_j$ with those $\underline{\beta}_j$, i.e., $\beta_j(1)$, $\beta_j(2),\ldots,\beta_j(n)$.

(ii) Intercomparison of normalized eigenvalues $\kappa_1,\ldots,\kappa_p$ of $\underline{K}$ with the normalized eigenvalues $\lambda_1,\ldots,\lambda_p$ of $\underline{L}$.

(iii) Intercomparison of the principal vectors in $\underline{E} = [\underline{e}_1 \ldots \underline{e}_p]$ and $\underline{F} = [\underline{f}_1 \ldots \underline{f}_p]$. Specifically, compare the p values $e_j(1)$, $e_j(2),\ldots,e_j(p)$ of $\underline{e}_j$ with those of $\underline{f}_j$, i.e., $f_j(1)$, $f_j(2),\ldots,f_j(p)$.

One possible useful method of intercomparison is the r-tile method, which we now turn to consider.

B. Intercomparisons by r-tiles Using Various Procedures for Reference Distribution

We shall illustrate one method of data intercomparison which is based on the *r-tile classification* of the range of values of each of the variates in the object of interest. To be specific, suppose we are to compare a p-dimensional unit vector $\underline{e}$ with another p-dimensional unit vector $\underline{f}$, when both are measured within a natural-basis frame defined by $\{\underline{u}_1,\ldots,\underline{u}_p\}$, i.e.,

$$\underline{e} = \sum_{j=1}^{p} e_j \underline{u}_j \tag{7.1}$$

$$\underline{f} = \sum_{j=1}^{p} f_j \underline{u}_j \tag{7.2}$$

Imagine the components e_j of $\underline{e}$ to be formed from variates randomly drawn from N(0,1) and then normalized. It is shown in Appendix E that e_j is distributed on [-1,1] in accordance with the probability density function described in

(E2.6). This density is the same for all p components of $\underline{e}$. Moreover, if $\underline{f}$ is constructed in the same way, then its components, too, will be distributed as defined in (E2.6). It is therefore possible to partition the common range [-1,1] of these components into r subintervals $[-1,x_1],[x_1,x_2],\ldots,[x_{r-1},1]$, so that the probability of a component falling in any one of these subintervals is 1/r. The right boundary x_j of the jth subinterval is defined by

$$\int_{-1}^{x_j} q(x)dx = j/r \quad , \quad j = 1,\ldots,r-1 \tag{7.3}$$

where q(x) is defined in (E2.6). The r classification subintervals so defined form the basis for the r-tile intercomparison method.

There are five broad procedures we may now use to apply the r-tile method. These methods supply the requisite reference distributions for various statistics arising in the r-tile method. These procedures are defined in DIT(II) as:

IOP: Ideal observation procedure

EOP: Empirical observation procedure

APP: Auto-cross permutation procedure

PPP: Pool-permutation procedure

CIP: Classic intercomparison procedure

We shall briefly consider these in turn.

(i) In IOP we would have available for our reference distribution constructions an adequate data base, i.e., a sufficiently rich collection $\underline{D}^{(i)},\underline{M}^{(i)}$ i = 1,...,N of comparable n×p data matrices. From each pair $\underline{D}^{(i)},\underline{M}^{(i)}$ we would find $\underline{E}^{(i)}$ and $\underline{F}^{(i)}$, and in particular we would choose for intercomparison two vectors $\underline{e}^{(i)}$ and $\underline{f}^{(i)}$. The first component $e_1^{(i)}$ of $\underline{e}^{(i)}$ would then fall in one of the r-tile subintervals defined via (7.3), the second component $e_2^{(i)}$ will fall in

one of the r-tile subintervals, and so on. All p components of $\underline{e}^{(i)}$ will be classified in this way. Similarly, the p components of $\underline{f}^{(i)}$ will be classifiable. Fig. 7.1 (upper) shows the results of classification of $\underline{e}^{(i)},\underline{f}^{(i)}$ for the case of r = 3, p = 7, and Fig. 7.1 (lower) shows another example for the case of r = 5, p = 7.

When both vectors' components are plotted in a common r-tile setting, as in Fig. 7.1 (upper), e.g., then we can tally the r-class errors they subtend. The 0-class error u_0 associated with Fig. 7.1 (upper) is 2, since two times out of the seven the components e_j,f_j landed in the same subinterval of [-1,1]. The 1-class error in (a) is $u_1 = 3$ since there are three occasions where e_j,f_j are one class apart; and the 2-class error $u_2 = 2$. The sum of these errors is $u_0+u_1+u_2 = 2+3+2 = 7 = p$. In case of Fig. 7.1 (lower), we can tally the r-tile errors similarly. Thus $u_0 = 2$, $u_1 = 2$, $u_2 = 1$, $u_3 = 1$, $u_4 = 1$. In general for an r-tile classification, the r-tile errors are $u_0,u_1,\ldots,u_{r-1}$, and these add up to p.

In an adequate (i.e., IOP) data setting, such as the present one, it is a straightforward matter to estimate the probability distribution functions of the j-class errors u_j. Let $u_j^{(i)}$ be the *j-class error* of $\underline{e}^{(i)},\underline{f}^{(i)}$ associated with the ith sample $\underline{D}^{(i)},\underline{M}^{(i)}$ from the data collection. Then, for fixed j we would arrange the $u_j^{(i)}$ in ascending order (after relabeling):

$$u_j^{(1)} < u_j^{(2)} < \cdots < u_j^{(r)} \qquad \begin{matrix} r = 100 \text{ usually,} \\ j = 1,\ldots,r-1 \end{matrix} \tag{7.4}$$

This would define the 05% and 95% critical values $u_j^{(05)},u_j^{(95)}$, say, if r = 100 (the usual value chosen in adequate settings). From these we could make r inferences about any unit vectors such as $\underline{e}',\underline{f}'$ that would subsequently come up for comparison. Thus we would find the j-class errors $u_0',\ldots,u_{r-1}'$ subtended by $\underline{e}',\underline{f}'$ in the manner explained above. Then we would compare u_0' with $u_0^{(05)},u_0^{(95)}$,

INTERCOMPARISON OF UNIT VECTORS $\underline{e}, \underline{f}$ IN E_7 USING TERCILE AND QUINTILE CLASSIFICATIONS OF THEIR COMPONENT RANGES

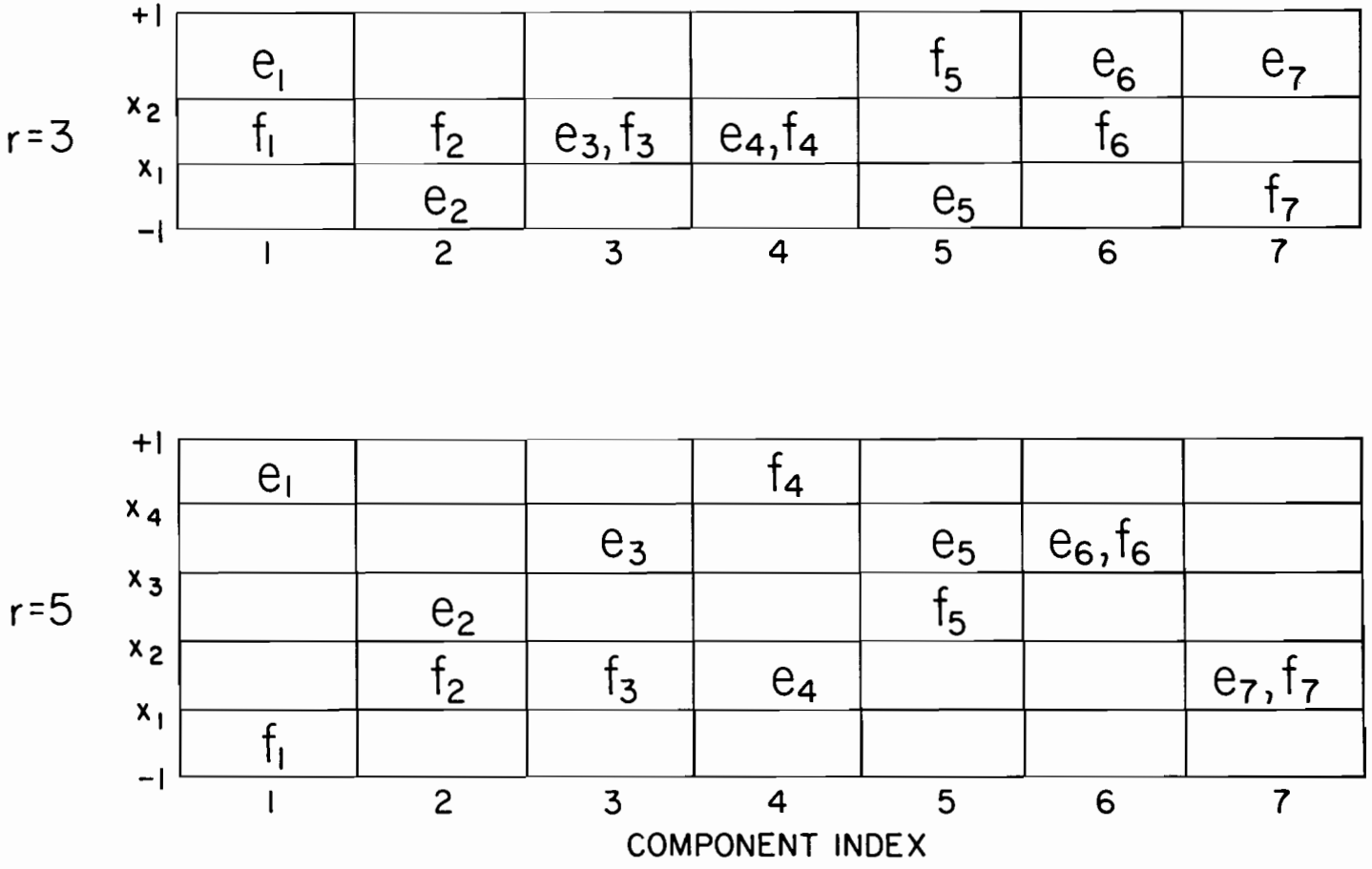


Fig. 7.1

and decide if $\underline{e}', \underline{f}'$ were significantly distant or close. For example, if $u_0^{(95)} < u_0'$, then we would say with confidence 95% that $\underline{e}', \underline{f}'$ are *significantly close* on the 0-class level. Furthermore, if $u_j' < u_j^{(05)}$, $j = 1,\ldots,r-1$, then $\underline{e}', \underline{f}'$ are with confidence 95% *significantly close* on the j-class level. The more j-class levels on which $\underline{e}', \underline{f}'$ are significantly close, then the greater is the belief that $\underline{e}', \underline{f}'$ are indeed close. The greatest significantly-close-type score achievable in an r-tile test is $u_0 = p$ and $u_j = 0$ for $j = 1,\ldots,r-1$.

When deciding on the *significantly-distant* option, we of course would want $u_o' < u_o^{(05)}$, and $u_j^{(95)} < u_j'$, $j = 1,\ldots,r-1$ and the greatest significantly-distant-type score would be $u_j = 0$, $j = 0,\ldots,r-2$; and $u_{r-1} = p$.

A measure of distance, natural to the r-tile method, is the *moment*

$$m = \sum_{j=1}^{r-1} j u_j \quad . \tag{7.5}$$

This gives the linear distance* between $\underline{e}$ and $\underline{f}$. When estimating the distribution functions for the u_j, that for the moment m should also be estimated. Clearly the smaller m is, the closer are the unit vectors $\underline{e}, \underline{f}$. Hence for statistically-close tests we would use the left tail of m's distribution.

If we had to choose only two measures of distance to gauge the closeness of $\underline{e}, \underline{f}$ when $r > 3$, then these would be u_0 and m. For $r = 3$, we would choose u_0 and m or u_1 and m.† One chooses r in accordance with available data (generally the more data available (i.e., the larger N is) the larger r can be when

---

* For r = 3, the idea of m can be obtained by thinking of distance between two corners in a rectangular street network where north-south blocks are twice the length of east-west blocks.

† For examples of the case r = 3 (terciles), see Preisendorfer and Mobley (1982), and Preisendorfer (1977).

compiling the distributions). Given enough data, one chooses r also on the basis of how finely the range [-1,1] is to be subdivided, and on how adept the researcher is in visualizing and making use of multiparameter measures of distance.

(ii) In EOP, APP, PPP, the data matrices $\underline{D}^{(i)}, \underline{M}^{(i)}$, $i = 1,\ldots,r$ are handled in the manners described in DIT(II). In each procedure the j-class errors and moment m for a given data/model pair $\underline{D},\underline{M}$ can be determined. The end result in each procedure is a reference distribution for u_j, $j = 0,\ldots,r-1$, and moment $m = \sum_{j=1}^{r-1} j u_j$. These reference distributions can then be applied to the problem of judging whether or not two given unit vectors $\underline{e}', \underline{f}'$ in the $\underline{E},\underline{F}$ frames of $\underline{D},\underline{M}$, respectively, are significantly distant or close. This decision process was described in detail in (i), just above. The reader can now see the rather large set of possibilities he may resort to in his quest for a reference background for the j-class errors u_j and moment m.

(iii) We consider finally the Classical Intercomparison Procedure. There are several possibilities here, too, that would be of interest to pursue in future research.

First of all for the purpose of finding the reference distributions of the u_j and m, one can postulate the Stochaster: a person or device that places a marker randomly in some r-tile slot for each of the p component indexes. Thus, in Fig. 7.1 (upper), the Stochaster will randomly place a point in one of the three classes above index 1. Then quite independently of that action the Stochaster would randomly place a point in one of the three classes above index 2, and so on, for all p indexes. The probability of the point falling in any one of the classes above each index is, by definition of

the Stochaster, 1/r. It then follows that the joint probability for a set $(u_0,\ldots,u_{r-1})$ of j-class errors subtended by two p-sets of marks placed by the Stochaster is*

$$P(u_0,u_1,\ldots,u_{r-1}) = \left(\frac{p!}{u_0!u_1!\ldots u_{r-1}!}\right) a_0^{u_0}\, a_1^{u_1} \ldots a_{r-1}^{u_{r-1}} \qquad (7.6)$$

where

$$a_0 = 1/r$$

$$a_j = 2(r-j)/r^2 \quad , \quad j = 1,\ldots,r-1$$

$$\sum_{j=o}^{r-1} a_j = 1 \quad , \quad \sum_{j=o}^{r-1} u_j = p$$

Therefore, by adopting the CIP, we completely open up the analytic possibilities regarding reference distributions for the u_j. In particular the distribution for u_j is:

$$P(u_j) = \frac{p!}{u_j!(p-u_j)!}\; a_j^{u_j}(1-a_j)^{p-u_j} \qquad (7.7)$$

$$j = 0,\ldots,r-1.$$

The distributions for u_j, so laboriously obtained by the other four procedures, as sketched above, are now generously supplied by classical statistical procedures, providing one can justify the underlying assumptions giving rise to (7.6), (7.7). In the present study this means asking whether the ith and jth components of randomly produced unit vectors (as in Appendix E) are statistically *independent*. For it is the hypothesized statistical *independence* of the

---

* See e.g. Preisendorfer (1977), p. 10.

Stochaster's component-to-component actions that allows (7.6) or (7.7) to be derived. Are the components of randomly-produced unit vectors pairwise independent? The answer clearly is "no," as a perusal of (E2.2) will immediately show. Here, then, we are looking into the *degree* of statistical dependence of the components of $\underline{e}$ (or of $\underline{f}$). If this degree is sufficiently small, then (7.6) and (7.7) may be nearly correct and descriptions of the u_j distributions therefore may be usefully applied to the unit vector intercomparison problem. Of course, we may simply use (as has already been done, e.g., in Preisendorfer and Mobley (1982)) the Stochaster's reference background supplied by (7.6) (for r = 3) to gauge the significance of separation of spatial patterns in general--patterns that themselves have nonzero spatial correlations within them. These observations thus lead us to the following list of research questions concerning reference distributions.

C. Research on Reference Distributions

Let two random unit vectors $\underline{e},\underline{f}$ be generated by the method of §2 in Appendix E. Thus their components are generated by means of random samplings of N(0,1).

(i) What is the joint probability density of the jth and kth components of $\underline{e}$?

(ii) What is the probability density function for the jth canonic direction angle ϕ_j of $\underline{e}$? (§7 of Appendix C)

(iii) Are the jth and kth canonic direction angles ϕ_j,ϕ_k of $\underline{e}$ statistically independent?

(iv) What is the probability density function of $\phi_k-\phi_k'$ where ϕ_k,ϕ_k' are the kth canonic direction angles of $\underline{e},\underline{f}$, respectively? (See discussion below Table C7.1 in Appendix C.)

(v) How well can the distributions in (i) be represented by normal distributions (of one or two variates, respectively)?

These questions may be sought analytically or numerically. We conjecture that in (ii) the distribution is uniform, and that the ϕ_j are independent of one another in (iii). The questions may be extended to normalized principal component vectors $\underline{\alpha},\underline{\beta}$, and to eigenvalue vectors $\underline{\kappa},\underline{\lambda}$ (Table 2.1 and (2.12)). Suitably phrased, these questions may be extended to p×p orthonormal frames of vectors represented either by vector components or by the $\frac{1}{2}p(p-1)$ canonic direction angles (Appendix C). The reader will observe that the main thrust of the questions is toward the canonic direction angles. In the likely event of an affirmative answer to question (iii), the theory of the Stochaster, embodied in (7.6), would be rigorously applicable, resulting in practical intercomparison tests for EOF's, their principal components, and even their normalized eigenvalue vectors $\underline{\kappa}$ and $\underline{\lambda}$.

8. References

Works in this series on Data Intercomparison Theory (NOAA Technical Memorandums, ERL-PMEL) are:

DIT(I): Minimal Spanning Tree Tests for Location and Scale Differences

DIT(II): Trinity Tests for Location, Spread and Pattern Differences

DIT(III): S-Phase and T-Phase Tests for Spatial Pattern and Temporal Evolution Differences

DIT(IV): Tercile Tests for Location, Spread and Pattern Differences

DIT(V): Case Study: Effects of Objective Analysis on a Tropical Pacific Sea Surface Temperature Set

Halmos, P. R. (1958) Finite-Dimensional Vector Spaces, D. Van Nostrand, N.Y.

Preisendorfer, R. W. (1977) Climate Forecast Verification via Multinomial Stochasters. SIO Ref 77-33, Scripps Inst. Ocn., La Jolla, CA.

Preisendorfer, R. W. (1979) Model Skill and Model Significance in Linear Regression Hindcasts. SIO Ref 79-12, Scripps Inst. Ocn., La Jolla, CA.

Preisendorfer, R. W., C. D. Mobley (1982) Climate Forecast Verifications, U.S. Mainland, 1974-82, NOAA Tech. Memo ERL PMEL-36.

Rao, C. R. (1973) Linear Statistical Inference and its Applications, 2nd Edition, N.Y.

APPENDIX A

Inner Products and Norms of Data Sets

1. Introduction

In this Appendix we outline the theory of inner products and norms of data sets, concepts which lie at the base of (euclidean-space) data intercomparison theory. In the view of the present Appendix, a pair of n×p data sets $\underline{D}$ and $\underline{M}$ in matrix form is visualizable as two vectors in E_{np}, which therefore have an inner product between them, and also their difference $\underline{D}-\underline{M}$ has an associated distance or norm. Inner product is intimately related to correlation, and one of the main conclusions the machinery of the present Appendix allows us to reach (cf. §3 of the main text) is that the SHAPE statistic and its descendants are analyzable into various forms of the correlation statistic. The latter has a well-established statistical-inferencial procedure, and so data intercomparison problems centering on the SHAPE-like statistics of data sets can be reduced to significance decisions about correlations (cf §§4,7 of the main text).

2. Inner Product

Let $\underline{X},\underline{Y}$ be two n×p matrices. Let us write

$$'(\underline{X},\underline{Y})' \text{ for } \sum_{j=1}^{p}\sum_{i=1}^{n} x(i,j)y(i,j) \qquad \text{(A2.1)}$$

where

$$\underline{X} = \{x(i,j):\ i = 1,\ldots,n;\ j = 1,\ldots,p\}$$

$$\underline{Y} = \{y(i,j):\ i = 1,\ldots,n;\ j = 1,\ldots,p\}$$

The number (X,Y) is the *inner product* of $\underline{X}$ and $\underline{Y}$. It is the direct generalization of the inner product of two n×1 matrices (vectors) $\underline{u},\underline{v}$:

$$(\underline{u},\underline{v}) = \sum_{i=1}^{n} u(i)v(i) = \underline{u}^T\underline{v} \tag{A2.2}$$

The last equality uses the customary notation for $(\underline{u},\underline{v})$ when we work with vectors. Equation (A2.1) can be placed into this classical form $\underline{u}^T\underline{v}$ provided we write

$$\text{'trace } \underline{A}' \quad \text{for} \sum_{i=1}^{p} a(i,i) \tag{A2.3}$$

where $\underline{A}$ is any p×p matrix $\{a(i,j): i,j = 1,\ldots,p\}$. Then we can write (A2.1) as

$$(\underline{X},\underline{Y}) = \text{trace } \underline{X}^T\underline{Y} \tag{A2.4}$$

where "T" as usual denotes the transpose operator. By direct computation we observe that:

$$\begin{aligned} &\text{trace } \underline{X}^T\underline{Y} = \text{trace } \underline{X}\,\underline{Y}^T \\ &= \text{trace } \underline{Y}\,\underline{X}^T = \text{trace } \underline{Y}^T\underline{X} \end{aligned} \tag{A2.5}$$

Hence (A2.4) is independent of the order of $\underline{X}$ and $\underline{Y}$ and of whether the transpose operation "T" is taken on $\underline{X}$ or $\underline{Y}$.

Either from the basic form (A2.1) or from (A2.4), we can prove the following properties enjoyed by inner products of data sets: let $\underline{X},\underline{Y},\underline{Z}$ be n×p data sets, and a an arbitrary real number. Then

$$(\underline{X},\underline{Y}) = (\underline{Y},\underline{X}) \tag{A2.6}$$

$$(a\underline{X},\underline{Y}) = a(\underline{X},\underline{Y}) \tag{A2.7}$$

$$(\underline{X}+\underline{Y},\underline{Z}) = (\underline{X},\underline{Z}) + (\underline{Y},\underline{Z}) \tag{A2.8}$$

The *norm* $\|\underline{X}\|^2$ is defined as

$$\|\underline{X}\|^2 \equiv (\underline{X},\underline{X}) \tag{A2.9}$$

We may think of $\|\underline{X}\|$ as the *length* of $\underline{X}$. The basic properties of inner product and length as we know them from E_3 and ordinary vectors in E_n hold also in E_{np}:

$$\|\underline{X}\| = 0 \quad \text{if and only if } \underline{X} = \underline{0} \tag{A2.10}$$

$$\|a\,\underline{X}\| = |a|\ \|\underline{X}\| \text{ , for any real number } a \tag{A2.11}$$

(triangle inequality) $$\|\underline{X}+\underline{Y}\| \leq \|\underline{X}\| + \|\underline{Y}\| \tag{A2.12}$$

(Schwarz inequality) $$|(\underline{X},\underline{Y})| \leq \|\underline{X}\|\ \|\underline{Y}\| \tag{A2.13}$$

$$\|\underline{X}-\underline{Y}\|^2 = \|\underline{X}\|^2 + \|\underline{Y}\|^2 - 2(X,Y) \tag{A2.14}$$

If $\|\underline{X}\| = \|\underline{Y}\| = 1$, then

$$0 \leq \|\underline{X}-\underline{Y}\|^2 = 2[1-(\underline{X},\underline{Y})] \leq 4 \tag{A2.15}$$

This implies (consistently with (A2.13)) that for unit length $\underline{X},\underline{Y}$,

$$|(\underline{X},\underline{Y})| \leq 1 \tag{A2.16}$$

Therefore if $\underline{X},\underline{Y}$ have unit length we can assign a real angle θ_{XY} between them, $0 \leq \theta_{XY} \leq \pi$, such that

$$(\underline{X},\underline{Y}) \equiv \cos\theta_{XY} \tag{A2.17}$$

In general, (A2.14) may be written

$$\|\underline{X}-\underline{Y}\|^2 = \|\underline{X}\|^2 + \|\underline{Y}\|^2 - 2\|\underline{X}\|\|\underline{Y}\|\cos\theta_{XY} \tag{A2.18}$$

where $\cos\theta_{XY}$ is $(\underline{X}/\|\underline{X}\|, \underline{Y}/\|\underline{Y}\|)$. Also, (A2.15) becomes, for unit length $\underline{X},\underline{Y}$,

$$\|\underline{X} - \underline{Y}\|^2 = 2[1-\text{trace}\underline{X}^T\underline{Y}] = 2[1-\text{trace}\underline{X}\,\underline{Y}^T] \qquad \text{(A2.19)}$$
$$= 2[1-(\underline{X},\underline{Y})] = 2[1-\cos\theta_{XY}]$$

This formula is characteristic of the norm of $\underline{X}-\underline{Y}$ for two unit vectors $\underline{X},\underline{Y}$ in E_{np}. It is the direct generalization of the norm of $\underline{u}-\underline{v}$ of two $n\times 1$ unit vectors in E_n:

$$\|\underline{u}-\underline{v}\|^2 = 2[1-\underline{u}^T\underline{v}] = 2[1-\text{trace}\,\underline{u}\,\underline{v}^T] \qquad \text{(A2.20)}$$
$$= 2[1-(\underline{u},\underline{v})] = 2[1-\cos\theta_{uv}]$$

The *correlation coefficient* of two average-centered* data sets $\underline{u} = \{u(i): i = 1,\ldots,n\}$, $\underline{v} = \{v(i): i = 1,\ldots,n\}$ is:

$$r(\underline{u},\underline{v}) = \frac{\sum_{i=1}^{n} u(i)v(i)}{\left(\sum_{i=1}^{n} u^2(i)\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} v^2(i)\right)^{\frac{1}{2}}} \qquad \text{(A2.21)}$$

and is therefore seen to be simply $\tilde{\underline{u}}^T\tilde{\underline{v}} = (\tilde{\underline{u}},\tilde{\underline{v}})$, where $\tilde{\underline{u}},\tilde{\underline{v}}$ are the unit vectors associated with $\underline{u},\underline{v}$, respectively. Thus *norm*, *inner product*, and *correlation* are intimately connected in (A2.19) and (A2.20).

Finally, we observe that if the $q\times n$ matrix $\underline{B}$ is such that $\underline{B}^T\underline{B} = \underline{I}_n$, and the $p\times q$ matrix $\underline{C}$ is such that $\underline{C}\,\underline{C}^T = \underline{I}_p$, then for all $n\times p$ matrices $\underline{X},\underline{Y}$,

$$(\underline{B}\,\underline{X}, \underline{B}\,\underline{Y}) = (\underline{X},\underline{Y}) \qquad \text{(A2.22)}$$

$$(\underline{X}\,\underline{C}, \underline{Y}\,\underline{C}) = (\underline{X},\underline{Y}) \qquad \text{(A2.23)}$$

---

* i.e. $\sum_{i=1}^{n} u(i) = 0$, $\sum_{i=1}^{n} v(i) = 0$.

These formulas follow at once from (A2.4), and amount to the assertions that inner products and norms are invariant under change of bases or rotations in data spaces. It is worth noting that these relations hold also for ordinary vectors. Thus, e.g., if $\underline{R}$ is a p×p matrix such that $\underline{R}^T\underline{R} = \underline{I}_p$ (i.e., if $\underline{R}$ is a rotation) and $\underline{u},\underline{v}$ are vectors in E_p, then (A2.22), (A2.23) become

$$(\underline{R}\ \underline{u})^T(\underline{R}\ \underline{v}) = \underline{u}^T\underline{v} \tag{A2.24}$$

$$\text{trace}(\underline{u}^T\underline{R})^T(\underline{v}^T\underline{R}) = \underline{u}^T\underline{v} \tag{A2.25}$$

A general related result, that is often useful, is:

$$\text{trace } \underline{B}^T\underline{X}\ \underline{B} = \text{trace } \underline{B}\ \underline{X}\ \underline{B}^T = \text{trace } \underline{X} \tag{A2.26}$$

for any p×p matrix $\underline{X}$ and any p×p matrix $\underline{B}$ of orthonormal vectors.

APPENDIX B

Singular Value Decompositions (SVD) of Data Sets

1. Introduction

The singular value decomposition (SVD) of a space-centered n×p data matrix $\underline{D} = \{d(t,x): \ t = 1,\ldots,n; \ x = 1,\ldots,p\}$ allows us to split the space-time structure of $\underline{D}$ neatly into three parts $\underline{A}'$, $\underline{K}^{\frac{1}{2}}$, and $\underline{E}$

$$\underline{D} = \underline{A}'\underline{K}^{\frac{1}{2}}\underline{E}^{T} \tag{B1.1}$$

which, respectively, describe the *temporal evolution* of the data set (via $\underline{A}'$), its *variance structure* (via $\underline{K}$), and the *spatial pattern* (via $\underline{E}$). The SVD forms the basis for the various intercomparison statistics described in the text. In view of the use of (B1.1) in calculating SHAPE, we shall develop the SVD of $\underline{D}$ for the case of the standardized form of $\underline{D}$ (cf. (B2.8)).

2. Standardized Data Sets

Let $\underline{D}' = \{d'(t,x): \ t = 1,\ldots,n; \ x = 1,\ldots,p\}$ be an n×p data matrix in *primitive form*. We *center* $\underline{D}'$ in space by finding the time averages

$$d_0(x) \equiv n^{-1} \sum_{t=1}^{n} d'(t,x) \ , \ x = 1,\ldots,p \tag{B2.1}$$

writing

$$'\underline{d}_0' \text{ for } [d_0(1), \cdots , d_0(p)]^{T} \tag{B2.2}$$

the *centroid* of $\underline{D}'$, and defining the *centroid matrix*:

$$\underline{D}_0 \equiv \begin{bmatrix} d_0(1) & d_0(2) & \cdots & d_0(p) \\ d_0(1) & d_0(2) & \cdots & d_0(p) \\ \vdots & \vdots & & \vdots \\ d_0(1) & d_0(2) & \cdots & d_0(p) \end{bmatrix} = \begin{bmatrix} \underline{d}_0^T \\ \underline{d}_0^T \\ \vdots \\ \underline{d}_0^T \end{bmatrix} \tag{B2.3}$$

Then the *space-centered** data set $\underline{D}$ is

$$\underline{D} \equiv \underline{D}' - \underline{D}_0 \tag{B2.4}$$

We write

$$'\sigma_D^2' \text{ for } \|\underline{D}' - \underline{D}_0\|^2 \tag{B2.5}$$

Thus σ_D^2 is the *norm* of $\underline{D}' - \underline{D}_0$ (cf. Appendix A). Alternate forms of σ_D^2 are

$$\sigma_D^2 = \sum_{t=1}^{n} \|\underline{d}'(t) - \underline{d}_0\|^2$$

$$= \sum_{x=1}^{p} \sum_{t=1}^{n} (d'(t,x) - d_0(x))^2 \tag{B2.6}$$

where we write:

$$'\underline{d}'(t)' \text{ for } [d'(t,1),\ldots,d'(t,p)]^T \text{ , } t = 1,\ldots,n. \tag{B2.7}$$

σ_D^2 is a measure of the scatter, or variance of $\underline{D}'$ about $\underline{D}_0$, i.e., of the n-point swarm $\{\underline{d}'(t): t = 1,\ldots,n\}$ about $\underline{d}_o$ in E_p. The *standardized form* of $\underline{D}'$ is defined by writing

---

* An entirely analogous development can be made for *time-centered* data sets, resulting in a generally dual theory of data intercomparison. The modifications of the present formalism to cover the time-centered case are readily made. Note that in DIT(II) for simplicity we dropped the primes from $d'(t,x)$, $\underline{d}'(t)$ and zeros from $\underline{d}_0(x)$, $\underline{d}_0$. Here we need the primes for the reader who wants to keep the distinction between centered and uncentered data.

$$'\tilde{\underline{D}}' \text{ for } \sigma_D^{-1}\,(\underline{D}'-\underline{D}_0) = \begin{bmatrix} (\underline{d}'(1)-\underline{d}_0)^T/\sigma_D \\ (\underline{d}'(2)-\underline{d}_0)^T/\sigma_D \\ \vdots \\ (\underline{d}'(n)-\underline{d}_0)^T/\sigma_D \end{bmatrix} \tag{B2.8}$$

3. SVD of Standardized Data Sets

Let $\tilde{\underline{D}}$ be an n×p data set in standardized form.* Then by the theory of Appendix A and (B2.8) we have

$$\|\tilde{\underline{D}}\|^2 = \text{trace } \tilde{\underline{D}}^T\tilde{\underline{D}} = 1. \tag{B3.1}$$

Let us write

$$'\tilde{\underline{S}}' \quad \text{for} \quad \tilde{\underline{D}}^T\tilde{\underline{D}} \tag{B3.2}$$

the p×p *scatter matrix* of $\tilde{\underline{D}}$. Let $\underline{e}_1,\ldots,\underline{e}_p$ be the p×1 eigenvectors of $\tilde{\underline{S}}$, and $\kappa_1^2,\ldots,\kappa_p^2$ their associated dimensionless eigenvalues, with $\kappa_j \geq 0$, $j = 1,\ldots,p$. Therefore, by definition of $\underline{e}_j$ and κ_j^2,

$$\tilde{\underline{S}}\,\underline{e}_j = \kappa_j^2\underline{e}_j \quad , \quad j = 1,\ldots,p \tag{B3.3}$$

Writing

$$'\tilde{\underline{K}}' \quad \text{for} \quad \text{diag}[\kappa_1^2,\ldots,\kappa_p^2] \tag{B3.4}$$

$$'\underline{E}' \quad \text{for} \quad [\underline{e}_1 \ \ldots \ \underline{e}_p] \qquad (p\times p) \tag{B3.5}$$

---

* As in the case of primed symbols, the tilded symbols are kept for readers who need to distinguish between dimensioned and dimensionless variables.

where

$$\underline{e}_j \equiv [e_j(1),\ldots,e_j(p)]^T \quad , \quad j = 1,\ldots,p \tag{B3.6}$$

we see that (B3.3) can take the matrix form:

$$\tilde{\underline{S}}\ \underline{E} = \underline{E}\ \tilde{\underline{K}} \tag{B3.7}$$

with

$$\underline{E}^T\underline{E} = \underline{E}\ \underline{E}^T = \underline{I}_p \tag{B3.8}$$

In view of (B3.1) and (A2.26),

$$\text{trace } \tilde{\underline{K}} = 1 \quad , \quad \text{i.e., } \sum_{j=1}^{p} \kappa_j^2 = 1 \tag{B3.9}$$

Because of the space-centering in (B2.4), $\tilde{\underline{D}}$ has at most n-1 independent rows. Hence the rank of $\tilde{\underline{D}}$, and therefore of $\tilde{\underline{S}}$ is $\min[n-1,p] = \rho$. Thus at most ρ of the κ_j^2 can be non zero. The reader should reflect on the necessity of centering a data set prior to forming the scatter matrix. (What is the interpretation of the diagonal elements, say, of the scatter matrix of a non-centered data set?) Also it should be noted that division of the centered data set by σ_D affects $\underline{K}$ but not $\underline{E}$. Finally, it should be verified that the eigenvectors associated with sets of equal eigenvalues (e.g. zero eigenvalues) are not uniquely determined.

The *principal decomposition* of $\tilde{\underline{D}}$ is defined by writing

$$'\tilde{\underline{A}}' \text{ for } \tilde{\underline{D}}\ \underline{E} \qquad (n\times p) \tag{B3.10}$$

and then writing the identity

$$\tilde{\underline{D}} = \tilde{\underline{D}}\ \underline{E}\ \underline{E}^T \tag{B3.11}$$

as

$$\tilde{\underline{D}} = \tilde{\underline{A}}\,\underline{E}^{T} \tag{B3.12}$$

Here $\tilde{\underline{A}}$ is the n×p matrix of *principal components* (which by standardization are dimensionless)

$$\tilde{\underline{A}} \equiv \begin{bmatrix} \tilde{a}_1(1) & \dots & \tilde{a}_p(1) \\ \vdots & & \vdots \\ \tilde{a}_1(n) & \dots & \tilde{a}_p(n) \end{bmatrix} \tag{B3.13}$$

On recalling (B2.8), (B3.12) can be written in vector form as

$$\underline{d}'(t) = \underline{d}_0 + \sigma_D \sum_{j=1}^{p} \tilde{a}_j(t)\underline{e}_j \quad , \quad t = 1,\dots,n \tag{B3.14}$$

or in scalar form as

$$d'(t,x) = d_0(x) + \sigma_D \sum_{j=1}^{p} \tilde{a}_j(t)e_j(x) \tag{B3.15}$$

$$t = 1,\dots,n: \quad x = 1,\dots,p.$$

The SVD of $\underline{D}$ is obtained by writing*

$$\text{'}\underline{A}\text{''} \quad \text{for} \quad \tilde{\underline{A}}\,\tilde{\underline{K}}^{-\frac{1}{2}} \qquad (n\times p) \tag{B3.16}$$

so that

$$\tilde{\underline{A}} = \underline{A}'\tilde{\underline{K}}^{\frac{1}{2}}$$

---

* If some κ_j is zero then replace "κ_j^{-1}" by "0" in (B3.16). This defines the *generalized inverse* of $\tilde{\underline{K}}^{\frac{1}{2}}$ and hence of $\tilde{\underline{D}}$.

where

$$\underline{A}' \equiv \begin{bmatrix} \alpha_1(1) & \dots & \alpha_p(1) \\ \vdots & & \vdots \\ \alpha_1(n) & \dots & \alpha_p(n) \end{bmatrix} \tag{B3.17}$$

(Observe that $\underline{A}'$ is independent of the normalization by σ_D, just as $\underline{E}$ is.)

Then (B3.12) *takes the desired SVD form*:

$$\tilde{\underline{D}} = \underline{A}'\tilde{\underline{K}}^{\frac{1}{2}}\underline{E}^T \tag{B3.18}$$

where

$$\underline{A}'^T\underline{A}' = \underline{I}_p \tag{B3.19}$$

which follows from (B3.16) and

$$\tilde{\underline{A}}^T\tilde{\underline{A}} = \tilde{\underline{K}} \ .$$

This in turn is deduced as follows (using (B3.2), (B3.7), (B3.8)):

$$\begin{aligned} \tilde{\underline{A}}^T\tilde{\underline{A}} &= (\tilde{\underline{D}}\ \underline{E})^T(\tilde{\underline{D}}\ \underline{E}) \\ &= \underline{E}^T(\tilde{\underline{D}}^T\tilde{\underline{D}})\underline{E} \\ &= \underline{E}^T\tilde{\underline{S}}\ \underline{E} \\ &= \underline{E}^T\underline{E}\ \tilde{\underline{K}} = \tilde{\underline{K}} \end{aligned} \tag{B3.20}$$

The vector and scalar versions of (B3.18) are

$$\underline{d}'(t) = \underline{d}_0 + \sigma_D \sum_{j=1}^{p} \kappa_j\alpha_j(t)\underline{e}_j \quad , \quad t = 1,\dots,n \tag{B3.21}$$

$$d'(t,x) = d_0(x) + \sigma_D \sum_{j=1}^{p} \kappa_j\alpha_j(t)e_j(x) \ , \tag{B3.22}$$

$$t = 1,\dots,n \ ; \ x = 1,\dots,p.$$

Because of space-centering (as in (B2.4)), $\sum_{t=1}^{n} \alpha_j(t) = 0$, and so the $\underline{\alpha}_j$, $j = 1,\ldots,p$ are confined to at most an n-1 dimensional subspace of E_n, i.e., to an E_ρ, $\rho = \min[n-1,p]$.

4. Connections Between Standardized Sets and Non Standardized Sets

We summarize here some of the connections between $\underline{D}$ and its standardized form $\underline{\tilde{D}}$ that may be useful in practice.

By (B2.4) and (B2.8) we have

$$\underline{D} = \sigma_D \underline{\tilde{D}} \tag{B4.1}$$

The *scatter matrix* $\underline{S}$ of $\underline{D}$ is by definition

$$\underline{S} \equiv \underline{D}^T \underline{D} \tag{B4.2}$$

and so by (B4.1), (B4.2) and (B3.2), we have

$$\underline{S} = \sigma_D^2 \underline{\tilde{S}} \tag{B4.3}$$

If ℓ_j, $j = 1,\ldots,p$ are the eigenvalues of $\underline{S}$, and κ_j^2, $j = 1,\ldots,p$, are those of $\underline{\tilde{S}}$, then

$$\ell_j = \sigma_D^2 \kappa_j^2 \ , \ j = 1,\ldots,p \tag{B4.4}$$

and conversely

$$\kappa_j = \sigma_D^{-1} \ell_j^{\frac{1}{2}} \ , \ j = 1,\ldots,p \tag{B4.5}$$

Hence

$$\underline{L} \equiv \mathrm{diag}[\ell_1,\ldots,\ell_p] = \sigma_D^2 \underline{\tilde{K}} \tag{B4.6}$$

and (B3.7) for $\underline{D}$ becomes via (B4.3), (B4.6),

$$\underline{S}\ \underline{E} = \underline{E}\ \underline{L} \tag{B4.7}$$

Thus $\underline{S}$ and $\tilde{\underline{S}}$ share the same eigenvectors.

The total scatter of the space-centered set $\underline{D}$ is given in the following equivalent forms:

$$\sum_{x=1}^{p}\sum_{t=1}^{n} d^2(t,x) = (\underline{D},\underline{D}) = \|\underline{D}\|^2 = \text{trace}\ \underline{D}^T\underline{D}$$

$$= \text{trace}\ \underline{S} = \sum_{j=1}^{p} \ell_j = \sigma_D^2 \tag{B4.8}$$

The first equality comes from (A2.1), the second from (A2.9), the third from (A2.4), the fourth from (B4.2), the fifth from (B4.7) via:

$$\underline{S} = \underline{E}\ \underline{L}\ \underline{E}^T \quad , \tag{B4.9}$$

and (A2.26). The sixth equality in (B4.8) comes from (B3.9) and (B4.6).

APPENDIX C

Canonic Rotation Angles Between Orthonormal Vector Frames

1. Introduction

The canonic rotation angles allow a reduction of the number of parameters in ORIEN (cf. §2 of the main text) by a factor of two. We now present some introductory comments on the theory of canonic rotation angles.

The SVD (B3.18) of an n×p data set $\underline{\tilde{D}}$ produces a p×p matrix $\underline{E}$ of eigenvectors:

$$\underline{E} = [\underline{e}_1 \ \ldots \ \underline{e}_p] \tag{C1.1}$$

which summarize the spatial patterns of the data set. If we had another n×p data set $\underline{\tilde{M}}$ of SVD:

$$\underline{\tilde{M}} = \underline{B}'\underline{\tilde{L}}^{\frac{1}{2}}\underline{F}^T \tag{C1.2}$$

then it may be of interest to compare $\underline{E}$ and the p×p matrix $\underline{F}$ for similarities, where

$$\underline{F} = [\underline{f}_1 \ \ldots \ \underline{f}_p] \ . \tag{C1.3}$$

A natural measure of distance between $\underline{E}$ and $\underline{F}$ is attained through the rotation transformation $\underline{R}$ that maps $\underline{E}$ into $\underline{F}$:

$$\begin{aligned} \underline{R} &= \underline{F}\,\underline{E}^T \\ &= [\underline{f}_1 \ \ldots \ \underline{f}_p] \begin{bmatrix} \underline{e}_1^T \\ \vdots \\ \underline{e}_p^T \end{bmatrix} \\ &= \sum_{j=1}^{p} \underline{f}_j \underline{e}_j^T \end{aligned} \tag{C1.4}$$

This $\underline{R}$ has the property that it maps $\underline{e}_k$ into $\underline{f}_k$:

$$\underline{R}\,\underline{e}_k = \left(\sum_{j=1}^{p} \underline{f}_j\underline{e}_j^T\right)\underline{e}_k = \sum_{j=1}^{p} \underline{f}_j(\underline{e}_j^T\underline{e}_k) = \underline{f}_k \qquad \text{(C1.5)}$$

for

$$k = 1,\ldots,p\ .$$

The norm between $\underline{e}_k$ and $\underline{f}_k$ is

$$\|\underline{e}_k - \underline{f}_k\|^2 = 2(1-\underline{e}_k^T\underline{f}_k) \qquad \text{(C1.6)}$$
$$k = 1,\ldots,p$$

and the norm between the frames $\underline{E},\underline{F}$ is

$$\begin{aligned}\|\underline{E}-\underline{F}\|^2 &= \text{trace}(\underline{E}-\underline{F})^T(\underline{E}-\underline{F}) \\ &= 2(1-\text{trace}\underline{E}^T\underline{F}) \\ &= 2\left(1-\sum_{j=1}^{p} \underline{e}_j^T\underline{f}_j\right) \qquad \text{(C1.7)}\end{aligned}$$

This, except for a normalizing factor (coming from SHAPE), is what in §3 we have called "ORIEN." Hence ORIEN is a natural measure of distance between the two frames, involving the p inner products $\underline{e}_j^T\underline{f}_j$, $j = 1,\ldots,p$. These p parameters can be used individually (as in (C1.6)) or collectively (as in (C1.7)) to measure the separations of interest. In the case of (C1.7) it is possible to reduce the numbers of parameters by half if we introduce the notion of a canonic rotation angle.

The remainder of this Appendix is devoted to the subject of canonic rotation angles between two orthonormal bases $\underline{E},\underline{F}$ of E_p. These angles are closely related to the eigenvalues of $\underline{R} = \underline{F}\,\underline{E}^T$.

2. Basic Properties of Rotations in E_p.

Matrices of the form $\underline{R} = \underline{F}\,\underline{E}^T$, where $\underline{F}$ and $\underline{E}$ are $p \times p$ matrices of orthonormal vectors, are called *rotations*. They are characterized by the property $\underline{R}^T\underline{R} = \underline{R}\,\underline{R}^T = \underline{I}_p$. We examine the structure of the $p \times 1$ eigenvectors $\underline{w}$ and associated eigenvalues λ of a rotation $\underline{R}$. Thus, by definition $\underline{R}$,$\underline{w}$, and λ are such that:

$$\underline{R}\ \underline{w} = \lambda \underline{w} \ . \tag{C2.1}$$

We know from (C1.5) that $\underline{R}$ preserves the length of $\underline{w}$ (cf. also (A2.24)). Hence (C2.1) cannot hold in general unless both λ and $\underline{w}$ are complex valued. (What must $\|\underline{w}\|$ be if $\underline{w}$ and λ in (C2.1) are real valued and $\lambda = 1$?) To prepare for such a possibility, we generalize the notion of inner product and norm to complex valued vector components:

$$\|\underline{w}\|^2 \equiv \underline{w}^*\underline{w} \equiv \bar{\underline{w}}^T\underline{w} \tag{C2.2}$$

Hence '*' means: take the complex conjugate (denoted by an overbar) as well as the transpose. In detail, if

$$\underline{w} = [w_1,\ldots,w_p]^T \ , \quad \underline{v} = [v_1,\ldots,v_p]^T \tag{C2.3}$$

then

$$\underline{w}^* = [\bar{w}_1,\ldots,\bar{w}_p] \ , \tag{C2.4}$$

and

$$\underline{w}^*\underline{v} = \sum_{j=1}^{p} \bar{w}_j v_j \ , \quad \|\underline{w}\|^2 = \sum_{j=1}^{p} |w_j|^2 \ . \tag{C2.5}$$

Suppose $\underline{w} \neq \underline{0}$ is an eigenvector of $\underline{R}$ with R's entries real valued as usual. Now, evaluate $(\underline{R}\ \underline{w})^*(\underline{R}\ \underline{w})$ two ways:

once directly via: $(\underline{R}\ \underline{w})^*(\underline{R}\ \underline{w}) = \underline{w}^*\underline{R}^*\underline{R}\ \underline{w} = \|\underline{w}\|^2$ (C2.6)

once indirectly via (C2.1): $(\lambda\underline{w})^*(\lambda\underline{w}) = |\lambda|^2\ \|\underline{w}\|^2$ (C2.7)

Since these two values are equal, we conclude

$$|\lambda|^2 = 1 \tag{C2.8}$$

Therefore:

A. *The eigenvalues* λ *of* $\underline{R}$ *are of the form* $\lambda = e^{i\theta}$, $0 \leq \theta < 2\pi$.

Observe that not only $\underline{w},\lambda$ satisfy (C2.1), but also the pair $\bar{\underline{w}},\bar{\lambda}$.

Therefore:

B. *The eigenvalues* λ *and their eigenvectors* $\underline{w}$ *come in distinct complex conjugate pairs when* $|\lambda| \neq 1$.

Suppose now that $\underline{w}_1,\lambda_1$ and $\underline{w}_2,\lambda_2$ are eigen pairs of $\underline{R}$ with $\underline{w}_1 \neq \underline{0}$, $\underline{w}_2 \neq \underline{0}$, and $\lambda_1 \neq \lambda_2$, so $\theta_1 \neq \theta_2$ in $\lambda_j = e^{i\theta_j}$, $j = 1,2$. We now evaluate $(\underline{R}\ \underline{w}_1)^*(\underline{R}\ \underline{w}_2)$ two ways:

once directly via: $(\underline{R}\ \underline{w}_1)^*(\underline{R}\ \underline{w}_2) = \underline{w}_1^*\underline{w}_2$

once indirectly via: $(\lambda_1\underline{w}_1)^*(\lambda_2\underline{w}_2) = \bar{\lambda}_1\lambda_2\ \underline{w}_1^*\underline{w}_2$

Since these two values are equal, we conclude

$$(e^{i(\theta_2-\theta_1)}-1)\ \underline{w}_1^*\underline{w}_2 = 0$$

Since $\theta_1 \neq \theta_2$, and their difference lies strictly in $(-2\pi,2\pi)$, it follows that $\underline{w}_1^*\underline{w}_2 = 0$. Hence:

C. *Eigenvectors of* $\underline{R}$ *belonging to distinct eigenvalues are orthogonal.*

In working with real data sets, the probability of encountering equal eigenvalues or real eigenvalues of $\underline{R}$ is practically zero. Therefore it is practically certain that, under real working conditions, a p×p rotation matrix with p complex, unit-magnitude eigenvalues will have a set of p

pairwise orthogonal eigenvectors. In order to simplify the present exposition we will work under this assumption unless otherwise noted.

We now examine the real and imaginary parts of the eigenvectors $\underline{w}$ of $\underline{R}$. Let

$$\underline{w} = \underline{x} + i\underline{y}$$

Here $\underline{x},\underline{y}$ are $p\times 1$ vectors with real components. Let $\underline{w}_1,\underline{w}_2$ belong respectively to eigenvalues λ_1,λ_2 such that $\lambda_1 \neq \lambda_2$ and $\lambda_1 \neq \bar{\lambda}_2$. Then for λ_1,λ_2

$$\begin{aligned} 0 = \underline{w}_1^*\underline{w}_2 &= (\underline{x}_1+i\underline{y}_1)^*(\underline{x}_2+i\underline{y}_2) \\ &= (\underline{x}_1^T-i\underline{y}_1^T)(\underline{x}_2+i\underline{y}_2) \\ &= (\underline{x}_1^T\underline{x}_2+\underline{y}_1^T\underline{y}_2) + i(\underline{x}_1^T\underline{y}_2-\underline{x}_2^T\underline{y}_1) \end{aligned} \qquad \text{(C2.9)}$$

On the other hand, for $\lambda_1,\bar{\lambda}_2$

$$0 = \underline{w}_1^*\bar{\underline{w}}_2 = (\underline{x}_1^T\underline{x}_2-\underline{y}_1^T\underline{y}_2) - i(\underline{x}_1^T\underline{y}_2-\underline{x}_2^T\underline{y}_1) \qquad \text{(C2.10)}$$

The real and imaginary parts of (C2.9), (C2.10) are separately zero. From these we conclude

$$\underline{x}_1^T\underline{x}_2 = 0 \qquad \underline{y}_1^T\underline{y}_2 = 0 \qquad \text{(C2.11)}$$

$$\underline{x}_1^T\underline{y}_2 = 0 \qquad \underline{x}_2^T\underline{y}_2 = 0 \qquad \text{(C2.12)}$$

Also, for any j, if $\underline{w}_j$ $\bar{\underline{w}}_j$ belong to $\lambda_j,\bar{\lambda}_j$ and $\lambda_j \neq \bar{\lambda}_j$ then

$$0 = \underline{w}_j^*\bar{\underline{w}}_j = (\underline{x}_j^T\underline{x}_j-\underline{y}_j^T\underline{y}_j) - 2i(\underline{x}_j^T\underline{y}_j)$$

whence

$$\underline{x}_j^T\underline{x}_j = \underline{y}_j^T\underline{y}_j \qquad \text{(C2.13)}$$

$$\underline{x}_j^T \underline{y}_j = 0 \qquad \text{(C2.14)}$$

Condition (C2.13) allows us to make all $\underline{x}_j, \underline{y}_j$ of unit length. (Since $\|\underline{w}\|^2 = \|\underline{x}\|^2 + \|\underline{y}\|^2 = 2\|\underline{x}\|^2 = 2\|\underline{y}\|^2$, in (C2.1), divide through by $\frac{1}{2}\|\underline{w}\|^2$.) The argument leading to (C2.11), (C2.12) also holds for general indexes j,k on the λ. Thus the interesting eigenvalues λ_j of a rotation are those that are non real, so that we have (C2.13), (C2.14); and those pairs λ_j, λ_k of eigenvalues that are distinct ($\lambda_j \neq \lambda_k$) and nonconjugate ($\lambda_j \neq \bar{\lambda}_k$), so that we have (C2.11), (C2.12). The real eigenvalues of $\underline{R}$ can only be 1 or -1, resulting in identity rotations or reflections, respectively.

Summarizing:

D. *If $\lambda_j, \underline{w}_j$ is an eigen pair of $\underline{R}$ with $\underline{w}_j \equiv \underline{x}_j + i\underline{y}_j$, then for any two non conjugate distinct eigenvalues λ_j, λ_k, (C2.11), (C2.12) hold, and for non real λ_j, (C2.13), (C2.14) hold. In other words, under the above conditions, $\underline{x}_j, \underline{y}_j$ are orthogonal for all j; and $\underline{x}_j, \underline{x}_k$ and $\underline{y}_j, \underline{y}_k$ are pairwise orthogonal for $j \neq k$.* Finally, *all $\underline{x}_j, \underline{y}_j$ can be normalized to unit length.*

Writing out (C2.1) in full complex form we find:

$$\underline{R}(\underline{x}_j + i\underline{y}_j) = [c_j - is_j](\underline{x}_j + i\underline{y}_j) \qquad \text{(C2.15)}$$

where* $c_j \equiv \cos\theta_j$, $s_j \equiv \sin\theta_j$.

Separating real and imaginary parts, we have

$$\underline{R}\,\underline{x}_j = c_j\underline{x}_j + s_j\underline{y}_j \qquad \text{(C2.16)}$$

$$\underline{R}\,\underline{y}_j = -s_j\underline{x}_j + c_j\underline{y}_j \; . \qquad \text{(C2.17)}$$

---

* The minus sign convention $c_j - is_j$ is chosen here and in (C2.19) below to permit rotation angles from $\underline{x}_i$ to $\underline{y}_i$ to be positive 90° clockwise, in the plane of $\underline{x}_i, \underline{y}_i$ as seen along the direction $-\underline{x}_i \times \underline{y}_i$.

These may be written as

$$\underline{R}\ [\underline{x}_j, \underline{y}_j] = [\underline{x}_j, \underline{y}_j] \begin{bmatrix} c_j & -s_j \\ s_j & c_j \end{bmatrix} \tag{C2.18}$$

which is the basis for the required real form of (C2.1). The θ_k are the real versions of the λ_k (cf. property A above), and are called *canonic rotation angles*.

We can now build up $\underline{R}$ out of pieces like (C2.18). Thus we have:

E. *Let $\underline{R}$ be a rotation in E_p. If the dimension p of E_p is even, i.e., $p = 2\ell$ for some positive integer ℓ, then there are ℓ pairs of real valued unit vectors $[\underline{x}_j, \underline{y}_j]$ in E_p, $j = 1, \ldots, \ell$, such that all 2ℓ vectors are mutually orthogonal, and ℓ angles $\theta_j, 0 \leq \theta_j < 2\pi$, such that*

$$\underline{R}[\underline{x}_1 \underline{y}_1\ \underline{x}_2 \underline{y}_2\ \cdots\ \underline{x}_\ell \underline{y}_\ell] =$$

$$[\underline{x}_1 \underline{y}_1\ \underline{x}_2, \underline{y}_2\ \cdots\ \underline{x}_\ell \underline{y}_\ell] \begin{bmatrix} c_1 & -s_1 & & & & \\ s_1 & c_1 & & & 0 & \\ & & c_2 & -s_2 & & \\ & & s_2 & c_2 & \ddots & \\ 0 & & & & c_\ell & -s_\ell \\ & & & & s_\ell & c_\ell \end{bmatrix} \tag{C2.19}$$

i.e., such that

$$\underline{R}\ \underline{W} = \underline{W}\ \underline{L} \tag{C2.20}$$

Therefore

$$\underline{R} = \underline{W}\ \underline{L}\ \underline{W}^T \tag{C2.21}$$

is the desired representation of the rotation $\underline{R}$ on E_p. Here $\underline{W}$ is a $p \times p$ matrix of eigenvectors of $\underline{R}$ defined in the context of (C2.19) such that

$$\underline{W}\,\underline{W}^T = \underline{W}^T\underline{W} = \underline{I}_p \tag{C2.22}$$

which follows from property D. We also have, by virtue of property D:

F. *Let $\underline{R}$ be a rotation in E_p. If the dimension p of E_p is odd, i.e., $p = 2\ell+1$ for some positive integer ℓ, then there are ℓ pairs of real valued unit vectors $[\underline{x}_j,\underline{y}_j]$ in E_p, $j = 1,\ldots,\ell$ such that all 2ℓ vectors are mutually orthogonal, and ℓ angles θ_j, $0 \leqq \theta_j < 2\pi$, along with a real valued unit vector $\underline{x}_{\ell+1}$ in E_p orthogonal to all 2ℓ others, such that:*

$$\underline{R}[\underline{x}_1\underline{y}_1\ \underline{x}_2\underline{y}_2\ \cdots\ \underline{x}_\ell\underline{y}_\ell\ \underline{x}_{\ell+1}] = [\underline{x}_1\underline{y}_1\ \underline{x}_2\underline{y}_2\ \cdots\ \underline{x}_\ell\underline{y}_\ell\ \underline{x}_{\ell+1}] \begin{bmatrix} c_1 & -s_1 & & & & & \\ s_1 & c_1 & & & 0 & & \\ & & c_2 & -s_2 & & & \\ & & s_2 & c_2 & \ddots & & \\ & & & & c_\ell & -s_\ell & \\ & 0 & & & s_\ell & c_\ell & \\ & & & & & & \pm 1 \end{bmatrix} \tag{C2.23}$$

i.e.,

$$\underline{R} = \underline{W}\,L\,\underline{W}^T \quad , \quad \underline{W}\,\underline{W}^T = \underline{W}^T\underline{W} = \underline{I}_p \tag{C2.24}$$

Here $\underline{x}_{\ell+1}$ is a unit eigenvector of $\underline{R}$ associated with eigenvalue ± 1 ($= e^{i\theta}$, with $\theta = 0$ or π). Note that for the case $p = 2\ell+1$, the characteristic equation of $\underline{R}$ is a polynomial of degree $2\ell+1$. The roots of this equation occur in conjugate pairs, and *all* must be of unit magnitude. The odd real root must therefore be +1 or -1.

This brings us to an important practical matter: that of finding a *unique* set of ℓ canonic rotation angles in (C2.19) or (C2.23). It is quite possible that two different computer programs, starting with data sets $\underline{D},\underline{M}$, could yield up $\underline{E},\underline{F}$ frames that give two different sets of θ_j, $j = 1,\ldots,\ell$. The following considerations will lead to *uniquely* defined sets of canonic angles θ_j, $j = 1,\ldots,[p/2]$.

3. Chirality of Orthonormal Frames

We shall show that it is possible to connect up orthonormal frames $\underline{E}$ and $\underline{F}$ by a rotation $\underline{R}$ only if the frames have the same "handedness"--i.e., left or right handedness--or "chirality", for short: a simple example will make this clear. Let $\underline{E},\underline{F}$ be of the form

$$\underline{E} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad , \quad \underline{F} = \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \tag{C3.1}$$

where $a^2+b^2 = 1$, $a > 0$, $b > 0$, and $c^2+d^2 = 1$, $c > 0$, $d > 0$. Hence $\underline{e}_1 = [a,b]^T$ and $\underline{e}_2 = [-b,a]^T$ are orthonormal vectors, as also are $\underline{f}_1 = [c,d]^T$, $\underline{f}_2 = [-d,c]^T$. Observe that the determinants of $\underline{E}$ and $\underline{F}$ are both +1. The rotation matrix that maps $\underline{E}$ into $\underline{F}$ is given by

$$\underline{R} = \underline{F}\,\underline{E}^T = \begin{bmatrix} c & -d \\ d & c \end{bmatrix}\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \tag{C3.2}$$

$$= \begin{bmatrix} ac + bd & -(ad-bc) \\ ad - bc & ac + bd \end{bmatrix} \equiv \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$$

The eigenvalues of $\underline{R}$ are given by the solutions λ of

$$(\alpha - \lambda)^2 + \beta^2 = 0 \tag{C3.3}$$

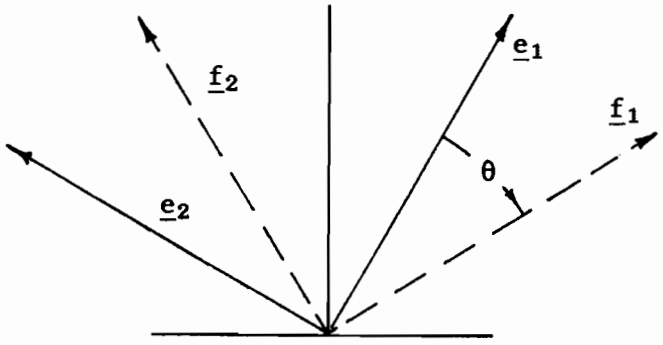
i.e.,

$$\lambda = \alpha \pm i\beta = e^{\pm i\theta} \tag{C3.4}$$

The angle θ is zero if and only if $\beta = 0$, and this is the case if and only if

$$ad = bc \quad , \quad \text{i.e.,} \quad a/b = c/d \tag{C3.5}$$

This is so if and only if $a = c$, $b = d$, i.e., the $\underline{E},\underline{F}$ frames are identical. A sketch of $\underline{e}_1,\underline{e}_2$ and $\underline{f}_1,\underline{f}_2$ is:



The meaning of θ is that of a rotation of $\underline{E}$ into $\underline{F}$.

Now suppose we multiply $\underline{e}_2$ by -1. Then we obtain a new orthonormal frame

$$\underline{E}' = \begin{bmatrix} a & b \\ b & -a \end{bmatrix} \tag{C3.6}$$

Observe that

$$\det\underline{E}' = -(a^2+b^2) = -1 \tag{C3.7}$$

whereas before

$$\det\underline{E} = a^2+b^2 = 1 \ . \tag{C3.8}$$

Moreover, the rotation matrix is now

$$\begin{aligned} \underline{R}' = \underline{F}(\underline{E}')^T &= \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \begin{bmatrix} a & b \\ b & -a \end{bmatrix} \\ &= \begin{bmatrix} ac & -bd \\ ad+bc & \end{bmatrix} \begin{bmatrix} ad+bc & \\ -(ac-bd) & \end{bmatrix} \equiv \begin{bmatrix} \gamma & \delta \\ \delta & -\gamma \end{bmatrix} \end{aligned} \tag{C3.9}$$

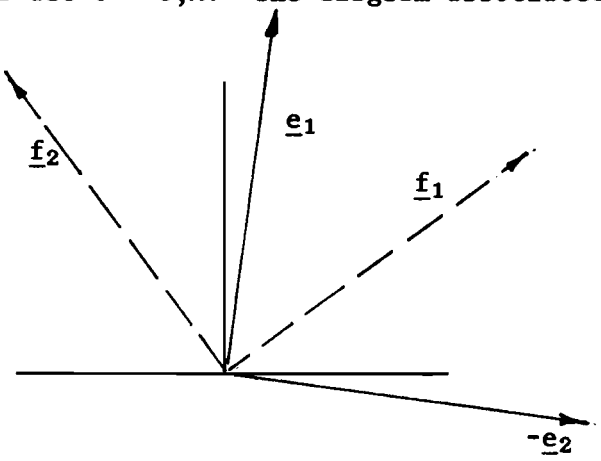
The eigenvalues of $\underline{R}'$ are given by the solutions λ of

$$-(\gamma-\lambda)(\gamma+\lambda)-\delta^2 = 0 \tag{C3.10}$$

i.e.,

$$\lambda = \pm(\gamma^2+\delta^2)^{\frac{1}{2}} = \pm 1 \qquad \text{(C3.11)}$$

The associated λ-angles are $\theta = 0,\pi$. The diagram associated with $\underline{E}',\underline{F}$ is:



We conclude that there is no rotation $\underline{R}$, with det $\underline{R} = +1$, in the plane of $\underline{e}_1,\underline{e}_2$, that maps $\underline{E}'$ into $\underline{F}$, so that $\underline{e}_1 \rightarrow \underline{f}_1$ and $-\underline{e}_2 \rightarrow \underline{f}_2$.

A numerical program encountering general p×p frames such as $\underline{E}',\underline{F}$, i.e., frames of different chirality, yields up degenerate angles, i.e., angles only in the form 0 or π. In reality, however, the associated eigenvector maps for $\underline{E}'$ and $\underline{F}$, as usually plotted in data displays, could be visually quite close. They may only differ by a sign, but that sign difference is immaterial for most uses of eigenvector plots. The way to prevent the degenerate θ case (*all* θ's are 0 or π) is to make certain, when determining the vector members of $\underline{E}$ and $\underline{F}$, that the determinants of these matrices are both of value +1 or both of value -1, i.e., that $\underline{E}$ and $\underline{F}$ have the same chirality. By simply changing the sign of one $\underline{e}_j$, or $\underline{f}_j$, the associated determinant's value undergoes a sign change, and no important change in physical properties of $\underline{e}_j$ or $\underline{f}_j$ is made. Now observe that the determinant of $\underline{R}$ is the product of the eigenvalues of $\underline{R}$

and so is easily reckoned as part of the canonic rotation angle program. Furthermore, observe, as above, that in general the determinant of an eigenframe $\underline{E}$ or $\underline{F}$ is ± 1. Hence, since

$$\det\underline{R} = (\det\underline{F})(\det\underline{E}^T) = \det(\underline{F})\det(\underline{E}) \qquad \text{(C3.12)}$$

we can insure $\det\underline{R} = +1$ by merely changing the sign of one vector, $\underline{e}_1$ (say), if necessary. Then, as the simple example above indicates, we would have $\underline{E},\underline{F}$ of the same handedness and we would obtain the associated non degenerate canonic rotation angles. Therefore a *true* rotation between $\underline{E}$ and $\underline{F}$ would be determinable.*

4. ORIEN via Canonic Rotation Angles

From §2, we have, by definition

$$\text{ORIEN}(\underline{E},\underline{F}) = \frac{1}{p}\,\|\underline{E}-\underline{F}\|^2$$

for the average norm between p×p eigenframes $\underline{E},\underline{F}$ of two n×p data sets $\underline{D},\underline{M}$, respectively. By (C1.7) we can write this as

$$\text{ORIEN}(\underline{E},\underline{F}) = \frac{2}{p}\,(1 - \sum_{j=1}^{p} \underline{e}_j{}^T\underline{f}_j) \qquad \text{(C4.1)}$$

Assuming $\underline{E},\underline{F}$ have the same chirality (i.e., they have determinants of like sign) we wish to express $\text{ORIEN}(\underline{E},\underline{F})$ in terms of the canonic rotation angles θ_j between the two frames. Towards this end, we compute

---

* The set of all p×p rotation matrices $\underline{R}$ with $\det\underline{R} = +1$ forms a group, while those that have $\det\underline{R} = -1$, do not. A *true* rotation is a matrix $\underline{R}$ with $\det\underline{R} = +1$.

$$\begin{aligned}
\|\underline{E}-\underline{F}\|^2 &= \text{trace } (\underline{E}-\underline{F})^T(\underline{E}-\underline{F}) \\
&= \text{trace } (E-\underline{R}\ \underline{E})^T(\underline{E}-\underline{R}\ \underline{E}) \\
&= \text{trace } E^T(\underline{I}_p-\underline{R}^T)(\underline{I}_p-\underline{R})\underline{E} \\
&= \text{trace } (\underline{I}_p-\underline{R}^T)(\underline{I}_p-\underline{R}) \\
&= \text{trace } [2\underline{I}_p-(\underline{R}+\underline{R}^T)]
\end{aligned}$$

Here we have used (A2.26) in going from the third to the fourth line. We now return to (C2.19), (C2.23) and use the representations of $\underline{R}$ given there. In the case of $p = 2\ell$,

$$\begin{aligned}
\text{trace}(\underline{R}+\underline{R}^T) &= \text{trace } \underline{W}(\underline{L}+\underline{L}^T)\underline{W}^T \\
&= \text{trace } (\underline{L}+\underline{L}^T) \\
&= 4 \sum_{j=1}^{\ell} c_j
\end{aligned} \tag{C4.2}$$

Also

$$\text{trace } 2\underline{I}_p = 2p = 4\ell \tag{C4.3}$$

Hence

$$\begin{aligned}
\text{ORIEN}(\underline{E},\underline{F}) = \frac{1}{p}\|\underline{E}-\underline{F}\|^2 &= \frac{4}{p}\sum_{k=1}^{\ell}(1-c_k) \quad , \\
&= \frac{4}{p}\sum_{k=1}^{\ell}(1-\cos\theta_k) \ , \ \ell = [p/2]
\end{aligned} \tag{C4.4}$$

which was to be shown. From the point of view of distance measure, we see that we can work with θ_k in the range $0 \leqq \theta_k \leqq \pi$, since $\|\underline{E}-\underline{F}\|^2$ attains a maximum when all θ_k are π.

In the case of $p = 2\ell+1$, since we want a true rotation, we take the $(2\ell+1)$st root in (C2.23) to be +1. On calculating the new counterparts to

(C4.2), (C4.3), we find (C4.4) once again. Hence (C4.4) holds for all values of p. This is reflected in the meaning of "[p/2]" in (C4.4) which says: "take the largest integer in p/2."

5. The Octant (Minimum-ORIEN) Condition

For the purpose of running an S-Phase test (§5 of the main text), one that can be duplicated by independent investigators using different eigenvalue-finding routines, we have used the following conventions. The goal of the conventions is to define two unique $\underline{E}$ and $\underline{F}$ frames belonging to given data matrices $\underline{D},\underline{M}$, which yield up the *same set of canonic rotation angles*.

A. Suppose, then, the numerical routine has returned two sets of eigenvectors $\underline{E} = [\underline{e}_1 \ \ldots \ \underline{e}_p]$, $\underline{F} = [\underline{f}_1 \ \ldots \ \underline{f}_p]$. The $\underline{e}_j$ and $\underline{f}_j$ are fixed except for signs. We use (C4.1), (C4.4) as guides to decide whether or not to change the sign of a given $\underline{e}_j$. The idea is to systematically switch the signs of the $\underline{e}_j$ so as to obtain a minimum distance between the frames $\underline{E}$ and $\underline{F}$. Therefore if some inner product in (C4.1) is negative, say $\underline{e}_j^T\underline{f}_j < 0$, then change $\underline{e}_j$ to $-\underline{e}_j$. The end result of such a systematic procedure is an ORIEN($\underline{E},\underline{F}$) of minimum value. Also, as a result, the $\underline{e}_j,\underline{f}_j$ pairs are each 90° or less apart. This is the *octant* or ***minimum ORIEN*** condition.

B. Once the octant condition has been imposed on $\underline{E},\underline{F}$, we must check to see that $\underline{E},\underline{F}$ have the same chirality.* If $(\det\underline{E})(\det\underline{F}) > 0$, then $\underline{E}$ and $\underline{F}$ have

---

* One can construct an example wherein $\underline{e}_j^T\underline{f}_j > 0$, $j = 1,\ldots,p$, but where $\underline{E},\underline{F}$ are of opposite chirality.

the same chirality and we can go on to find the canonic rotation angles. If $(\det\underline{E})(\det\underline{F}) < 0$, then pick for sign reversal that $\underline{e}_j$ which will increase ORIEN($\underline{E}$,$\underline{F}$) the least, using (C4.1) as a guide. *The end result will be a unique pair of eigenframes $\underline{E}$,$\underline{F}$ of the same chirality and of minimum possible distance apart. In this way two independent researchers using the same n×p $\underline{D}$, $\underline{M}$ sets will produce identical sets of canonic rotation angles θ_j, j = 1,...,[p/2].*

6. Representation of Rotated Vectors

A. Returning to (C2.21) and (C2.24), we convert these matrix statements into vector form. Thus (C2.21) becomes for $p = 2\ell$

$$\underline{R} = \sum_{j=1}^{\ell} [c_j(\underline{x}_j\underline{x}_j^T + \underline{y}_j\underline{y}_j^T) - s_j(\underline{x}_j\underline{y}_j^T - \underline{y}_j\underline{x}_j^T)] \qquad \text{(C6.1)}$$

where

$$c_j \equiv \cos\theta_j \quad , \quad s_j \equiv \sin\theta_j \quad , \quad j = 1,\ldots,\ell$$

The action of $\underline{R}$ on the eigenvectors $\underline{x}_j$,$\underline{y}_j$ is, for j = 1,...,ℓ

$$\underline{R}\,\underline{x}_j = c_j\underline{x}_j + s_j\underline{y}_j \qquad \text{(C6.2)}$$

$$\underline{R}\,\underline{y}_j = -s_j\underline{x}_j + c_j\underline{y}_j \qquad \text{(C6.3)}$$

so that

$$\underline{R}\,(\underline{x}_j + i\underline{y}_j) = e^{-i\theta}j(\underline{x}_j + i\underline{y}_j) \qquad \text{(C6.4)}$$

For the case of $p = 2\ell+1$, in (C2.24), set $c_{\ell+1} = 1$, $s_{\ell+1} = 0$, $\underline{y}_{\ell+1} = 0$ and replace ℓ by $\ell+1$ everywhere in (C6.1). Thus (C2.24) becomes

$$\underline{R} = \sum_{j=1}^{\ell} [c_j(\underline{x}_j\underline{x}_j^T + \underline{y}_j\underline{y}_j^T) - s_j(\underline{x}_j\underline{y}_j^T - \underline{y}_j\underline{x}_j^T)] \pm \underline{x}_{\ell+1}\underline{x}_{\ell+1}^T \tag{C6.5}$$

The plus sign is chosen for true rotations.

Now let $\underline{z}$ be an arbitrary vector. Then the image of $\underline{z}$ under $\underline{R}$ for $p = 2\ell+1$, is

$$\underline{w} = \underline{R}\,\underline{z} = \sum_{j=1}^{\ell} [c_j(\underline{x}_j\xi_j + \underline{y}_j\eta_j) - s_j(\underline{x}_j\eta_j - \underline{y}_j\xi_j)] \pm \underline{x}_{\ell+1}\xi_{\ell+1}$$

$$= \sum_{j=1}^{\ell} [(c_j\xi_j - s_j\eta_j)\underline{x}_j + (c_j\eta_j + s_j\xi_j)\underline{y}_j] \pm \underline{x}_{\ell+1}\xi_{\ell+1} \tag{C6.6}$$

where

$$\xi_j = \underline{z}^T\underline{x}_j \quad , \quad \eta_j = \underline{z}^T\underline{y}_j \quad , \quad j = 1,\ldots,\ell \tag{C6.7}$$

For the case of $p = 2\ell$, drop the last term in (C6.6).

B. A special case of (C6.1) or (C6.5) occurs when the vectors $\underline{x}_j, \underline{y}_j$ are from a natural basis of some E_p in which we are working. Let $\{\underline{u}_1,\ldots,\underline{u}_p\}$ be a natural basis of E_p (i.e., the kth element of $\underline{u}_j$ is the Kronecker delta δ_{kj}). Then if $p = 2\ell$, (C6.1) becomes

$$\underline{R} = \sum_{j=1}^{\ell} [c_j(\underline{u}_{2j-1}\underline{u}_{2j-1}^T + \underline{u}_{2j}\underline{u}_{2j}^T) - s_j(\underline{u}_{2j-1}\underline{u}_{2j}^T - \underline{u}_{2j}\underline{u}_{2j-1}^T)] \tag{C6.8}$$

If $p = 2\ell+1$, then add $\underline{u}_{2\ell+1}\underline{u}_{2\ell+1}^T$ to the right side of (C6.8).

From this we see that the action of $\underline{R}$ on a vector

$$\underline{z} = \sum_{j=1}^{p} z_j \underline{u}_j \tag{C6.9}$$

is one of rotations in successive planes formed by adjacent pairs $\underline{u}_{2j-1}, \underline{u}_{2j}$ of vectors, $j = 1,\ldots,\ell$.

C. The rotations in the natural frame $\{\underline{u}_1,\ldots,\underline{u}_p\}$ need not be confined to adjacent members of the basis. A rotation in E_p within the plane of the arbitrary vector pair $\underline{u}_u, \underline{u}_v$ is given by the $p\times p$ matrix.

$$\underline{R}_{uv} \equiv \begin{array}{c} \\ u \\ \\ \\ v \\ \\ \\ \end{array} \left[\begin{array}{ccccccc} 1 \ddots 1 & & & & & & \\ & c_j & & & -s_j & & \\ & & 1 & & & & \\ & & & \ddots 1 & & & \\ & s_j & & & c_j & & \\ & & & & & 1 & \\ & & & & & & \ddots 1 \end{array}\right] \tag{C6.10}$$
$$\qquad\qquad\qquad u \qquad\qquad\qquad v$$

Thus the c_j's occur on the main diagonal at locations u and v, and the s_j's occur, as shown, to fill out the rectangle of four symbols. All other entries of the main diagonal are 1 and all remaining off-diagonal entries are zero, indicating zero rotation within all other possible coordinate planes. There are $\frac{1}{2}p(p-1)$ linearly independent matrices of this structure, and these form the basis of a vector space of rotations within the frame $\{\underline{u}_1,\ldots,\underline{u}_p\}$. If the determinants of these matrices are 1 (choose the + sign of the last diagonal entry in (C2.23) when $p = 2\ell+1$) then the vector space is a group under matrix multiplication (and hence we will have a $\frac{1}{2}p(p-1)$ dimensional field of rotations).

7. Canonic Direction Angles of a Unit Vector in E_p

In this section we develop the theory of canonic direction angles of a unit vector in E_p. These angles are direct generalizations of the polar coordinate angles of a unit vector in E_3. They may be of use in tests of significance of differences between pairs of unit vectors. Such tests arise, e.g., in principal vector (EOF) intercomparisons between two data sets. Suggestions for new tests, based on canonic direction angles, for vectors and for orthonormal frames, are made in §7 of the main text.

A. Let $\underline{e} = [e_1,\ldots,e_p]^T$ be a unit vector in E_p. Let $\{\underline{u}_1,\ldots,\underline{u}_p\}$ be a natural basis of E_p (i.e., $\underline{u}_j$ has Kronecker deltas as components δ_{kj}). For the purpose of finding the canonic direction angles of $\underline{e}$, we will subject $\underline{e}$ to a sequence of rotations $\underline{R}_k(\phi_k)$ of the form (C6.8):

$$\underline{R}_k(\phi_k) = \underline{R}(\underline{u}_k,\underline{u}_{k+1};\phi_k) + \underline{I}(\underline{u}_{k+2},\ldots,\underline{u}_p) \qquad (C7.1)$$
$$k = 1,\ldots,p-1.$$

Here

$$\underline{R}(\underline{u}_k,\underline{u}_{k+1};\phi_k) \equiv c_k[\underline{u}_k\underline{u}_k^T + \underline{u}_{k+1}\underline{u}_{k+1}^T] - s_k[\underline{u}_k\underline{u}_{k+1}^T - \underline{u}_{k+1}\underline{u}_k^T] \qquad (C7.2)$$

$$\underline{I}(\underline{u}_{k+2},\ldots,\underline{u}_p) \equiv \underline{u}_{k+2}\underline{u}_{k+2}^T + \underline{u}_{k+3}\underline{u}_{k+3}^T + \cdots + \underline{u}_{p-1}\underline{u}_{p-1}^T + \underline{u}_p\underline{u}_p^T \qquad (C7.3)$$

$$c_k \equiv \cos\theta_k \ , \ s_k \equiv \sin\phi_k \qquad (C7.4)$$

To keep track of successive mappings of $\underline{e}$, let us place a subscript "1" on it to start the sequence, so: $\underline{e}_1 \equiv \underline{e}$. Then in (C7.1) set k = 1 and find

$$\underline{e}_2 \equiv \underline{R}_1(\phi_1)\underline{e}_1 \qquad (C7.5)$$

where

$$\underline{e}_1 \equiv \sum_{j=1}^{p} e_j \underline{u}_j \qquad \text{(C7.6)}$$

The vector $\underline{e}_2$ may be written as:

$$\underline{e}_2 = (c_1[\underline{u}_1\underline{u}_1^T + \underline{u}_2\underline{u}_2^T] - s_1[\underline{u}_1\underline{u}_2^T - \underline{u}_2\underline{u}_1^T])\underline{e}_1 + \underline{I}(\underline{u}_3,\ldots,\underline{u}_p)\underline{e}_1$$

$$= (c_1e_1 - s_1e_2)\underline{u}_1 + (c_1e_2 + s_1e_1)\underline{u}_2 + \sum_{j=3}^{p} e_j\underline{u}_j \qquad \text{(C7.7)}$$

We want to find that value of ϕ_1 in (C7.7) which will make the coefficient of $\underline{u}_1$ zero. Thus we require

$$c_1e_1 - s_1e_2 = 0 \qquad \text{(C7.8)}$$

It follows that

$$c_1 = e_2/\gamma_2 \quad , \quad s_1 = e_1/\gamma_2 \qquad \text{(C7.9)}$$

where

$$\gamma_2 = (e_1^2 + e_2^2)^{\frac{1}{2}} \geq 0 \qquad \text{(C7.10)}$$

Using the principal branch $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$ of the arctan(x) function where x is in the range $(-\infty,\infty)$, we have ϕ_1 defined in the range, $-\frac{1}{2}\pi \leq \phi_1 < \frac{3}{2}\pi$:

$$\phi_1 = \begin{cases} \frac{1}{2}\pi - \arctan(e_2/e_1) & \text{when } \begin{cases} e_1 > 0, & e_2 \geq 0 \\ \text{or} & \\ e_1 > 0, & e_2 < 0 \end{cases} \\ 0 & \text{when } e_1 = 0, \quad e_2 > 0 \\ 0 & \text{when } e_1 = 0, \quad e_2 = 0 \\ \pi & \text{when } e_1 = 0, \quad e_2 < 0 \\ -\frac{1}{2}\pi - \arctan(e_2/e_1) & \text{when } \begin{cases} e_1 < 0, & e_2 \leq 0 \\ \text{or} & \\ e_1 < 0, & e_2 > 0 \end{cases} \end{cases} \qquad \text{(C7.11)}$$

Using this value of ϕ_1 in (C7.7), we obtain

$$\underline{e}_2 = \gamma_2 \underline{u}_2 + \sum_{j=3}^{p} e_j \underline{u}_j \qquad \text{(C7.12)}$$

Setting k = 2 in (C7.1), then we have

$$\underline{e}_3 = \underline{R}_2(\phi_2)\underline{e}_2 \qquad \text{(C7.13)}$$

Using the representation of $\underline{R}_2(\phi_2)$ we find

$$\underline{e}_3 = (c_2\gamma_2 - s_2 e_3)\underline{u}_2 + (c_2 e_3 + s_2\gamma_2)\underline{u}_3 + \sum_{j=4}^{p} e_j \underline{u}_j \qquad \text{(C7.14)}$$

We require the coefficient of $\underline{u}_2$ in (C7.14) to vanish. Thus we require

$$c_2\gamma_2 - s_2 e_3 = 0 \qquad \text{(C7.15)}$$

It follows that

$$c_2 = e_3/\gamma_3 \quad , \quad s_2 = \gamma_2/\gamma_3 \qquad \text{(C7.16)}$$

where

$$\gamma_3 = (e_1^2 + e_2^2 + e_3^2)^{\frac{1}{2}} \geq 0 \ . \qquad \text{(C7.17)}$$

From (C7.16) we require

$$\phi_2 = \begin{cases} \arccos(e_3/\gamma_3) & \gamma_3 > 0 \\ 0 & \gamma_3 = 0 \end{cases} \qquad \text{(C7.18)}$$

so that ϕ_2 is in the range $0 \leqq \phi_2 \leqq \pi$.

The choices of the zero ϕ_1, ϕ_2 values, when $\gamma_2 = 0$, $\gamma_3 = 0$, respectively, in (C7.11) and (C7.18), are made on the common-sense basis that no rotation is needed when no component of finite size of the current $\underline{e}_j$ falls on the current $\underline{u}_j$ basis vector.

If we go through the stage once more, now finding ϕ_3, $\underline{R}_3(\phi_3)$ and $\underline{e}_4$, we see that the case just completed for k = 2 is repeated in all essential steps. Hence the kth stage yields ϕ_k, $\underline{R}_k(\phi_k)$ and $\underline{e}_{k+1}$ of the form:

$$\underline{e}_{k+1} = \underline{R}_k(\phi_k)\underline{e}_k \quad , \quad k = 1,\ldots,p-1 \quad , \quad \text{with } \underline{e}_1 \equiv \underline{e} \text{ given} \qquad \text{(C7.19)}$$

$$\underline{R}_k(\phi_k) = \underline{R}(\underline{u}_k, \underline{u}_{k+1}; \phi_k) + \underline{I}(\underline{u}_{k+2}, \ldots, \underline{u}_p), \text{ as in (C7.2), (C7.3), (C7.4)} \qquad \text{(C7.20)}$$

$$\phi_k = \begin{cases} \arccos[e_{k+1}/\gamma_{k+1}], \ \gamma_{k+1} = [e_1^2 + \cdots + e_{k+1}^2]^{\frac{1}{2}} \ , \ k = 2,\ldots,p-1, \\ 0 \ , \ \gamma_{k+1} = 0 \end{cases} \qquad \text{(C7.21)}$$

$$\phi_1 \text{ as in (C7.11)} \qquad \text{(C7.22)}$$

$$\underline{e}_k = \gamma_k \underline{u}_k + \sum_{j=k+1}^{p} e_j \underline{u}_j \quad , \quad k = 1,\ldots,p-1 \qquad \text{(C7.23)}$$

$$\underline{e}_p = \underline{u}_p \qquad \text{(C7.24)}$$

The angles ϕ_k, k = 1,...,p-1, are the desired *canonic direction angles* associated with the unit vector $\underline{e}$ in E_p.

B. The process of finding the ϕ_k is shown schematically in Fig. C7.1. The amounts of rotation required to reduce the $\underline{u}_k$-component of $\underline{e}_k$ and its two successors to zero is shown in parts (a), (b), (c) of the figure. The final three stages of rotation are shown in parts (d), (e), (f). The vector $\underline{e}$ has therefore been subject to p-1 successive rotations (some possibly of zero degrees) which eventually place $\underline{e}$'s image along $\underline{u}_p$. The p-1 numbers ϕ_k, $k = 1,\ldots,p-1$ associated to $\underline{e}$ in this way constitute a form of address of ϕ_k with respect to the natural basis $\{\underline{u}_1,\ldots,\underline{u}_p\}$. Therefore, like the components e_j of $\underline{e}$, the ϕ_k may change from one basis to another. However, if we have two vectors $\underline{e},\underline{f}$ that are close in the sense of the norm $\|\underline{e}-\underline{f}\|^2$, then the two sets of angles $\phi_1,\ldots,\phi_{p-1}$ and $\phi_1',\ldots,\phi_{p-1}'$ associated with them, with respect to a fixed natural basis, will generally have the p-1 values $\phi_k-\phi_k'$ small and be directly comparable. The matter is not quite this simple, but perhaps some form of systematic comparison of ϕ_k and ϕ_k', $k = 1,\ldots,p-1$ for a fixed natural basis will lead to a method of judging whether two unit vectors are significantly close or distant, resulting in a method that is more powerful (in the technical sense) than one based on simple one-number indexes such as correlations or norms.

C. If we apply the formulas (C7.21), (C7.22) to the individual natural basis elements we may find their canonic direction angles. For example let $\underline{u}_1 = (1,\ldots,0)$ be substituted in place of the general vector $\underline{e}$ used to develop the theory. In terms of the $\underline{e}$-components we have $e_1 = 1$, $e_2 = 0,\ldots,e_p = 0$. From (C7.11) and these values of e_j, we find $\phi_1 = \frac{1}{2}\pi$. From (C7.21), $\phi_2 = \cdots = \phi_{p-1} = \frac{1}{2}\pi$. Going next to $\underline{u}_2 = (0,1,\ldots,0)$, we have $e_1 = 0$, $e_2 = 1,\ldots,e_p = 0$, and find $\phi_1 = 0$, $\phi_2 = \cdots = \phi_{p-1} = \frac{1}{2}\pi$. Table C7.1 summarizes the canonic direction angles, so found, for all the basis elements.

CANONIC DIRECTION ANGLES FOR A UNIT VECTOR IN E_p. STAGES k, k+1, k+2, ..., p-2, p-1, p

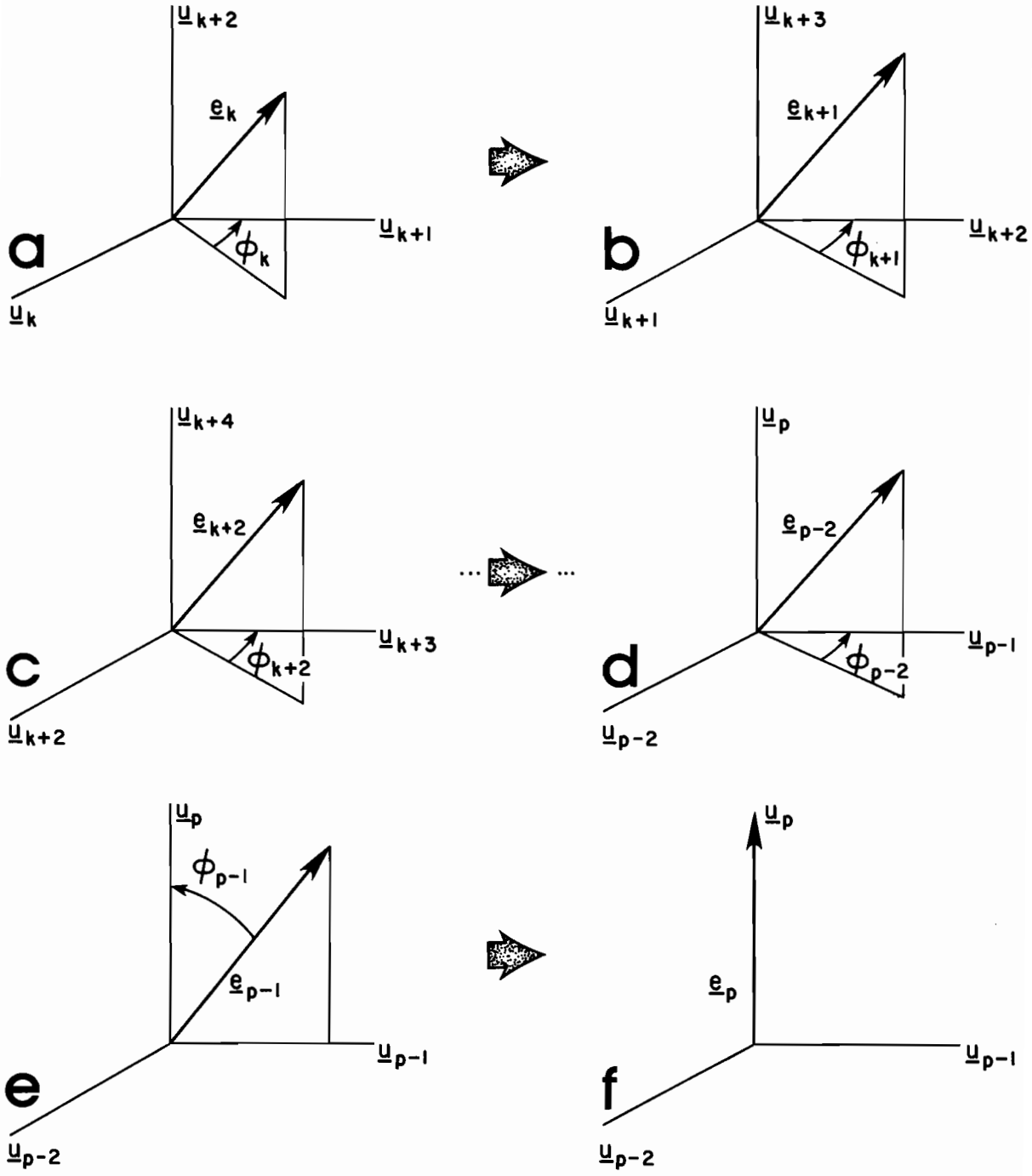


Fig. C7.1

Table C7.1

Canonic Direction Angles of Natural Basis Vectors $\underline{u}_j$

Basis Vector	Angle Index: 1	2	3	···	k···		p-2	p-1
$\underline{u}_1$	½π	½π	½π		···		½π	½π
$\underline{u}_2$	0	½π	½π		···		½π	½π
⋮	⋮	⋮	⋮				⋮	⋮
$\underline{u}_k$	0	0	0	···	½π	···	½π	½π
⋮	⋮	⋮	⋮		⋮		⋮	⋮
$\underline{u}_{p-1}$	0	0	0	···	0	···	0	½π
$\underline{u}_p$	0	0	0	···	0	···	0	0

A study of this table, with an eye toward learning to gauge "distance" between two unit vectors in terms of canonic directions, reveals the following possibilities. We know that $\|\underline{u}_j - \underline{u}_k\|^2 = 2(1-\underline{u}_j^T\underline{u}_k) = 2$ for all $j,k = 1,\ldots,p$. Therefore each $\underline{u}_j$ is the same distance from another $\underline{u}_k$, namely $2^{\frac{1}{2}}$. Moreover, every distinct pair $\underline{u}_j,\underline{u}_k$ subtends the same angle, namely ½π. In particular $\underline{u}_1$ and $\underline{u}_2$ subtend the same angle, namely ½π, as subtended by $\underline{u}_1,\underline{u}_p$. On comparing the canonic direction angles of $\underline{u}_1,\underline{u}_2$, we see they differ in the first angle by an amount ½π. They do not differ in the other angles following the first. On comparing $\underline{u}_1,\underline{u}_p$, we see they differ in all angles. But yet the equations $\|\underline{u}_1-\underline{u}_p\|^2 = 2 = \|\underline{u}_1-\underline{u}_2\|^2$ say that the distances between $\underline{u}_1,\underline{u}_p$ and $\underline{u},\underline{u}_2$ in terms of canonic direction should be the same. If we are to use the canonic direction angles as measures of distance, then for a given pair of unit vectors we may proceed along the set of angle indexes of their canonic direction angles *taking differences of the angles and tallying the amounts and*

number of new differences encountered. In this way, every pair of basis elements differ by the fixed amount $\frac{1}{2}\pi$. More generally, if $\phi_1,\ldots,\phi_{p-1}$ and $\phi'_1,\ldots,\phi'_{p-1}$ are the canonic direction angles of unit vectors $\underline{e}$ and $\underline{f}$, then we would form the p-1 differences: $\phi_1-\phi'_1$, $\phi_2-\phi'_2$, $\phi_3-\phi'_3,\ldots,$ $\phi_{p-1}-\phi'_{p-1}$. If $\phi_1-\phi'_1 \neq \phi_2-\phi'_2$, then we retain $\phi_1-\phi_2$ for later use. If $\phi_1-\phi'_1 = \phi_2-\phi'_2$ then we discard $\phi_1-\phi'_1$ and retain $\phi_2-\phi'_2$. If $\phi_2-\phi'_2 \neq \phi_3-\phi'_3$, then we retain $\phi_2-\phi'_2$ for later use. However, if $\phi_2-\phi'_2 = \phi_3-\phi'_3$, we discard $\phi_2-\phi'_2$, and go on to compare $\phi_3-\phi'_3$ with $\phi_4-\phi'_4$, and so on down the line. The number of these retained differences and their amounts may be r-tile classified, as suggested in §7 of the main text, and used as multiparameter gauges of separation of the unit vectors $\underline{e}$ and $\underline{f}$.

D. We now establish the reverse connection between the canonic direction angles ϕ_k and the components $e_1,\ldots,e_p$ of the given vector $\underline{e}$. The direct connection is given in (C7.21),(C7.22), going from the e_j to the ϕ_k. We will show how the e_j are obtained from the ϕ_k.

From (C7.19) we have

$$\underline{e}_k = \underline{R}^T(\phi_k)\underline{e}_{k+1} \quad , \quad k = 1,\ldots,p-1 \tag{C7.25}$$

where

$$\underline{R}^T(\phi_k) = \underline{R}(-\phi_k) \tag{C7.26}$$

Hence in (C7.2)

$$\underline{R}^T(\phi_k) = c_k[\underline{u}_k\underline{u}_k^T + \underline{u}_{k+1}\underline{u}_{k+1}^T] + s_k[\underline{u}_k\underline{u}_{k+1}^T - \underline{u}_{k+1}\underline{u}_k^T] + \underline{I}(\underline{u}_{k+2},\ldots,\underline{u}_p) \tag{C7.27}$$

Starting with k = p-1, since $\underline{e}_p = \underline{u}_p$ (cf. (C7.24)), we find

$$\underline{e}_{p-1} = \underline{R}_{p-1}^{T}(\phi_{p-1})\underline{e}_{p}$$

$$= c_{p-1}[\underline{u}_{p-1}\underline{u}_{p-1}^{T} + \underline{u}_{p}\underline{u}_{p}^{T}]\underline{u}_{p} + s_{p-1}\,[\underline{u}_{p-1}\underline{u}_{p}^{T} - \underline{u}_{p}\underline{u}_{p-1}^{T}]\underline{u}_{p} \qquad \text{(C7.28)}$$

whence

$$\underline{e}_{p-1} = c_{p-1}\underline{u}_{p} + s_{p-1}\underline{u}_{p-1} \qquad \text{(C7.29)}$$

Next,

$$\underline{e}_{p-2} = \underline{R}_{p-2}^{T}(\phi_{p-2})\underline{e}_{p-1}$$

$$= c_{p-2}[\underline{u}_{p-2}\underline{u}_{p-2}^{T} + \underline{u}_{p-1}\underline{u}_{p-1}^{T}]\,\underline{e}_{p-1} + s_{p-2}[\underline{u}_{p-2}\underline{u}_{p-1}^{T} - \underline{u}_{p-1}\underline{u}_{p-2}^{T}]\underline{e}_{p-1}$$

$$= c_{p-2}s_{p-1}\underline{u}_{p-1} + s_{p-1}s_{p-2}\underline{u}_{p-2} \qquad \text{(C7.30)}$$

From this pattern we generally expect, for $j = 1,\ldots,p-1$,

$$\underline{e}_{p-j} = c_{p-j}s_{p-1}s_{p-2}\cdots s_{p-(j-1)}\underline{u}_{p-(j-1)} + s_{p-1}s_{p-2}\cdots s_{p-j}\underline{u}_{p-j} \qquad \text{(C7.31)}$$

In particular, on setting $j = p-3$, $j = p-2$ and $j = p-1$, we find

$$\underline{e}_{3} = c_{3}s_{p-1}s_{p-2}\cdots s_{4}\underline{u}_{4} + s_{p-1}s_{p-2}\cdots s_{3}\underline{u}_{3} \qquad \text{(C7.33)}$$

$$\underline{e}_{2} = c_{2}s_{p-1}s_{p-2}\cdots s_{3}\underline{u}_{3} + s_{p-1}s_{p-2}\cdots s_{2}\underline{u}_{2} \qquad \text{(C7.34)}$$

$$\underline{e}_{1} = c_{1}s_{p-1}s_{p-2}\cdots s_{2}\underline{u}_{2} + s_{p-1}s_{p-2}\cdots s_{1}\underline{u}_{1} \qquad \text{(C7.35)}$$

We may tabulate these results:

Table C7.2

Components of a Unit Vector $\underline{e}$ in E_p in terms of its Canonic Direction Angles

Basis Vector	Component of $\underline{e}$ along $\underline{u}_j$
$\underline{u}_p$	$e_p = \cos\phi_{p-1}$
$\underline{u}_{p-1}$	$e_{p-1} = \sin\phi_{p-1} \cos\phi_{p-2}$
$\underline{u}_{p-2}$	$e_{p-2} = \sin\phi_{p-1} \sin\phi_{p-2} \cos\phi_{p-3}$
$\vdots$	$\vdots$
$\underline{u}_{p-j}$	$e_{p-j} = \sin\phi_{p-1} \sin\phi_{p-2} \cdots \sin\phi_{p-j} \cos\phi_{p-(j-1)}$
$\vdots$	$\vdots$
$\underline{u}_3$	$e_3 = \sin\phi_{p-1} \sin\phi_{p-2} \cdots \sin\phi_3 \cos\phi_2$
$\underline{u}_2$	$e_2 = \sin\phi_{p-1} \sin\phi_{p-2} \cdots \sin\phi_3 \sin\phi_2 \cos\phi_1$
$\underline{u}_1$	$e_1 = \sin\phi_{p-1} \sin\phi_{p-2} \cdots \sin\phi_3 \sin\phi_2 \sin\phi_1$

The canonic direction angles of a unit vector $\underline{e}$ in E_p are therefore the generalizations to E_p of the polar coordinates of a unit vector in E_3. Here the range of ϕ_1 is $-\frac{1}{2}\pi \leq \phi_1 < \frac{3}{2}\pi$, and $0 \leq \phi_k \leq \pi$, $k = 2,\ldots,p-1$.

E. It is easy to check, from Table C7.2, that $e_1^2+\cdots+e_p^2 = 1$. This indicates that the p components of the unit vector $\underline{e}$ may be arbitrarily chosen subject only to the unit length constraint. A practical way to do this is to select p numbers ε_j from $(-\infty,\infty)$ according to some rule, and then set $e_j = \varepsilon_j/(\varepsilon_1^2+\cdots+\varepsilon_p^2)^{\frac{1}{2}}$, for $j = 1,\ldots,p$. One such rule, the *normal-distribution rule* (pick the ε_j randomly from N(0,1)) is discussed in Appendix E, along with the distribution of e_j it induces. On the other hand, if we choose the canonic directions ϕ_j

arbitrarily from their respective domains, then each choice will result in a unit vector $\underline{e}$, using Table C7.2 to find the e_j from the ϕ_j. It is plausible that if the components e_j are generated using the normal-distribution rule, then each canonic direction angle ϕ_j is *uniformly* distributed over its range and independently of the other ϕ_j. Conversely, by constructing $\underline{e}$ via independent random selections of ϕ_j from a *uniformly* distributed population of angles over their ranges, the components of $\underline{e}$ will be distributed as if they were generated from the normal-distribution rule. (It would then appear that the normal-distribution rule is a sufficient but not necessary rule to generate uniformly distributed canonic rotation angles.) We will not seek proofs here of all these intuitively plausible assertions (see §7C).

8. Canonic Direction Angles of an Orthonormal Frame in E_p

We come now to the final main topic of this Appendix, namely the assignment of canonic direction angles to each member of an orthonormal frame in E_p, where E_p has some chosen fixed natural basis to be used to find the angles. The purpose of these angles is to provide another gauge of distance between two p×p orthonormal frames $\underline{E},\underline{F}$, within the given E_p, a gauge that does not compress the $\frac{1}{2}p(p-1)$ pieces of information inherent in each frame into a single number, as does the ORIEN statistic. However, the utility of these angles will not be explored much beyond their derivation, which is included here for future reference.

A. Let $\underline{E} = [\underline{e}_1\underline{e}_2 \cdots \underline{e}_p]$ be an orthonormal frame in E_p relative to a given natural basis. For the present purpose, let us relabel these as $\underline{e}_{11},\underline{e}_{12},\ldots,\underline{e}_{1p}$ to designate the first step in a sequence of p-1 steps. Starting with $\underline{e}_{11}$ and using (C7.21), (C7.22), find $\underline{e}_{11}$'s canonic direction angles: $\phi_{11},\phi_{12},\ldots,\phi_{1,p-1}$.

Also, by means of (C7.20) and using these angles, find $\underline{R}_{1k}(\phi_{1k})$, $k = 1,\ldots,p-1$. Then form the matrix product $\underline{S}_1 \equiv R_{1,p-1}(\phi_{p-1}) \cdots \underline{R}_{11}(\phi_1)$. This matrix, by our constructions in par A, maps $\underline{e}_{11}$ into $\underline{u}_p$. Moreover, $\underline{S}_1$ maps the remaining p-1 elements $e_{12},\ldots,\underline{e}_{1p}$ of the frame $\underline{E}$ into new p-1 vectors: $\underline{e}_{2,k-1} \equiv \underline{S}_1 \underline{e}_{1,k}$, $k = 2,\ldots,p$ which are all in the p-1 dimensional subspace spanned by $\underline{u}_1,\ldots,\underline{u}_{p-1}$.

B. We now may start over with the p-1 vectors $\underline{e}_{2,k}$, $k = 1,\ldots,p-1$ in the p-1 space of the $\underline{u}_j$'s, $j = 1,\ldots,p-1$. Thus, using (C7.21), (C7.22), find $\underline{e}_{21}$'s canonic direction angles: $\phi_{21},\phi_{22},\ldots,\phi_{2,p-2}$. Also, by means of (C7.20), and using these angles, find $\underline{R}_{2k}(\phi_{2k})$, $k = 1,\ldots,p-2$. Then form the matrix product $\underline{S}_2 \equiv R_{2,p-2}(\phi_{2,p-2}) \cdots \underline{R}_{21}(\phi_{21})$. This matrix maps $\underline{e}_{21}$ into $\underline{u}_{p-1}$ and it maps the remaining p-2 elements $\underline{e}_{22},\ldots,\underline{e}_{2,p-1}$ into new elements: $\underline{e}_{3,k-1} = \underline{S}_2\underline{e}_{2,k}$, $k = 2,\ldots,p-1$ in a subspace spanned by $\underline{u}_1,\ldots,\underline{u}_{p-2}$.

C. The general jth step, $j = 1,\ldots,p-1$, starts with $\underline{e}_{j1},\underline{e}_{j2},\ldots,\underline{e}_{j,p-(j-1)}$. From $\underline{e}_{j1}$ we find $\phi_{j1},\phi_{j2},\ldots,\phi_{j,p-j}$. Using these p-j angles, we construct $\underline{R}_{jk}(\phi_{jk})$, $k = 1,\ldots,p-j$. Then form the product $\underline{S}_j = \underline{R}_{j,p-j}(\phi_{j,p-j}) \cdots \underline{R}_{j1}(\phi_{j1})$. $\underline{S}_j$ by construction maps $\underline{e}_{j1}$ into $\underline{u}_{p-(j-1)}$, and we define a new set of p-j transformed basis elements via $\underline{e}_{j+1,k-1} \equiv \underline{S}_j\underline{e}_{jk}$, $k = 2,\ldots,p-(j-1)$ contained in the subspace spanned by $\underline{u}_1,\ldots,\underline{u}_{p-j}$.

D. The results of these p-1 steps are summarized in Table C8.1.

Table C8.1

Canonic Direction Angles of a p×p Orthonormal Frame $[\underline{e}_1 \ \ldots \ \underline{e}_p]$

$$\Phi: \quad \begin{matrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1,p-3} & \phi_{1,p-2} & \phi_{1,p-1} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2,p-3} & \phi_{2,p-2} & \\ \vdots & \vdots & & \vdots & & \\ \phi_{j1} & \phi_{j2} & \cdots & \phi_{j,p-j} & & \\ \vdots & \vdots & & & & \\ \phi_{p-2,1} & \phi_{p-2,2} & & & & \\ \phi_{p-1,1} & & & & & \end{matrix} \qquad \text{(C8.1)}$$

These angles, as in the individual-vector case of §7 of this Appendix, are dependent on the chosen, fixed natural basis $\{\underline{u}_1,\ldots,\underline{u}_p\}$ of E_p. However, in a given basis, two orthonormal frames $\underline{E},\underline{F}$ will have a uniquely assigned ordered set Φ of $\frac{1}{2}p(p-1)$ canonic direction angles that may be intercompared row by corresponding row and index by corresponding index, much as in the vector case of §7 of this Appendix. In this way we can develop a multiparameter gauge of the closeness of two frames $\underline{E},\underline{F}$ of a different character than ORIEN. Under random sampling from N(0,1) to generate an n×p data set $\underline{D}$, we would expect the ϕ_{jk} in (C8.1) to be uniformly distributed over their respective ranges. If this is so, then the r-tile method of intercomparing vectors in E_p (§7 of the main text) would be extendable to intercomparisons of $\underline{E}$ and $\underline{F}$.

APPENDIX D

Canonic Correlation Angles Between Orthonormal Vector Frames

1. Introduction

The canonic correlation angles do for principal components what the canonic rotation angles do for principal vectors: they provide an economical description of the separation of the sets of principal component time series. We now present some introductory comments on the theory of canonic correlation angles.

The SVD (B3.18) of an n×p data set $\underline{D}$ produces an n×p matrix $\underline{A}'$ of principal components

$$\underline{A}' = [\underline{\alpha}_1 \ \ldots \ \underline{\alpha}_p] \tag{D1.1}$$

which summarize the temporal evolution of the data set. If we had another n×p data set $\underline{\tilde{M}}$, of SVD as in (C1.2), then it may be of interest to compare $\underline{A}'$ and the n×p matrix $\underline{B}'$ for similarities, where

$$\underline{B}' = [\underline{\beta}_1 \ \ldots \ \underline{\beta}_p] \ . \tag{D1.2}$$

A natural measure of distance between the frames $\underline{A}'$ and $\underline{B}'$ is COREL, as defined in §2 of the main text:

$$\begin{aligned} \text{COREL} &= \frac{1}{p} \|\underline{A}'-\underline{B}'\|^2 \\ &= \frac{2}{p}(1- \sum_{j=1}^{p} \underline{\alpha}_j^T \underline{\beta}_j) \end{aligned} \tag{D1.3}$$

In view of the developments in Appendix C, one is led to inquire whether we can have something here analogous to the canonic rotation angle θ_k. In Appendix C we had eigenframes $\underline{E},\underline{F}$ of $\underline{\tilde{D}}$ and $\underline{\tilde{M}}$ and the required rotation was $\underline{R} = \underline{F}\,\underline{E}^T$, whence the canonic rotation angles θ_k. In analogy to $\underline{R}$, one is tempted to

define a new rotation $\underline{S}$ of E_n into E_n in the present setting by simply writing "$\underline{S}$" for $\underline{B}'\underline{A}'^T$. However, when we test this for the rotational property: $\underline{S}^T\underline{S} = \underline{S}\,\underline{S}^T = \underline{I}_n$, we obtain $\underline{S}^T\underline{S} = \underline{A}'\underline{A}'^T$, which is generally not $\underline{I}_n$ (cf. $\underline{A}'^T\underline{A}' = \underline{I}_p$, by (B3.19)). We also obtain $\underline{S}\,\underline{S}^T = \underline{B}'\underline{B}'^T$, which again is generally not $\underline{I}_n$ (cf. $\underline{B}'^T\underline{B}' = \underline{I}_p$). Some investigating shows that there is generally no *unique* rotation in E_n that maps the orthonormal set $\underline{A}'$ into $\underline{B}'$. This is because generally we require $n-1 \geqq p$ for non-degenerate eigenvalue structures* in the SVD.

Suppose, however, that we remain within the subspace of E_n spanned by $\underline{A}'$ and that spanned by $\underline{B}'$. We then rotate $\underline{A}'$ within the subspace spanned by the vectors of $\underline{A}'$. At the same time we rotate $\underline{B}'$ within the subspace spanned by $\underline{B}'$. All the while these rotations are going on we monitor COREL in (D1.3), where the $\underline{\alpha}_j$ and $\underline{\beta}_j$ are now elements of the rotated $\underline{A}'$ and rotated $\underline{B}'$. There will generally exist an orientation of $\underline{A}'$ and an orientation of $\underline{B}'$ within their respective subspaces that together will minimize COREL. For data intercomparison purposes this minimum COREL will give a measure of the closeness of the subspaces spanned by $\underline{A}'$ and $\underline{B}'$ an hence the closeness of the temporal evolutions of $\underline{\tilde{D}}$ and $\underline{\tilde{M}}$. We now proceed to find this minimum COREL.

2. Minimizing COREL

Following the program outlined above, let $\underline{R},\underline{S}$ be n×n rotation matrices such that

$$\underline{R}\,\underline{\alpha}_j = \underline{a}_j \quad , \quad \underline{S}\,\underline{\beta}_j = \underline{b}_j \quad , \quad j = 1,\ldots,p \qquad \text{(D2.1)}$$

---

* Thus it follows that the space E_n is generally too big for the two orthonormal frames $\underline{A}' = [\underline{\alpha}_1 \cdots \underline{\alpha}_p]$ and $\underline{B}' = [\underline{\beta}_1 \cdots \beta_p]$ individually to span it, leaving many ways in which each of these frames can be extended to bases for E_n. For each extension there is a generally different rotation between the extended bases.

where $\underline{a}_j$ is in the subspace of E_n spanned by $[\underline{\alpha}_1,\ldots,\underline{\alpha}_p]$, and $\underline{b}_j$ is in the span of the $[\underline{\beta}_1,\ldots,\underline{\beta}_p]$. The latter sets are defined in Appendix B. Thus, we require that $\underline{a}_j,\underline{b}_j$ be linear combinations of the $\underline{\alpha}_j,\underline{\beta}_j$, respectively. Thus, for suitable numbers r_{jk}, and s_{jk}, we have:

$$\underline{a}_j = \underline{R}\,\underline{\alpha}_j = \sum_{k=1}^{p} r_{jk}\,\underline{\alpha}_k \quad , \quad j = 1,\ldots,p$$

$$\underline{b}_j = \underline{S}\,\underline{\beta}_j = \sum_{k=1}^{p} s_{jk}\,\underline{\beta}_k \quad , \quad j = 1,\ldots,p \tag{D2.2}$$

Moreover, we explicitly require that $\underline{R},\underline{S}$ do not change the lengths of the $\underline{\alpha}_j,\underline{\beta}_j$ when the latter are rotated. Thus we require of r_{jk},s_{jk}, that

$$1 = \underline{a}_j^T\underline{a}_j = \Big(\sum_{k=1}^{p} r_{jk}\,\underline{\alpha}_k\Big)^T\Big(\sum_{\ell=1}^{p} r_{j\ell}\,\underline{\alpha}_\ell\Big) = \sum_{k=1}^{p} r_{jk}^2 \tag{D2.3}$$

$$1 = \underline{b}_j^T\underline{b}_j = \sum_{k=1}^{p} s_{jk}^2 \quad , \quad j = 1,\ldots,p\ . \tag{D2.4}$$

Hence for the two rotated sets $A = [\underline{a}_1,\ldots,\underline{a}_p]$, $\underline{B} = [\underline{b}_1,\ldots,\underline{b}_p]$,

$$\begin{aligned}
\text{COREL}\ (\underline{A},\underline{B}) &= \frac{1}{p}\,\|\underline{A}-\underline{B}\|^2 \\
&= \frac{2}{p}\sum_{j=1}^{p}(1-\underline{a}_j^T\underline{b}_j) \\
&= \frac{2}{p}\sum_{j=1}^{p}\Big[1-\Big(\sum_{\ell=1}^{p} r_{j\ell}\underline{\alpha}_\ell\Big)^T\Big(\sum_{m=1}^{p} s_{jm}\underline{\beta}_m\Big)\Big] \\
&= \frac{2}{p}\sum_{j=1}^{p}\Big[1-\sum_{\ell=1}^{p}\sum_{m=1}^{p} r_{j\ell}s_{jm}\,\underline{\alpha}_\ell^T\underline{\beta}_m\Big]
\end{aligned} \tag{D2.5}$$

Equation (D2.5) shows that COREL($\underline{A}$,$\underline{B}$), for fixed $\underline{\alpha}_\ell$,$\underline{\beta}_m$, may be considered as a function of the $r_{j\ell}$ and s_{jm}. We seek a minimum of COREL($r_{j\ell}$,s_{jm}) subject to the conditions (D2.3), (D2.4). These conditions can be written as

$$\sum_{j=1}^{p} \sum_{\ell=1}^{p} r_{j\ell}^2 = p \tag{D2.6}$$

$$\sum_{j=1}^{p} \sum_{\ell=1}^{p} s_{jm}^2 = p \tag{D2.7}$$

Hence, if λ,μ are appropriate Lagrangian multipliers, we wish to simultaneously minimize the set of numbers:

$$\phi(r_{j\ell},s_{jm}) = \text{COREL}(r_{j\ell},s_{jm}) + \lambda \sum_{j=1}^{p} \sum_{\ell=1}^{p} r_{j\ell}^2 + \mu \sum_{j=1}^{p} \sum_{\ell=1}^{p} s^2{}_{jm} \tag{D2.8}$$

$j,\ell,\ m = 1,\ldots,p$

A necessary condition for this minimization is

$$\frac{\partial\phi(r_{j\ell},s_{jm})}{\partial r_{j\ell}} = 0 \quad , \quad j,\ell,\ m = 1,\ldots,p \tag{D2.9}$$

$$\frac{\partial\phi(r_{j\ell},s_{jm})}{\partial s_{jm}} = 0 \quad , \quad j,\ell,\ m = 1,\ldots,p \tag{D2.10}$$

Applying these operations to (D2.8) we find, for $j,\ell,\ m = 1,\ldots,p$,

$$\frac{\partial\phi}{\partial r_{j\ell}} = -\frac{2}{p} \sum_{m=1}^{p} s_{jm}\underline{\alpha}_\ell^T\underline{\beta}_m + \frac{2}{p}\lambda\, r_{j\ell} = 0 \tag{D2.11}$$

$$\frac{\partial\phi}{\partial s_{jm}} = -\frac{2}{p} \sum_{\ell=1}^{p} r_{j\ell}\underline{\alpha}_\ell^T\underline{\beta}_m + \frac{2}{p}\mu\, s_{jm} = 0 \tag{D2.12}$$

Using (D2.2) we can reduce the summations to:

$$\underline{\alpha}_{\ell}^{T}\underline{b}_{j} = \lambda\, r_{j\ell} \qquad j,\ell,m = 1,\ldots,p \tag{D2.13}$$

$$\underline{\beta}_{m}^{T}\underline{a}_{j} = \mu\, s_{jm} \tag{D2.14}$$

Multiplying (D2.13) by $\underline{\alpha}_{\ell}$ and (D2.14) by $\underline{\beta}_{m}$, summing over ℓ,m, respectively, and using (D2.2), we find: for j = 1,...,p,

$$\left(\sum_{\ell=1}^{p} \underline{\alpha}_{\ell}\underline{\alpha}_{\ell}^{T}\right) \underline{b}_{j} = \lambda_{j}\underline{a}_{j} \qquad j = 1,\ldots,p \tag{D2.15}$$

$$\left(\sum_{m=1}^{p} \underline{\beta}_{m}\underline{\beta}_{m}^{T}\right) \underline{a}_{j} = \mu_{j}\underline{b}_{j} \tag{D2.16}$$

Here we have added subscripts to the μ,λ, as suggested by the algebra. Write

$$'\underline{P}_{\alpha}' \text{ for } \sum_{\ell=1}^{p} \underline{\alpha}_{\ell}\underline{\alpha}_{\ell}^{T} \quad (= \underline{A}'\underline{A}'^{T}) \quad (n\times n) \qquad j = 1,\ldots,p \tag{D2.17}$$

$$'\underline{P}_{\beta}' \text{ for } \sum_{m=1}^{p} \underline{\beta}_{m}\underline{\beta}_{m}^{T} \quad (= \underline{B}'\underline{B}'^{T}) \quad (n\times n) \tag{D2.18}$$

Then (D2.15), (D2.16) become

$$\underline{P}_{\alpha}\underline{b}_{j} = \lambda_{j}\underline{a}_{j} \tag{D2.19}$$

$$\underline{P}_{\beta}\underline{a}_{j} = \mu_{j}\underline{b}_{j} \tag{D2.20}$$

These are the required equations governing $\underline{a}_{j},\underline{b}_{j}$. They form a pair of simultaneous eigenvalue problems for the $\underline{a}_{j},\underline{b}_{j}$, λ_{j} and u_{j}. The $\underline{a}_{j},\underline{b}_{j}$ satisfying these equations minimize COREL[$\underline{A},\underline{B}$].

3. Canonic Correlation Angles

A. To prepare for the numerical solution of (D2.19), (D2.20) and some interpretations, observe that $\underline{P}_\alpha, \underline{P}_\beta$ are *projections** on the spaces spanned by the $\underline{\alpha}$'s and $\underline{\beta}$'s respectively.

That is:

$$\underline{P}_\alpha^T = \underline{P}_\alpha \quad , \quad \underline{P}_\beta^T = \underline{P}_\beta \tag{D3.1}$$

and in particular

$$\underline{P}_\alpha \underline{a}_j = \underline{a}_j \quad , \quad \underline{P}_\beta \underline{b}_j = \underline{b}_j \quad , \quad j = 1,\ldots,p \tag{D3.2}$$

and so, also,

$$\underline{a}_j^T \underline{P}_\alpha = \underline{a}_j^T \quad , \quad \underline{b}_j^T \underline{P}_\beta = \underline{b}_j^T \quad , \quad j = 1,\ldots,p \tag{D3.3}$$

From these properties it follows that, by operating on (D2.19) with $\underline{a}_j^T$ and recalling (D2.3),

$$\underline{a}_j^T (\underline{P}_\alpha \underline{b}_j) = \lambda_j \, \underline{a}_j^T \underline{a}_j = \lambda_j \quad , \quad j = 1,\ldots,p$$

Similarly, operating on (D2.20) with $\underline{b}_j^T$ and recalling (D2.4),

$$\underline{b}_j^T (\underline{P}_\beta \underline{a}_j) = \mu_j \underline{b}_j^T \underline{b}_j = \mu_j \quad , \quad j = 1,\ldots,p$$

The left sides of these equations simplify by (D3.3) and are clearly equal to $\underline{a}_j^T \underline{b}_j$ so that we find

$$\underline{a}_j^T \underline{b}_j = \lambda_j = \mu_j \quad , \quad j = 1,\ldots,p \tag{D3.4}$$

---

* In our introductory remarks we encountered $\underline{P}_\alpha = \underline{A}'\underline{A}'^T$ and $\underline{P}_\beta = \underline{B}'\underline{B}'^T$ in our search for rotations in E_n. We now see the true nature of these matrices, i.e., they are *projections* and not rotations.

By this, the geometric interpretation of μ_j is that of the cosine of an angle between $\underline{a}_j$ and $\underline{b}_j$. This angle is ψ_j, the jth *canonic correlation angle*. "Correlation" is used in the sense established by (A2.21), now in the temporal domain (since $\underline{a}_j, \underline{b}_j$ are n×1 vectors and "n" denotes in this study the number of samples in time of a geophysical field).

B. We now reduce (D2.19), (D2.20) to forms readily evaluated numerically. Operating on the left of each side of (D2.19) with $\underline{P}_\beta$, and similarly on (D2.20) with $\underline{P}_\alpha$, and using these equations also to simplify the results, we find

$$\underline{P}_\beta \underline{P}_\alpha \, \underline{b}_j = \mu_j^2 \underline{b}_j \qquad j = 1,\ldots,p, (p+1,\ldots,n) \tag{D3.5}$$

$$\underline{P}_\alpha \underline{P}_\beta \, \underline{a}_j = \mu_j^2 \underline{a}_j \tag{D3.6}$$

We may therefore solve either one of these for the respective $\mu_j, \underline{b}_j$ or $\mu_j, \underline{a}_j$. The other set of vectors then follows from (D2.19), (D2.20), as the case may be. Observe that μ_j enters only as μ_j^2. Hence we take $\mu_j > 0$, and the associated angles ψ_j are in the range $0 \leq \psi_j \leq \pi/2$. We will comment on the extended j-range in (D3.5), (D3.6), in §4, below.

To begin the calculations, consider (D3.6). Then by definition of matrix products

$$(\underline{P}_\alpha \underline{P}_\beta)_{ij} = \sum_{\ell=1}^{n} (\underline{P}_\alpha)_{i\ell} \, (\underline{P}_\beta)_{\ell j} \quad , \quad i,j = 1,\ldots,n. \tag{D3.7}$$

From (D2.17), (D2.18), for $i,j,\ell = 1,\ldots,n$,

$$(\underline{P}_\alpha)_{i\ell} = \sum_{m=1}^{p} \alpha_m(i) \, \alpha_m(\ell) \tag{D3.8}$$

$$(\underline{P}_\beta)_{\ell j} = \sum_{m=1}^{p} \beta_m(\ell) \, \beta_m(j) \tag{D3.9}$$

where $\underline{\alpha}_m = [\alpha_m(1),\ldots,\alpha_m(n)]^T$, m = 1,...,p, as defined in (B3.17). Hence the n×n matrix $\underline{P}_\alpha\underline{P}_\beta$ needed in (D3.6) is fully defined in terms of the orthonormalized principal components of the data sets $\underline{\tilde{D}},\underline{\tilde{M}}$. The remaining details of solution of (D3.6) proceed as in any other eigenvalue problem.

4. Orthogonal Projections $\underline{P}_\alpha$, $\underline{P}_\beta$ on E_n

Observe that the routine solution of (D3.6), e.g., yields n (rather than just p) vectors $\underline{a}_j$. We began our analysis in (D2.1) restricting attention only to the first p vectors $\underline{a}_j$, j = 1,...,p. Generally the remaining vectors $\underline{a}_{p+1},\ldots,\underline{a}_n$ in (D3.6) are outside of the A′-space and are associated with zero μ_j's. For example, suppose $\underline{a}_{p+1}$, is a non zero vector in the orthogonal complement* of the $\underline{A}'$-space. There are two cases (see Fig. D4.1 drawn for the case n-1 = 3, p = 2): (i) either $\underline{a}_{p+1}$ is in the $\underline{B}'$-space or (ii) $\underline{a}_{p+1}$ is in the orthogonal complement of the $\underline{B}'$-space. In case (i), $\underline{P}_\beta\underline{a}_{p+1} = \underline{a}_{p+1}$ by (D3.2), and then $\underline{P}_\alpha(\underline{P}_\beta\underline{a}_{p+1}) = \underline{P}_\alpha\underline{a}_{p+1} = 0$ by hypothesis and (D2.17). Hence $\mu_{p+1} = 0$. In case (ii), $(\underline{P}_\beta\underline{a}_{p+1}) = 0$ directly (by (D2.18)) and so $\mu_{p+1} = 0$. Hence in general the $\underline{a}_j$,j = p+1,...,n belong to zero eigenvalues μ_j. On the other hand, those μ_j in the j range 1,...,p, that are zero, are of important physical interest. An estimate of the number of such μ_j will be made in §6 of this Appendix.

We observe further that those $\underline{a}_j,\underline{a}_k$ belonging to distinct μ_j,μ_k are necessarily orthogonal. To see this, multiply each side of (D3.6) in turn by $\underline{a}_j^T$ and $\underline{a}_k^T$, for an arbitrary choice of j, k = 1,...,p:

---

* E_n can be represented as a direct sum of the space spanned by $\underline{A}'$ and another subspace of E_n, all of whose vectors are orthogonal to all of those in $\underline{A}'$. This subspace is the *orthogonal complement* of $\underline{A}'$. See, e.g. Fig. D4.1.

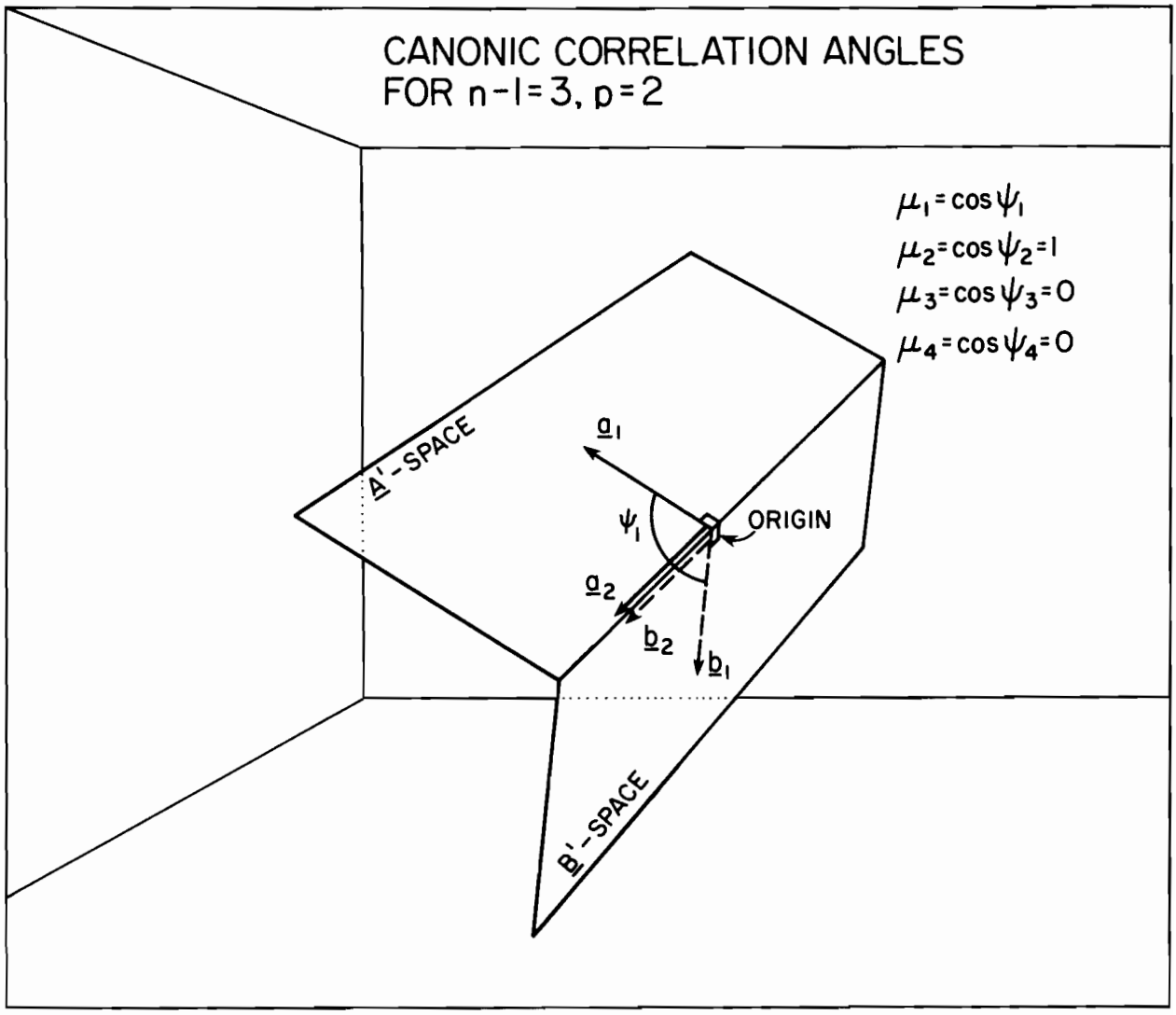

CANONIC CORRELATION ANGLES
FOR n-1=3, p=2
$\mu_1 = \cos\psi_1$
$\mu_2 = \cos\psi_2 = 1$
$\mu_3 = \cos\psi_3 = 0$
$\mu_4 = \cos\psi_4 = 0$
A'-SPACE
$\underline{a}_1$
ψ_1
ORIGIN
$\underline{a}_2$
$\underline{b}_2$
$\underline{b}_1$
B'-SPACE

Fig. D4.1

$$\underline{a}_j^T (\underline{P}_\alpha \underline{P}_\beta \underline{a}_k) = \mu_k^2 \underline{a}_j^T \underline{a}_k \tag{D4.1}$$

$$\underline{a}_k^T (\underline{P}_\alpha \underline{P}_\beta \underline{a}_j) = \mu_j^2 \underline{a}_k^T \underline{a}_j \tag{D4.2}$$

Now, the left sides of these equations are equal. To see this, we have from (D4.1), (D3.3), and (D3.1):

$$(\underline{a}_j^T \underline{P}_\alpha) \underline{P}_\beta \underline{a}_k = \underline{\alpha}_j^T \underline{P}_\beta \underline{a}_k = \underline{\alpha}_k^T \underline{P}_\beta \underline{\alpha}_j$$

(The last equality follows from the fact that the transpose of a scalar results again in that scalar.) Further from (D4.2),

$$(\underline{a}_k^T \underline{P}_\alpha) \underline{P}_\beta \underline{a}_j = \underline{\alpha}_k^T \underline{P}_\beta \underline{a}_j$$

Thus the left sides of (D4.1), (D4.2) are equal. Therefore, from (D4.1) (D4.2)

$$(\mu_j^2 - \mu_k^2) \, \underline{a}_j^T \underline{a}_k = 0$$

whence

$$\underline{a}_j^T \underline{a}_k = 0 \quad , \quad j \neq k \, .$$

Since the $\underline{a}_j$ are by construction unit vectors (cf. (D2.3)), we may write the preceding result as:

$$\underline{a}_j^T \underline{a}_k = \delta_{jk} \quad , \quad j,k = 1,\ldots,p. \tag{D4.3}$$

In an exactly similar manner, we can show

$$\underline{b}_j^T \underline{b}_k = \delta_{jk} \quad , \quad j,k = 1,\ldots,p \tag{D4.4}$$

when the associated eigenvalues μ_j, μ_k are distinct.

Finally, and most interestingly, we have

$$\underline{a}_k^T \underline{b}_j = \mu_j \delta_{jk} \quad j,k = 1,\ldots,p. \tag{D4.5}$$

To see this, observe that from (D3.6) and (D4.3), we have, for j, k = 1,...,p,

$$\underline{a}_k^T (\underline{P}_\alpha \underline{P}_\beta \underline{a}_j) = \mu_j^2 \underline{a}_k^T \underline{a}_j = \mu_j^2 \delta_{jk}$$

The left side simplifies via (D3.3), (D2.20) for j, k = 1,...,p,

$$(\underline{a}_k^T \underline{P}_\alpha)\ \underline{P}_\beta \underline{a}_j = \underline{\alpha}_k^T \underline{P}_\beta \underline{a}_j = \underline{a}_k^T \mu_j \underline{b}_j$$

On comparing these two results, (D4.5) follows.

When j = k in (D4.5), we return to (D3.4).

5. COREL via Canonic Correlation Angles

It is difficult to make a picture that illustrates (D4.3), (D4.4), (D4.5) in a non-trivial setting. However, the sketch in Fig. D4.1 may be of help in visualizing the relations for the case n-1 = 3, p = 2. Note that we use "n-1 = 3" instead of "n = 3". This is in accordance with the fact that the space of the vectors $\underline{A}', \underline{B}'$ is actually an n-1 dimensional subset of E_n. (For more discussion of this, see §6 below.)

Continuing to examine Fig. D4.1, we see that in the $\underline{A}'$-space we have $\underline{a}_1, \underline{a}_2$ orthogonal to each other, as also are $\underline{b}_1, \underline{b}_2$ in the $\underline{B}'$-space. The rotating-minimizing action on COREL (cf. (D2.8), (D2.9), (D2.10)) forces $\underline{a}_2$ and $\underline{b}_2$ to coincide at the intersection of the two subspaces, and constrains $\underline{a}_1, \underline{b}_1$ to remain normal to $\underline{a}_2, \underline{b}_2$, respectively. As a result, the canonic correlation angle ψ_1 measures the angle of separation of the two subspaces, a situation,

by virtue of (D4.5), to be generally the case. In this way we see that (D2.5) for the minimum value of COREL becomes, via (D3.4):

$$\begin{aligned} \text{COREL}(\underline{A},\underline{B}) &= \frac{2}{p}\sum_{j=1}^{p}(1-\underline{a}_j^T\underline{b}_j) \\ &= \frac{2}{p}\sum_{j=1}^{p}(1-\mu_j) \\ &= \frac{2}{p}\sum_{j=1}^{p}(1-\cos\psi_j) \leq \text{COREL}(\underline{A}',\underline{B}'). \end{aligned} \tag{D5.1}$$

Observe that, in computing the ψ_j, we are using the $\underline{a}_j$ and $\underline{b}_j$ of the the minimizing frames $\underline{A},\underline{B}$ instead of the vectors of the original data-derived frames $\underline{A}',\underline{B}'$.

6. Estimating the Number of Expected Zero Canonic Correlation Angles

A. When computing the canonic correlations μ_j, some of them in the range $1 \leq j \leq p$ will necessarily be of unit magnitude, indicating a corresponding zero correlation angle ψ_j. The more of these ψ_j that are zero, the closer do the subspaces spanned by $\underline{A}'$ and $\underline{B}'$ lie to each other, and the more of the time series $\{\alpha_j(t): t = 1,\ldots,n\}$, $j = 1,\ldots,p$ that can be represented by the $\{\beta_j(t): t = 1,\ldots,n\}$, $j = 1,\ldots,p$. In what follows we give two ways of estimating the numbers of ψ_j expected to be zero. One way is exact via numerical procedures, the other is approximate by general reasoning.

B. Exact Determination of s

From the general theory of finite dimensional vector spaces we have the following theorem (Halmos, 1958, p. 19). Let P,Q be two subspaces of an n-dimensional vector space. Let the dimensions of P,Q be respectively p,q.

APPENDIX D

Let r be the dimension of the union of P and Q (the set of all linear combinations of the pooled set of vectors in P and Q), and let s be the dimension of the intersection of P and Q (the subspace of all vectors common to P and Q). Then

$$p+q = r+s \ . \tag{D6.1}$$

It can be shown that the dimension s of the intersection of P and Q is the expected number of zero correlation angles.

This fact can be illustrated in Fig. D4.1. First of all we recall (Appendix B, below (B3.22)) that space-centering the data sets reduces the span of the $\underline{A}', \underline{B}'$ vectors to a common n-1 dimensional subspace of E_n. For, if we have $\{\underline{u}(t): \ t = 1,\ldots,n\}$ as the orthonormal basis of E_n in which the data values of $\underline{\alpha}_j, \underline{\beta}_j$ are represented, then

$$\underline{\alpha}_j = \sum_{t=1}^{n} \alpha_j(t)\, \underline{u}(t) \quad , \quad j = 1,\ldots,p \tag{D6.2}$$

But by space centering,

$$\sum_{t=1}^{n} \alpha_j(t) = 0 \quad , \quad j = 1,\ldots,p \tag{D6.3}$$

Hence we can write (D6.2), for j = 1,...,p, as

$$\underline{\alpha}_j = \sum_{t=1}^{n-1} \alpha_j(t)[\underline{u}(t)-\underline{u}(n)] \tag{D6.4}$$

In a similar way we have, for j = 1,...,p,

$$\underline{\beta}_j = \sum_{t=1}^{n-1} \beta_j(t)\, [\underline{u}(t)-\underline{u}(n)] \tag{D6.5}$$

The set of vectors $\{\underline{u}(t)-\underline{u}(n): \ t = 1,\ldots,n-1\}$ is easily seen to be linearly independent, and therefore forms a basis of an n-1 dimensional subset of E_n. In Fig. D4.1, the common n-1 dimensional setting for $\underline{A}', \underline{B}'$ is drawn as three dimensional, and the latter basis sets have p = 2. In (D6.1), therefore, we have p = q = 2. As drawn in the Figure, we estimate visually, that r = 3. Hence s = (p+q)-r = 4-3 = 1. The two vectors $\underline{a}_2, \underline{b}_2$ are shown lying in the one dimensional intersection of the A′-space (= P) and the B′-space (= Q). Thus $\mu_2 = \cos\psi_2 = 1$, and $\psi_2 = 0$ between $\underline{a}_2$ and $\underline{b}_2$.

We now return to the main thread of the argument leading to an exact determination of s. The procedure is numerical. Start with the sets $\underline{A}'_p \equiv \{\underline{\alpha}_1 \cdots \underline{\alpha}_p\}$ and $\underline{B}'_p \equiv \{\underline{\beta}_1 \cdots \underline{\beta}_p\}$. Hence in (D6.1), p = q. The number ρ of linearly independent vectors in $\underline{A}'_p$ is assumed to be p and less than n-1. (If p = n-1, then r = n-1 and s = n-1.) Note that, in general, ρ = min[n-1,p]. Recall that, by construction, the $\underline{\alpha}_j$ are pairwise orthogonal and of unit length. We now determine r by starting with the estimate r = p. Then we take $\underline{\beta}_1$ from $\underline{B}'_p$ and find its representation in the $\underline{A}'_p$ frame:

$$\underline{\beta}'_1 = \sum_{j=1}^{p} (\underline{\beta}_1^T \underline{\alpha}_j)\, \underline{\alpha}_j \ . \qquad \text{(D6.6)}$$

Next, find $\|\underline{\beta}_1 - \underline{\beta}'_1\|$. If this is zero (say to within* ε), then we conclude that $\underline{\beta}_1$ is in the span of $\underline{A}'_p$. If $\|\underline{\beta}_1 - \underline{\beta}'\| > \varepsilon$, then $\underline{\beta}_1$ is not in the span of $\underline{A}'_p$ and we add $\underline{\beta}_1$ to the $\underline{A}'_p$-frame, and thereby produce a new orthonormal basis $\underline{A}'_{p+1}$, using the Gram-Schmidt procedure. As a result of these examinations, we update $\underline{A}'_p$ to $\underline{A}'_{p+1}$ or leave it as $\underline{A}'_p$, as the case may be. Next we take $\underline{\beta}_2$ from

---

* For example $\varepsilon = 10^{-10}$.

$\underline{B}'_p$ and form its representation in the updated $\underline{A}'$-frame, find $\|\underline{\beta}_2 - \underline{\beta}'_2\|$, and decide to augment the $\underline{A}'$-frame or not. In this manner, we eventually go through $\underline{B}'_p$, pick up, say, a new vectors beyond $\underline{A}'_p$ and the end result is an augmented $\underline{A}'$-frame of dimension p+a. This p+a is, by construction, precisely r. Then from (D6.1), since p = q,

$$s = 2p-r = p-a \tag{D6.7}$$

is our required exact determination of the number of zero canonic correlation angles.

C. Approximate Estimate of s

The approximate estimate of s depends on the following two intuitive insights into the growth of the dimension r with dimension p. The first insight is as follows. When 2p is small relative to (n-1), then (with probability 1) the union of the subspaces spanned by two randomly chosen $\underline{A}'$ and $\underline{B}'$, each of dimension p, has dimension r = 2p, and so by (D6.7), we have s = 0. As p grows, r continues to keep pace with 2p and s remains zero until we reach 2p = n-1. The second insight is as follows. At this point, with probability 1, r attains the limit n-1 and it cannot grow larger than n-1; but of course p can continue on to n-1. In this latter range, then, (D6.7) is of the form s = 2p-(n-1). These two intuitions can be summarized succinctly by the formula:

$$s \equiv (\text{number of } \psi_j \text{ equal to zero}) = \max[0,\ 2p-(n-1)] \tag{D6.8}$$

Since the total number of ψ_j of interest is p (recall from §4 of this Appendix that n-p are necessarily 90°, corresponding to $\mu_j = 0$), we also can say:

$$t \equiv (\text{number of } \psi_j \text{ greater than zero}) = p - \max[0,2p-(n-1)] \tag{D6.9}$$

The following table gives a simple illustration of these rules of thumb for the case n = 10.

p	number of ψ_j equal to 0° s = max[0,2p-(n-1)]	number of nonzero ψ_j t = p-s	number of ψ_j equal to 90° n-p
1	0	1	9
2	0	2	8
3	0	3	7
4	0	4	6
5	1	4	5
6	3	3	4
7	5	2	3
8	7	1	2
9	9	0	1

From this we see that the desirable research setting is one in which n-1 exceeds 2p, so that s = 0. In this way there is no geometrically forced coalescence of the spaces spanned by $\underline{A}'$ and $\underline{B}'$. Then, if COREL($\underline{A}$,$\underline{B}$) is significantly small under these conditions, the corresponding members of the $\underline{A}'$,$\underline{B}'$ frames of the data sets would be close in temporal evolution, or at least could closely represent one another by appropriate linear combinations.

APPENDIX E

The Distribution of a Component of a Random Unit Vector in E_p and its Link with the Distribution of the Correlation Coefficient

1. Introduction

The discussion in this Appendix provides a reference probability distribution for the significance test in §7 dealing with the comparison of two eigenvectors of a given pair of data matrices, $\underset{\sim}{\tilde{D}}$ and $\underset{\sim}{\tilde{M}}$. The discussion, interestingly, also informally provides a link between such a distribution and that of the classical correlation coefficient between two samples from a population of uncorrelated gaussian variates.

2. The Distribution of a Component of a Random Unit Vector in E_p

We construct a random unit vector in E_p, $p \geq 2$, as follows. We randomly draw p samples from N(0,1). Let these be γ_j, $j = 1,\ldots,p$. Then the set of random variates $\{\gamma_j/(\sum_{k=1}^{p} \gamma_k^2)^{\frac{1}{2}}: \ j = 1,\ldots,p\}$ comprises the components of a random unit vector in E_p. The squares of these components, namely

$$y_j = \gamma_j^2 / \sum_{k=1}^{p} \gamma_k^2 \qquad \text{(E2.1)}$$

are such that the γ_j^2, $j = 1,\ldots,p$ are independent chi-square variates, each of a single degree of freedom. Moreover, if we write (E2.1) as:

$$y_j = \frac{\gamma_j^2}{\gamma_j^2 + \delta_j^2} \quad , \quad j = 1,\ldots,p \ . \qquad \text{(E2.2)}$$

where

$$\delta_j^2 = \sum_{\substack{k=1 \\ k\neq j}}^{p} \gamma_k^2 \qquad \text{(E2.3)}$$

then we recognize y_j as a random variate with the beta distribution of 1 and p-1 degrees of freedom. That is, in general, if x_1 is a chi-square variate with k_1 degrees of freedom and x_2 is an independent chi-square variate with k_2 degrees of freedom, then the pdf of $y = x_1/(x_1+x_2)$ has the form*

$$p(y) = \frac{\Gamma(\frac{1}{2}(k_1+k_2))}{\Gamma(\frac{1}{2}k_1)\Gamma(\frac{1}{2}k_2)}\, y^{\frac{1}{2}k_1-1}\,(1-y)^{\frac{1}{2}k_2-1} \qquad \text{(E2.4)}$$

In the case of (E2.2), since γ_j^2 and δ_j^2 are independent, and $k_1 = 1$ and $k_2 = p-1$, we find,

$$p(y_j)dy_j = \frac{\Gamma(\frac{1}{2}p)}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}(p-1))}\, y_j^{-\frac{1}{2}}(1-y_j)^{\frac{1}{2}(p-3)}dy_j \qquad \text{(E2.5)}$$

for $j = 1,\ldots,p$, and $0 \leq y_j \leq 1$.

We are at present interested not in y_j but $x \equiv \pm y_j^{\frac{1}{2}}$, the components of the random vector. Moreover, it is clear that the pdf of y_j is independent of j. With this change of variables to x, so that $2xdx = dy_j = 2y_j^{\frac{1}{2}}dx$, (E2.5) becomes

$$\begin{aligned} p(y_j)dy_j &= \frac{\Gamma(\frac{1}{2}p)}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}(p-1))}\,(1-x^2)^{\frac{1}{2}(p-3)}dx \\ &\equiv q(x)dx \end{aligned} \qquad \text{(E2.6)}$$

---

* For a derivation of this well known pdf, from first principles, see Preisendorfer (1979), Equation (A50). Otherwise, see Rao (1973), p. 167.

where $-1 \leqq x \leqq 1$, and "x" denotes any of the components $\pm y_j^{\frac{1}{2}}$. In this way we attain the required distribution q(x) of the projection x of a random unit vector on a fixed unit vector. The factor "2" in the differential connection above was accounted for by going from the range $0 \leqq y_j \leqq 1$ to the range $-1 \leqq x \leqq 1$. The integral of q(x) over the latter range is unity. By symmetry, the first moment of x is 0, and by a simple calculation using (E2.6), the second moment of x (its variance) is 1/p. This shows that the "widths" of the pdfs in Figs. 4.1-4.8 go to zero approximately as $1/p^{\frac{1}{2}}$.

3. The Link with the Correlation Coefficient

The classical correlation coefficient arises in the following typical sampling operation. Draw p random samples $x_1,\ldots,x_p$ from N(0,1) and again another p random samples $y_1,\ldots,y_p$ from N(0,1), and then form the correlation coefficient:

$$r = \left(\sum_{j=1}^{p} x_j y_j\right)/\left(\sum_{k=1}^{p} x_k^2\right)^{\frac{1}{2}}\left(\sum_{k=1}^{p} y_k^2\right)^{\frac{1}{2}} \qquad \text{(E3.1)}$$

This then is equivalent to the exercise in §2 above of finding the inner product $\underline{x}^T\underline{y}$ of the two associated unit vectors formed via each x_j or y_j being divided by the square root of the sum of squares of the x_k or y_k, respectively, as shown in (E3.1). A moment's thought will show that this inner product of two random unit vectors $\underline{x},\underline{y}$ is equivalent, in the derivation of §2, to finding the component of $\underline{x}$ along a fixed unit vector $\underline{y}$, provided we reduce the number of independent components of $\underline{x}$ by 1, i.e., from p to p-1. (In effect $\underline{x}$ now lies in a variable 2 dimensional plane determined by it and $\underline{y}$. Hence, as seen from $\underline{y}$'s vantage, $\underline{x}$ has p-1 degrees of freedom, for $p \geqq 3$.) Thus we have at once from (E2.6), on replacing "p" by "p-1":

$$q(r) = \frac{\Gamma(\frac{1}{2}(p-1))}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}(p-2))} (1-r^2)^{\frac{1}{2}(p-4)} \quad , \quad p \geq 3 \qquad \text{(E3.2)}$$

for $-1 \leq r \leq 1$, which is the classical pdf for the sample correlation coefficient for p pairs of random samples from an uncorrelated bivariate normal distribution of zero population means.

NOAA ERL technical reports, technical memoranda, and data reports published by authors at Pacific Marine Environmental Laboratory in Seattle, Washington, are listed below. Microfiche copies are available from the USDOC, National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, Virginia 22161 (703-487-4650). Hard copies of some of these publications are available from the ERL Library in Boulder, Colorado (303-497-3271). Hard copies of some of the technical reports are sold by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (202-275-9251).

NOAA Technical Report Series

ERL82-POL1 Naugler, Frederic P. (1968)
Bathymetry of a region (PORL-421-2) North of the Hawaiian Ridge, pre-NTIS.

ERL93-POL2 Grim, Paul J. (1968)
Seamap deep-sea channel, Jan. 1969, 2 824 50 060, pre-NTIS.

ERL118-POL3 Le Mehaute, Bernard (1969)
An introduction to hydrodynamics and water waves, 2 vols. 725 pp.
NTIS: PB192 065, PB192 066.

ERL146-POL4 Rea, David K. (1970)
Bathymetry and magnetics of a region (POL-421-3) 29° to 35°N, 155° to 165°W.
NTIS: COM-71-00173.

ERL191-POL5 Reed, R.K. (1970)
Results from some parachute drogue measurements in the central North Pacific Ocean, 1961-1962, 9 pp.
NTIS: COM-71-50020.

ERL214-POL6 Lucas, William H. (1971)
Gravity anomalies and their relation to major tectonic features in the North Central Pacific, 19 pp.
NTIS: COM-71-50409.

ERL229-POL7 Halpern, David (1972)
Current meter observations in Massachusetts Bay, 36 pp.
NTIS: AD-745 465.

ERL230-POL8 Lucas, William H. (1972)
South Pacific RP-7-SU-71 Pago Pago to Callao to Seattle.
NTIS: COM-72-50454.

ERL231-POL9 Halpern, David (1972)
Description of an experimental investigation on the the response of the upper ocean to variable winds, 51 pp.
NTIS: COM-72-50452.

ERL232-POL10 Stevens, H. R., Jr. (1972)
RP-1-OC-71 Northeast Pacific geophysical survey, 91 pp.
NTIS: COM-72-50677.

ERL234-POL11 Lucas, William H. (1972)
Juan de Fuca Ridge and Sovanco fracture zone.
RP-5-OC-71, 39 pp.
NTIS: COM-72-50854.

ERL240-POL12 Halpern, David (1972)
Wind recorder, current meter and thermistor chain measurements in the northeast Pacific-August/September 1971, 37 pp.
NTIS: COM-73-50107.

ERL247-POL13 Cannon, G. A. and Norman P. Laird (1972)
Observations of currents and water properties in Puget Sound, 1972, 42 pp.
NTIS: COM-73-50402.

ERL252-POL14 Cannon, G. A., N. P. Laird, T. V. Ryan (1973)
Currents observed in Juan de Fuca submarine canyon and vicinity, 1971. 57 pp.
NTIS: COM-73-50401.

ERL258-POL15 Lucas, William H., and Richard R. Uhlhorn (1973)
Bathymetric and magnetic data from the northeast Pacific 40° to 58°N, 125° to 160°W. 9 pp.
NTIS: COM-73-50577.

ERL259-POL16 Ryan, T. V., N. P. Laird, G. A. Cannon (1973)
RP-6-OC-71 Data Report: Oceanographic conditions off the Washington coast, October-November 1971, 43 pp.
NTIS: COM-73-50922

ERL260-POL17 Cannon, Glenn A. (1973)
Observations of currenta in Puget Sound, 1970, 77 pp.
NTIS: COM-73-50666/9.

ERL261-POL18 Stevens, H. R. Jr., (1973)
RP-1-OC-70 Southeast Pacific geophysical survey, 60 pp.
NTIS: not available.

ERL271-POL19 Reed, Ronald K., and David Halpern (1973)
STD observationa in the northeast Pacific, September-October 1972, 58 pp.
NTIS: COM-73-50923/4.

ERL292-PMEL20 Reed, R. K. (1973)
Distribution and variation of physical properties along the SEAMAP standard section, 16 pp.
NTIS: COM-74-50334/3.

ERL323-PMEL21 Erickson, B. H. (1975)
Nazca plate program of the international decade of ocean exploration--OCEANOGRAPHER Cruise-RP 2-OC-73, 78 pp.
NTIS: COM-7540911/6.

ERL325-PMEL22 Halpern, D., J. M. Helseth, J. R. Holbrook, and R. M. Reynolds (1975)
Surface wave height measurements made near the Oregon coast during August 1972, and July and August 1973, 168 pp.
NTIS: COM-75-10900/9.

ERL327-PMEL23 Laird, N. P., and Jerry A. Galt (1975)
Observations of currents and water properties in Puget Sound, 1973, 141 pp.
NTIS: COM-73-50666/9.

ERL333-PMEL24 Schumacher, J. D., and R. M. Reynolds (1975)
STD, current meter, and drogue observations in Rosario Strait, January-March 1974, 212 pp.
NTIS: COM-75-11391/0.

ERL339-PMEL25 Galt, J. A. (1975)
Development of a simplified diagnostic model for interpretation of oceanographic data.
NTIS: PB-247 357/7.

ERL352-PMEL26 Reed, R. K., (1975)
An evaluation of formulas for estimating clear-sky insolation over the ocean, 25 pp.
NTIS: PB-253 055/8.

ERL384-PMEL27 Garwood, Roland (1977)
A general model of the ocean mixed layer using a two-component turbulent kinetic energy budget with mean turbulent field closure, 81 pp.
NTIS: PB-265 434/1.

ERL390-PMEL28 Hayes, S. P., and W. Zenk (1977)
Observations of the Antarctic Polar Front by a moored array during FDRAKE-76, 47 pp.
NTIS: PB-281 460/6.

ERL390-PMEL29 Hayes, S. P., and W. Zenk (1977)
Observations of the Antarctic Polar Front by a moored array during FDRAKE-76, 49 pp.
NTIS: PB-281 460/6.

ERL403-PMEL30 Chester, Alexander J. (1978)
Microzooplankton in the surface waters of the Strait of Juan de Fuca, 26 pp.
NTIS: PB 297233/AS.

ERL404-PMEL31 Schumacher, J. D., R. Sillcox, D. Dreves, and R. D. Muench (1978)
Winter circulation and hydrography over the continental shelf of the northwest Gulf of Alaska, 16 pp.
NTIS: PB 296 914/AS.

ERL407-PMEL32 Overland, J. E., M. H. Hitchman, and Y. J. Han (1979)
A regional surface wind model for mountainous coastal areas, 34 pp.
NTIS: PB 80 146 152.

ERL412-PMEL33 Holbrook, J. R., R. D. Muench, D. G. Kachel, and C. Wright (1980)
Circulation in the Strait of Juan de Fuca: Recent oceanographic observations in the Eastern Basin, 42 pp.
NTIS: PB 81-135352.

ERL415-PMEL34 Feely, R. A., and G. J. Massoth (1982)
Sources, composition, and transport of suspended particulate matter in lower Cook Inlet and northern Shelikof Strait, Alaska, 28 pp.
NTIS: PB 82-193263

ERL417-PMEL35 Baker, E. T. (1982)
Suspended particulate matter in Elliott Bay, 44 pp.
NTIS: PB 82-246943.

ERL419-PMEL36 Pease, C. J., S. A. Schoenberg, J. E. Overland (1982)
A climatology of the Bering Sea and its relation to sea ice extent, 29 pp.
NTIS: not yet available.

ERL422-PMEL37 Reed, R. K. (1982)
Energy fluxes over the eastern tropical Pacific Ocean, 1979-1982, 15 pp.
NTIS: PB 83 138305

NOAA Data Report Series

ERL PMEL-1 Mangum, L., N. N. Soreide, B. D. Davies, B. D. Spell, and S. P. Hayes (1980)
CTD/O_2 measurements during the equatorial Pacific Ocean climate study (EPOCS) in 1979, 643 pp.
NTIS: PB 81 211203.

ERL PMEL-2 Katz, C. N., and J. D. Cline (1980)
Low molecular weight hydrocarbon concentrations (C_1-C_4), Alaskan continental shelf, 1975-1979, 328 pp.
NTIS: PB 82 154211.

ERL PMEL-3 Taft, B. A., and P. Kovala (1981)
Vertical sections of temperature, salinity, thermosteric anomaly, and zonal geostrophic velocity from NORPAX shuttle experiment, part 1, 98 pp.
NTIS: PB 82 163106.

ERL PMEL-4 Pullen, P. E., and H. Michael Byrne (1982)
Hydrographic measurements during the 1978 cooperative Soviet-American tsunami expedition, 168 pp.
NTIS: not yet available.

ERL PMEL-5 Taft, B.A., P. Kovala, and A. Cantos-Figuerola (1982)
Vertical sections of temperature, salinity, thermosteric anomaly and zonal geostrophic velocity from NORPAX Shuttle Experiment--Part 2, 94 pp.
NTIS:

ERL PMEL-6 Katz, C.N., J.D. Cline, and K. Kelly-Hansen (1982)
Dissolved methane concentrations in the southeastern Bering Sea, 1980 and 1981, 194 pp.
NTIS:

NOAA Technical Memorandum Series

ERL PMEL-1 Sokolowski, T. J. and G. R. Miller (1968)
Deep sea release mechanism, Joint Tsunami Research Effort, pre-NTIS.

ERL PMEL-2 Halpern, David (1972)
STD observations in the northeast Pacific near 47°N, 128°W (August/September 1971), 28 pp.
NTIS: COM-72-10839.

ERL PMEL-3 Reynolds, R. Michael and Bernard Walter, Jr. (1975)
Current meter measurements in the Gulf of Alaska--Part 1: Results from NEGOA moorings 60, 61, 62A, 28 pp.
NTIS: PB-247 922/8.

ERL PMEL-4 Tracy, Dan E. (1975)
STD and current meter observations in the north San Juan Islands, October 1973.
NTIS: PB-248 825/2.

ERL PMEL-5 Holbrook, James R. (1975)
STD measurements off Washington and Vancouver Island during September 1973.
NTIS: PB-249 918/4.

ERL PMEL-6 Charnell, R. L. and G. A. Krancus (1976)
A processing system for Aanderaa current meter data, 53 pp.
NTIS: PB-259 589/0.

ERL PMEL-7 Moljeld, Harold O. and Dennis Mayer (1976)
Formulas used to analyze wind-driven currents as first-order autoregressive processes, 22 pp.
NTIS: PB-262 463/3.

ERL PMEL-8 Reed, R. K. (1976)
An evaluation of cloud factors for estimating insolation over the ocean, 23 pp.
NTIS: PB-264 174/4.

ERL PMEL-9 Nakamura, A. I. and R. R. Harvey (1977)
Versatile release timer for tree vehicle instrumentation over the ocean, 21 pp.
NTIS: PB 270321/AS.

ERL PMEL-10 Holbrook, James R. and David Halpern (1977)
A compilation of wind, current, bottom pressure, and STD/CTD measurements in the northeast Gulf of Alaska, February-May 1975.
NTIS: PB 270285.

ERL PMEL-11 Nakamura, A. I. and R. R. Harvey (1978)
Conversion from film to magnetic cassette recording for the Geodyne 102 current meter, 17 pp.
NTIS: PB-283 349/9.

ERL PMEL-12 Hayes, S. P., J. Glenn, N. Soreide (1978)
A shallow water pressure-temperature gage (PTG): Design, calibration, and operation, 35 pp.
NTIS: PB 286 754/7.

ERL PMEL-13 Schumacher, J. D., R. K. Reed, M. Grigsby, D. Dreves (1979)
Circulation and hydrography near Kodiak Island, September to November 1977, 52 pp.
NTIS: PB 297421/AS.

ERL PMEL-14 Pashinski, D. J., and R. L. Charnell (1979)
Recovery record for surface drift cards released in the Puget Sound-Strait of Juan de Fuca system during calendar years 1976-1977, 32 pp.
NTIS: PB 299047/AS.

ERL PMEL-15 Han, Y.-J. and J. A. Galt (1979)
A numerical investigation of the Bering Sea circulation using a linear homogeneous model, 40 pp.
NTIS: PB 299884/AS.

ERL PMEL-16 Loomis, Harold G. (1979)
A primer on tsunamis written for boaters in Hawaii, 10 pp.
NTIS: PB80-161003.

ERL PMEL-17 Muench, R. D. and J. D. Schumacher (1980);
(Hayes, Charnell, Lagerloef, and Pearson, contributors)
Some observations of physical oceanographic conditions on the northeast Gulf of Alaska continental shelf, 90 pp.
NTIS: PB81-102584.

ERL PMEL-18 Gordon, Howard R., ed. (1980)
Ocean remote sensing using lasers, 205 pp.
NTIS: PB80-223282.

ERL PMEL-19 Cardone, V. J. (1980)
Case studies of four severe Gulf of Alaska storms, 58 pp.
NTIS: PB81-102519.

ERL PMEL-20 Overland, J. E., R. A. Brown, and C. D. Mobley (1980)
METLIB--A program library for calculating and plotting marine boundary layer wind fields, 82 pp.
NTIS: PB81-141038.

ERL PMEL-21 Halo, S. A., C. H. Pease, and R. W. Lindsay (1980)
Physical environment of the eastern Bering Sea, March 1979, 127 pp.
NTIS: PB81-148496.

ERL PMEL-22 Muench, R. D., and J. D. Schumacher (1980)
Physical oceanographic and meteorological conditions in the northwest Gulf of Alaska, 147 pp.
NTIS: PB81-199473.

ERL PMEL-23 Wright, Cathleen (1980)
Observations in the Alaskan Stream during 1980, 34 pp
NTIS: PB81-207441.

ERL PMEL-24 McNutt, L. (1980)
Ice conditions in the eastern Bering Sea from NOAA and LANDSAT imagery: Winter conditions 1974, 1976, 1977, 1979, 179 pp.
NTIS: PB81-220188.

ERL PMEL-25 Wright, C., and R. K. Reed (1980)
Comparison of ocean and inland rainfall in the tropical South Pacific, Atlantic, and Indian Oceans, 17 pp.
NTIS: PB81-225401.

ERL PMEL-26 Katz, C. N. and J. D. Cline (1980)
Processes affecting distribution of low-molecular-weight aliphatic hydrocarbons in Cook Inlet, Alaska, 84 pp.
NTIS: not yet available.

ERL PMEL-27 Feely, R. A., G. J. Massoth, A. J. Paulson (1981)
Distribution and elemental composition of suspended matter in Alaskan coastal waters, 119 pp.
NTIS: PB82-124538.

ERL PMEL-28 Muench, R. D., J. D. Schumacher, and C. A. Pearson (1980)
Circulation in the lower Cook Inlet, Alaska, 26 pp.
NTIS: PB82-126418.

ERL PMEL-29 Pearson, C. A. (1981)
Guide to R2D2--Rapid retrieval data display, 148 pp.
NTIS: PB82-150384.

ERL PMEL-30 Hamilton, S. E., and J. D. Cline (1981)
Hydrocarbons associated with suspended matter in the Green River, Washington, 116 pp.
NTIS: PB82-148677.

ERL PMEL-31 Reynolds, R. M., S. A. Macklin, and T. R. Heister (1981)
Observations of South Alaskan coastal winds, 49 pp.
NTIS: PB82-164823.

ERL PMEL-32 Pease, C. H., and S. A. Salo (1981)
Drift characteristics of northeastern Bering Sea ice during 1980, 79 pp.
NTIS: PB 83 112466

ERL PMEL-33 Ikeda, Motoyoshi (1982)
Eddies detached from a jet crossing over a submarine ridge: A study using a simple numerical model, 38 pp.
NTIS: PB82-217563.

ERL PMEL-34 Liu, Cho-Teng (1982)
Tropical Pacific sea surface temperature measured by SEASAT microwave radiometer and by ships, 160 pp.
NTIS: not yet available.

ERL PMEL-35 Lindsay, R.W., and A.L. Comiskey (1982):
Surface and upper-air observations in the eastern Bering Sea, 90 pp.
NTIS: not yet available.

ERL PMEL-36 Preisendorfer, R., and C. E. Mobley (1982)
Climate forecast verifications off the U. S. mainland, 1974-1982, 225 pp.
NTIS: not yet available.