

## Optimal Accelerometer Layouts for Data Recovery in Signature Verification

*Current experimental use of pen acceleration data for signature verification has prompted the mathematical theory of our recent paper on the subject, expounding motion recovery techniques for a special pen with imbedded accelerometers. This continuation seeks to optimize the instrument layout as a mechanical filter which serves to extract the kinematic observables from the experimental noise. Our early sections, through various algebraic postulates, determine a simplest function of the layout geometry expressing the relative desirability of an accelerometer configuration. Any nontrivial layout demands some off-axial instruments, but natural "feel" requires few such placements. Hence, our middle sections prove mathematical results under various geometric assumptions which yield optimal layouts with one, two, or three off-axial instruments. Our final sections discuss the further constraint that all accelerometers measure just three directions. They assume not more than two off-axial instruments and obtain the best configurations of these tridirectional types.*

### Introduction

The usual, intuitive methods for signature verification compare the visual characteristics of handwritten samples, but some modern approaches to the problem observe various physical quantities during the pen motion [1-5]. Indeed, Herbst and Morrissey [6] have patented verification techniques via pen acceleration data, while Herbst and Liu [7, 8] have achieved high selectivity through imbedded accelerometers. Also, Liu, Herbst, and Anthony [9] describe additional tests with still more accurate results. Our work provides a mathematical theory for kinematic measurements [10] and specifically relates pen motion to accelerometer readings. Its development includes possible shortcuts in trajectory computation [11] and even suggests instrument layouts for maximum sensitivity [12]. This continuation seeks to optimize the accelerometer layout as a mechanical filter which serves to recover the kinematic observables from the experimental noise. Some previous authors study particular instrument layouts [13, 14], but our present arguments furnish the *best* configurations. A recent note outlines our mathematical procedure [15], and our latest disclosure amplifies the design specifications [16]. However, no planned IBM product uses the resulting configurations.

This work contains two distinct parts, which reinforce one another. Sections 2 and 3 constitute the first part; they marshal several plausible arguments to choose a particular *design function* which ranks all accelerometer layouts. Sections 4 through 11 constitute the second part; they assume just six accelerometers and find all optimal layouts which have certain natural symmetries. Thus the earlier choice justifies the later optimization, while the outcome of the later special case shows the value of the earlier general theory. The second part treats the most interesting case, but other designs may need greater generality; so the first part considers  $n$  accelerometers. These two tasks require wholly different methods. Indeed, the first part invokes possibly unfamiliar results from invariant theory, algebraic geometry, and matrix theory, but it applies this information in a fairly direct way; the second part involves mostly standard results from convex analysis, linear algebra, and Euclidean geometry, but it exploits these theorems in a somewhat nontrivial way. Hence some repeatedly used facts from linear algebra are called "lemmas" for convenient reference, and the most important conclusions toward our pen designs are called "theorems" for the same reason.

Copyright 1980 by International Business Machines Corporation. Copying is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract may be used without further permission in computer-based and other information-service systems. Permission to *republish* other excerpts should be obtained from the Editor.

An accelerometer fixed in a moving rigid body measures the local acceleration projected on some constant internal direction [17, p. 17]. If a layout comprises  $n$  such instruments, then its geometry defines an  $n \times 6$  matrix  $\mathbf{A}$ . Thus Section 2 proposes a design function  $\delta(\mathbf{A})$  which expresses the desirability of a configuration, and invariant theory describes the most general such function which yields the same results under all relevant symmetries. Scale invariance, in Section 3, excludes two candidates for this function. Then a further plausible remark specifies the zero set of  $\delta(\mathbf{A})$ , and a powerful recent theorem [18] isolates a unique simplest choice with these zeros:

$$\delta(\mathbf{A}) = [\det(\mathbf{A}^T \mathbf{A})]^{1/2} \quad (\mathbf{T} = \text{transpose}). \quad (1)$$

But the layouts with worst possible  $\delta(\mathbf{A})$  include all those with no off-axial instruments. Therefore, our earlier disclosure [12] supplied an additional site for some accelerometers, namely, an attached ring circling the pen at its nonwriting end.

Convexity arguments, at the outset of Section 4, limit instrument locations to the pen point, the far end, and this ring. Moreover,  $\delta(\mathbf{A}) \equiv 0$  unless  $n \geq 6$ . Hence the second part assumes just six accelerometers and tries various fixed locations. Then the detailed analysis includes all reasonable placements and finds the optimal directions. To locate maximizing values for six unknown unit vectors demands some ingenuity beyond mere partial differentiation. But some instruments may have coincident locations, and the chosen  $\delta(\mathbf{A})$  is a multilinear function; while our "lemmas" will eliminate some unknowns either when several accelerometers have the same position inside the pen or when  $\delta(\mathbf{A})$  is a singular bilinear form in some two vectors. The "theorems" in Sections 5, 6, 7 treat all layouts with at most four distinct locations; their results, in Section 8, are the best configurations with at most three off-axial instruments, assuming that accelerometers on the ring take positions with equiangular spacing. The *tridirectional arrangements* of Section 9 restrict all measured components to three directions, the ordered multiplicities for these directions being either (3, 2, 1) or (2, 2, 2) for nontrivial layouts. Section 11, through further auxiliary theorems, finds the best tridirectional arrangements with one or two off-axial instruments. The (3, 2, 1) arrangements achieve the prior maxima; the (2, 2, 2) arrangements suffer the constant handicap

$$\max \delta(\mathbf{A}_{(2,2,2)}) / \max \delta(\mathbf{A}_{(3,2,1)}) = 3\sqrt{3}/8 = 0.64952. \quad (2)$$

This analysis retains our prior notation. Let  $\mathbf{R}$  be the real number field, with elements  $\rho, \sigma, \tau, \dots$ . Indeed, for the time variable  $\tau$ , let  $(\cdot)'$  indicate the derivative. Let  $\mathbf{E}^3$  be real Euclidean 3-space, with elements  $\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots$ . Moreover, for some orthonormal basis  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ , let

$$\mathbf{a} = (\alpha_1, \alpha_2, \alpha_3)^T, \quad \mathbf{b} = (\beta_1, \beta_2, \beta_3)^T, \quad \mathbf{c} = (\gamma_1, \gamma_2, \gamma_3)^T \quad (3)$$

represent the vector components. Here the superscript  $\mathbf{T}$  signifies the transpose, or real adjoint. Also  $\mathbf{a} \cdot \mathbf{b}$  denotes the inner product, or "dot product," of two vectors; while  $\mathbf{a} \times \mathbf{b}$  denotes the outer product, or "cross product," of these vectors [19, pp. 181, 227]. The inner product defines the standard Euclidean norm:  $|\mathbf{a}| = (\mathbf{a} \cdot \mathbf{a})^{1/2}$ . The standard product identities [10, Section 3], for any parallel  $\mathbf{a}, \mathbf{b}$ , yield the following useful relation among otherwise arbitrary vectors:

$$\det(\mathbf{a} \times \mathbf{r}, \mathbf{b} \times \mathbf{s}, \mathbf{c} \times \mathbf{t}) = \det(\mathbf{a}, \mathbf{r}, \mathbf{s}) \det(\mathbf{b}, \mathbf{c}, \mathbf{t}). \quad (4)$$

For a moving pen, conceived as a general rigid body, we refer positions to an internal coordinate system, fixed in the body itself, and we relate this to an external coordinate system, fixed in the surrounding environment [20, Chapter 4]. Specifically, we introduce a right-handed orthonormal basis  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  for the internal coordinates, where we align the vector  $\mathbf{e}_3$  up the pen axis. Moreover, we locate the coordinate origin at the pen point, and we assign the vector displacement  $\mathbf{y}$  to an arbitrary internal position. If  $\mathbf{p}$  is the external vector displacement of the pen point and  $\mathbf{q}$  is the external vector displacement of the point  $\mathbf{y}$ , then the basic relation

$$\mathbf{q}(\tau) = \mathbf{p}(\tau) + \mathbf{U}(\tau)\mathbf{y} \quad (5)$$

connects the internal with the external description, where the orthogonal transformation  $\mathbf{U}(\tau)$  carries internal to external vectors. This orthogonal matrix  $\mathbf{U}(\tau)$  defines a skew-symmetric matrix  $\mathbf{W}(\tau)$  by

$$\mathbf{U}' = \mathbf{U}\mathbf{W} \text{ for all times } \tau. \quad (6)$$

The resulting matrix  $\mathbf{W}(\tau)$  defines an angular velocity  $\mathbf{w}(\tau)$  by

$$\mathbf{W}\mathbf{c} = \mathbf{w} \times \mathbf{c} \text{ for all vectors } \mathbf{c}. \quad (7)$$

## 2. Invariance properties of design function

Here we suppose a *design function* for instrument layouts, and we describe its invariance properties under symmetry transformations. Let a pen contain fixed linear accelerometers at internal positions  $\mathbf{y}_1, \dots, \mathbf{y}_n$ . Let these instruments measure respective acceleration components along unit vectors  $\mathbf{u}_1, \dots, \mathbf{u}_n$ . If the vector  $\Psi = (\psi_1, \dots, \psi_n)^T$  represents the corresponding measurements  $\psi_i(\tau)$  and the vector  $\Phi = (\phi_1, \dots, \phi_n)^T$  concatenates the quadratic forms  $\phi_i(\mathbf{w}) = (\mathbf{y}_i \times \mathbf{w}) \cdot (\mathbf{u}_i \times \mathbf{w})$ , then the kinematic variables  $\mathbf{p}, \mathbf{p}', \mathbf{U}, \mathbf{w}$  obey the basic system [10]

$$\mathbf{A}\mathbf{V}^T \begin{pmatrix} \mathbf{p}'' + \mathbf{g} \\ \mathbf{w}' \end{pmatrix} = \Phi(\mathbf{w}) + \Psi(\tau), \quad (8)$$

where the  $n \times 6$  matrix  $\mathbf{A}$  and the  $6 \times 6$  matrix  $\mathbf{V}$  have the particular forms

$$\mathbf{A}^T = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \\ \mathbf{y}_1 \times \mathbf{u}_1 & \cdots & \mathbf{y}_n \times \mathbf{u}_n \end{pmatrix}, \quad \mathbf{V}(\tau) = \begin{pmatrix} \mathbf{U}(\tau) & 0 \\ 0 & \mathbf{I} \end{pmatrix} \quad (9)$$

Here the constant vector  $\mathbf{g}$  denotes the external gravitational field. Also, a unique solution for the trajectory  $\mathbf{p}(\tau)$  requires the full rank of the matrix  $\mathbf{A}$ :

$$n \geq \text{rank}(\mathbf{A}) = 6; \quad (10)$$

and a least-squares combination of measurements provides a system of normal equations:

$$\mathbf{A}^T \mathbf{A} \mathbf{V}^T \begin{pmatrix} \mathbf{p}'' \\ \mathbf{w}' \end{pmatrix} + \mathbf{g} = \mathbf{A}^T \Phi(\mathbf{w}) + \mathbf{A}^T \Psi(\tau). \quad (11)$$

Clearly, the matrix  $\mathbf{A}$  must contain the essential geometry of the layout, and the ideal geometry should offer the most accurate solution of (11). Therefore, the matrix  $\mathbf{A}$  should produce our criterion for an optimal layout, and this unknown criterion must refine the nondegeneracy assertion of (10). However, (10), by standard theorems [19, pp. 78-80], requires simply the existence of  $(\mathbf{A}^T \mathbf{A})^{-1}$ . Also a measurement error  $\Delta \Psi$ , by (11), produces a change

$$\begin{pmatrix} \Delta \mathbf{p}'' \\ \Delta \mathbf{w}' \end{pmatrix} = \mathbf{V}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \Delta \Psi \quad (12)$$

in  $\mathbf{p}''$  and  $\mathbf{w}'$ , while the orthogonal matrix  $\mathbf{V}$  of (9) preserves the Euclidean norms of  $\Delta \mathbf{p}''$  and  $\Delta \mathbf{w}'$ . Thus a desirable instrument configuration should yield a "small" generalized inverse  $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ . To compare accelerometer layouts for different matrices  $\mathbf{A}$ , we express their desirability via a real-valued *design function*  $\delta(\mathbf{A})$ . We cannot rigorously isolate a unique  $\delta(\mathbf{A})$ , but we shall so delimit the possibilities that we can finally distinguish a unique simplest choice. Here we broaden the domain of the matrix  $\mathbf{A}$ , and we deduce some restrictions from various symmetries. Specifically, we permit accelerometers with different sensitivities, yielding measurements with different weights. Accordingly, we allow matrices  $\mathbf{A}$  as arguments having vectors  $\mathbf{u}_i$  of arbitrary magnitude.

Sensitivity and distance admit independent units and measurements. Thus a change of sensitivity scale yields a transformation  $\mathbf{y}_i \rightarrow \mathbf{y}_i$ ,  $\mathbf{u}_i \rightarrow \alpha \mathbf{u}_i$ , for all  $i$ , which implies a mapping

$$\mathbf{A} \rightarrow \alpha \mathbf{A} \text{ for some positive } \alpha. \quad (13)$$

Also a change of distance scale yields a transformation  $\mathbf{y}_i \rightarrow \alpha \mathbf{y}_i$ ,  $\mathbf{u}_i \rightarrow \mathbf{u}_i$ , for all  $i$ , which implies a mapping

$$\mathbf{A} \rightarrow \mathbf{A} \cdot \text{diag}(1, 1, 1, \alpha, \alpha, \alpha) \text{ for some positive } \alpha. \quad (14)$$

However, these scalings should not invert the relative desirability of any two layouts, whence each mapping

should simply multiply  $\delta(\mathbf{A})$  by a corresponding power  $\alpha^k$ . The design function  $\delta(\mathbf{A})$ , in other words, should be a homogeneous function of either the  $\mathbf{u}_i$  or the  $\mathbf{y}_i$ . If we translate the internal origin, then we do not alter the directions  $\mathbf{u}_i$ . Still, we take all  $\mathbf{y}_i$  into  $\mathbf{y}_i + \mathbf{a}$ , for an arbitrary 3-vector  $\mathbf{a}$ , and we map

$$\mathbf{A} \rightarrow \mathbf{A} \begin{pmatrix} \mathbf{I} & \mathbf{S}(\mathbf{a}) \\ 0 & \mathbf{I} \end{pmatrix} \quad (15)$$

for a corresponding  $3 \times 3$  skew matrix  $\mathbf{S}(\mathbf{a})$ . Specifically, the origin may thus become any  $\mathbf{y}_i$ ; moreover, these translations should not affect  $\delta(\mathbf{A})$ ; hence, the design function can involve only differences  $\mathbf{y}_i - \mathbf{y}_j$ . Alternatively, the origin for the internal coordinates may be, more symmetrically, the average  $(\mathbf{y}_1 + \cdots + \mathbf{y}_n)/n$  of the displacements  $\mathbf{y}_i$ . Our further remarks assume this last normalization.

If we renumber the  $n$  accelerometers in a different order, then we map

$$\mathbf{A} \rightarrow (n \times n \text{ permutation matrix}) \cdot \mathbf{A}. \quad (16)$$

If we reverse the direction  $\mathbf{u}_j$  of the  $j$ th instrument, then we map

$$\mathbf{A} \rightarrow \text{diag}(1, \cdots, 1, -1(j\text{th position}), 1, \cdots, 1) \cdot \mathbf{A}. \quad (17)$$

Clearly this yields a layout with identical capabilities, because it changes only the sign of the  $j$ th measurement. If we rotate the basis  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  via a proper orthogonal  $\mathbf{U}$ , then we multiply all  $\mathbf{y}_i$  and  $\mathbf{u}_i$  by this same transformation  $\mathbf{U}$ . Moreover this takes  $\mathbf{y}_i \times \mathbf{u}_i$  into  $\mathbf{U}(\mathbf{y}_i \times \mathbf{u}_i)$ , whence it maps

$$\mathbf{A} \rightarrow \mathbf{A} \begin{pmatrix} \mathbf{U} & 0 \\ 0 & \mathbf{U} \end{pmatrix}^T. \quad (18)$$

If we reverse all  $\mathbf{u}_i$ , via (17), then  $\mathbf{A}$  becomes  $-\mathbf{A}$ ; if we reverse all  $\mathbf{y}_i$  and  $\mathbf{u}_i$ , then  $\mathbf{A}$  becomes  $\mathbf{A} \cdot \text{diag}(-1, -1, -1, 1, 1, 1)$ . However, the replacement (18), together with the last two possibilities, generates all mappings

$$\mathbf{A} \rightarrow \mathbf{A} \begin{pmatrix} \pm \mathbf{U} & 0 \\ 0 & \pm \mathbf{U} \end{pmatrix} \quad (19)$$

with two independent  $\pm$  signs. All these transformations preserve the values  $\delta(\mathbf{A})$ .

The symmetries of the last paragraph form a compact group of transformations. Thus a sufficient goal for our investigation is a complete set of *polynomial* invariants. Indeed, any invariant function  $\delta(\mathbf{A})$  under the mappings (16)-(19) is simply an arbitrary function of the polynomial invariants  $\delta_1(\mathbf{A}), \cdots, \delta_k(\mathbf{A})$  [21, 22]. Moreover, the columns of  $\mathbf{A}^T$  contain the vectors  $\mathbf{u}_i$ ,  $\mathbf{y}_i \times \mathbf{u}_i$  for all instruments, and the action of (19) produces all images  $\mathbf{U}\mathbf{u}_i$ ,

$U(\mathbf{y}_i \times \mathbf{u}_i)$  under the *full* orthogonal group. The invariant combinations of these images involve only generic products of the forms [23, pp. 52-56]

$$(\mathbf{u}_i \times \mathbf{u}_j), (\mathbf{y}_i \times \mathbf{u}_j) \cdot \mathbf{u}_j, (\mathbf{y}_i \times \mathbf{u}_j) \cdot (\mathbf{y}_j \times \mathbf{u}_j). \quad (20)$$

Also, the independent  $\pm$  signs on the expression (19) require even numbers of the factors  $(\mathbf{y}_i \times \mathbf{u}_j) \cdot \mathbf{u}_j$ ; and our discussion of (13) and (14) imposes homogeneity in the  $\mathbf{y}_i$  and the  $\mathbf{u}_i$ . Indeed, reversals of individual  $\mathbf{u}_i$  imply even numbers of each single  $\mathbf{u}_i$ , and relabeling of accelerometers demands symmetry in all indices. Collectively, these symmetries restrict the form of any function  $\delta(\mathbf{A})$ , though our remarks permit the construction of many invariant polynomials.

### 3. Systematic choice of design function

Here we impose some plausible requirements beyond the stated invariance properties, and we discern a "simplest" design function under these combined hypotheses. Thus, the unknown  $\delta(\mathbf{A})$  is specifically a real-valued function, and we can choose its sign so that  $\delta(\mathbf{A})$  becomes larger as  $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$  becomes "smaller." Also, a noninvertible  $\mathbf{A}^T \mathbf{A}$  is clearly worse, and we can add a constant so that

$$\delta(\mathbf{A}) = 0 \text{ when rank } (\mathbf{A}) < 6. \quad (21)$$

Indeed, any nonsingular  $\mathbf{A}^T \mathbf{A}$  is presumably better, whence we may sharpen (21) so that

$$\delta(\mathbf{A}) > 0 \text{ when rank } (\mathbf{A}) = 6. \quad (22)$$

Now, linear algebra suggests three candidates for the design function, where the preceding remarks guide our normalization in each case. One measure of a "small" inverse  $(\mathbf{A}^T \mathbf{A})^{-1}$  is the denominator in Cramer's formula [19, p. 112]:

$$\delta_0(\mathbf{A}) = \det(\mathbf{A}^T \mathbf{A}). \quad (23)$$

This determinant, of course, is the product of the eigenvalues. Another possibility is

$$\delta_1(\mathbf{A}) = \text{smallest eigenvalue of } \mathbf{A}^T \mathbf{A}, \quad (24)$$

since its reciprocal is the subordinate  $\ell^2$  norm of  $(\mathbf{A}^T \mathbf{A})^{-1}$ , and this norm bounds the  $\ell^2$  error amplification in (12) [19, p. 201; 24, p. 56]. A third alternative is

$$\delta_2(\mathbf{A}) = \text{smallest eigenvalue/largest eigenvalue}, \quad (25)$$

since its reciprocal is the *condition number* of  $\mathbf{A}^T \mathbf{A}$ , and this quantity reflects the computational difficulty of inversion [24, p. 89].

Clearly, all candidates are unitary invariants of  $\mathbf{A}^T \mathbf{A}$ , whence these functions have the required behavior under (13) and (16)-(19). However, another symmetry excludes

the last two alternatives. The matrix  $\mathbf{A}$ , under the mapping (14), takes the form  $\mathbf{A}(\alpha) = \mathbf{A}\mathbf{Q}(\alpha)$ , with

$$\mathbf{Q}(\alpha) = \text{diag}(1, 1, 1, \alpha, \alpha, \alpha). \quad (26)$$

Hence, the largest eigenvalue of  $\mathbf{A}(\alpha)^T \mathbf{A}(\alpha) = \mathbf{Q}(\alpha)^T \mathbf{A}^T \mathbf{A} \mathbf{Q}(\alpha)$  is approximately constant for small  $\alpha$ , and asymptotically  $(\text{const.})\alpha^2$  for large  $\alpha$  [24, Chapter 2]. The inverse matrix, with the reciprocal eigenvalues, has the form  $[\mathbf{A}(\alpha)^T \mathbf{A}(\alpha)]^{-1} = \mathbf{Q}(\alpha^{-1})[\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{Q}(\alpha^{-1})$  by direct computation. Thus the smallest eigenvalue of  $\mathbf{A}(\alpha)^T \mathbf{A}(\alpha)$  is approximately constant for large  $\alpha$  and asymptotically  $(\text{const.})\alpha^2$  for small  $\alpha$ . Therefore, the functions  $\delta_1(\mathbf{A})$  and  $\delta_2(\mathbf{A})$  lack the required homogeneity in  $\alpha$ , whereas the expressions (14) and (15) yield the desired behavior for  $\delta_0(\mathbf{A})$ . Again, these arguments permit other, less plausible, alternatives, but the following results support further the special choice  $\delta_0(\mathbf{A})$ .

#### • Lemma 1

Let  $\mathbf{A}$  be a square matrix; that is,  $n = 6$ . If  $\mathbf{u}_1, \dots, \mathbf{u}_6, \mathbf{y}_1, \dots, \mathbf{y}_6$  have independent complex variables as components, then  $\det(\mathbf{A})$  is an irreducible polynomial over the complex numbers.

#### Proof

The polynomial  $\det(\mathbf{A})$ , by definition (9), is a linear function of each  $(\mathbf{u}_i, \mathbf{y}_i \times \mathbf{u}_i)$ . Hence, if this polynomial has nontrivial factorization  $\pi_1(\mathbf{A}) \cdots \pi_k(\mathbf{A})$ , then necessarily only one factor can involve each pair  $(\mathbf{u}_i, \mathbf{y}_i \times \mathbf{u}_i)$ . However, all permutations of the indices  $\{1, \dots, 6\}$  preserve  $\det(\mathbf{A})$  up to a sign, so that all relabelings of the vectors  $\mathbf{u}_i, \mathbf{y}_i$  express  $\det(\mathbf{A})$  as a nontrivial product. Moreover any complex polynomial must have *unique* irreducible factors [25, pp. 70-74]; therefore, the polynomial  $\det(\mathbf{A})$  can have no *nonlinear* irreducible factors. Thus, symmetry under permutations requires

$$\det(\mathbf{A}) = \text{constant} \cdot \prod_{i=1}^6 [\mathbf{a} \cdot \mathbf{u}_i + \mathbf{b} \cdot \mathbf{y}_i \times \mathbf{u}_i] \quad (27)$$

for some complex vectors  $\mathbf{a}, \mathbf{b}$ ; and sign reversal of all  $\mathbf{y}_i$  produces a similar formula with corresponding factors  $[\mathbf{a} \cdot \mathbf{u}_i - \mathbf{b} \cdot \mathbf{y}_i \times \mathbf{u}_i]$ , whence unique factorization of  $\det(\mathbf{A})$  implies

$$\mathbf{a} \cdot \mathbf{u}_i - \mathbf{b} \cdot \mathbf{y}_i \times \mathbf{u}_i = \gamma_i [\mathbf{a} \cdot \mathbf{u}_i + \mathbf{b} \cdot \mathbf{y}_i \times \mathbf{u}_i] \quad (28)$$

for some complex constants  $\gamma_i$ . Putting  $\mathbf{y}_i = 0$ , we obtain  $\gamma_i = 1$ ; letting  $\mathbf{y}_i$  be arbitrary, we find  $\mathbf{b} = 0$ . Hence the  $\mathbf{y}_i$  do not affect (27), and this independence provides a contradiction.  $\square$

#### • Theorem 1

Let  $n = 6$  for the matrix  $\mathbf{A}$  and  $\delta(\mathbf{A})$  satisfy (21)-(22) for real arguments. If  $\delta(\mathbf{A})$  is an analytic function of  $\mathbf{u}_1, \dots,$

$u_6, y_1, \dots, y_6$  for all  $u_i, y_i$  sufficiently near zero, then  $\delta(A)$  is the product of some nonnegative analytic function with  $\det(A^T A) = [\det(A)]^2$ .

#### Proof

The function  $\det(A)$ , arbitrarily near  $A = 0$ , takes both positive and negative values for real  $u_i$  and  $y_i$ . Moreover this function, by Lemma 1, can have no nontrivial analytic factors over the complex numbers; otherwise the leading terms of their power series would yield a nontrivial polynomial factorization of  $\det(A)$ . If  $\det(A) = 0$  for any real arguments, then  $\delta(A) = 0$  by assumption (21). Therefore,  $\det(A)$ , by a suitable Nullstellensatz [18, Proposition 4.2], divides the function  $\delta(A)$ ; while  $\delta(A)/\det(A)$ , by assumption (22), takes positive or negative values according as  $\det(A)$ , at the same point, takes positive or negative values. But the zero set of  $\det(A)$  includes no open subset of matrices. Hence  $\delta(A)/\det(A) = 0$  when  $\det(A) = 0$ . Thus  $\delta(A)/\det(A)$ , by the same Nullstellensatz, again has  $\det(A)$  as a factor.  $\square$

Clearly Theorem 1, for six accelerometers, offers a new fact supporting  $\delta_0(A)$ ; namely, this polynomial, even without symmetry restrictions, provides the "simplest" function satisfying (21)–(22). Also, Theorem 1 implies a weaker result when  $n > 6$ . Given any subset  $J = \{j_1, \dots, j_6\}$  of  $\{1, \dots, n\}$ , define the submatrix  $A_J$  of  $A^T$  by taking the six columns with indices  $j_1, \dots, j_6$ . If some  $\delta(A)$ , for any admissible  $n$ , is a polynomial of least total degree, then each  $A^T$  with only six nonzero columns yields an expression constant  $\cdot [\det(A_J)]^2$  for some set  $J$ . Thus  $\delta(A) = \text{constant} \cdot \sum_J [\det(A_J)]^2 + (\text{other terms})$  by permutation symmetry, while

$$\delta_0(A) = \det(A^T A) = \sum_J [\det(A_J)]^2 \quad (29)$$

by the Binet-Cauchy formula [26, p. 9]. Hence the "simplest" possibility is again  $\delta_0(A)$ , though a higher-degree alternative is, similarly,  $\delta_3(A) = \sum_J [\det(A_J)]^4$ . Here, despite this nonuniqueness, we prefer  $\delta_0(A)$  and, for analytical convenience, we define

$$\delta(A) = \delta_0(A)^{1/2} = [\det(A^T A)]^{1/2}. \quad (30)$$

#### • Theorem 2

Any one of the following implies the vanishing of  $\delta(A)$ : (1) the  $y_i$  describe collinear positions; (2) the  $u_i$  span a two-dimensional subspace; (3) some  $n - 2$  of the  $y_i$  are equal; (4) some  $n - 2$  of the  $u_i$  are parallel.

#### Proof

Shifting the origin in some cases, we find  $\text{rank}(A) < 6$  by inspection.  $\square$

#### 4. Auxiliary results for eventual optimization

Here we describe the variable domains, to specify our optimization problem; and we apply some convexity theo-

rems to circumscribe the admissible locations. Then we restrict further study to six-instrument layouts, and we collect some auxiliary facts for convenient reference. The chosen  $\delta(A)$ , in particular, is a linear function of each  $|u_i|$ . Hence we may specialize our problem to its original form, and we shall consider only vectors  $u_i$  of unit length.

Also, nontrivial maxima, by Theorem 2, Part 1, demand some accelerometers off the pen axis, whereas outlying instruments near the pen point impede the natural flow of handwriting motion. Thus we have proposed encircling the pen axis at its far end by a ring, and we have envisioned attaching this ring to the shaft by fins or spokes [12]. But the domain  $D$  for any vector  $y_i$  is the space available for instruments; and the set  $D$ , after this enlargement, is the pen volume together with the ring. If  $\text{co}(D)$  denotes the closed convex hull of  $D$  and  $\text{ex}(D)$  comprises the extreme points of  $\text{co}(D)$ , then  $\text{co}(D)$ , for a finite pen, is clearly a compact set in  $E^3$ , and  $\text{ex}(D)$ , by a standard result [27, p. 165], is already a subset of  $D$ . Indeed  $\text{ex}(D)$ , by the Krein-Milman theorem [27, p. 166], has  $\text{co}(D)$  as its closed convex hull. However  $\delta_0(A)$ , by (29), is a convex function of each  $y_i$ , because  $[\det(A_J)]^2$ , by inspection, is a convex function for each  $J$ . Hence our  $\delta(A)^2$ , for each  $y_i$ , will assume its maximum on  $\text{ex}(D)$ , though nonextremal points may also furnish maxima [27, p. 343]; and our search for optimal  $y_i$  may suppose their locations in  $\text{ex}(D)$ . This preferred subset of the total domain includes the ring together with the pen point; a later limitation on off-axial instruments supplements these two possibilities with the far end. Thus we may try successively these few locations for the  $y_i$ , and we need seek only the best directions for the  $u_i$ .

Unfortunately, the last problem, for  $n$  accelerometers, probably exceeds the resources of purely analytical methods. However, explicit results in a special case nicely display the potential of our theory. Moreover, the limited volume inside a pen clearly justifies the fewest accelerometers as the most interesting case. Hence all further discussion assumes  $n = 6$ , and various detailed arguments, each labeled a "theorem," yield optimal layouts. Now  $A$  is a square matrix and

$$\delta(A) = |\det(A)|. \quad (31)$$

The following lemmas include some well-known facts, but their formal statement permits convenient later references. If  $\omega(v)$  represents any real-valued function of the vector  $v$ , then  $\hat{v}$  denotes a generic point with maximal  $\omega(v)$ .

#### • Lemma 2

If  $a$  is an arbitrary 3-vector and  $\omega(v) = |a \cdot v|$ , then

$$\{a \cdot v : |v| = 1\} = [-|a|, +|a|] \quad (32)$$

and  $\max \{\omega(\mathbf{v}): |\mathbf{v}| = 1\} = |\mathbf{a}|$ . If  $\mathbf{a} = 0$ , then  $\hat{\mathbf{v}}$  is any unit vector. If  $\mathbf{a} \neq 0$ , then  $\hat{\mathbf{v}} = \pm \mathbf{a}/|\mathbf{a}|$ .

*Proof*

Obvious.  $\square$

• **Lemma 3**

If  $\mathbf{y}_j = \mathbf{y}_k$  for any distinct  $j, k$ , then  $\delta(\mathbf{A})$  involves  $\mathbf{u}_j, \mathbf{u}_k$  only through  $\mathbf{u}_j \times \mathbf{u}_k$ . Hence the vectors  $\hat{\mathbf{u}}_j, \hat{\mathbf{u}}_k$ , for a nonconstant function  $\delta(\mathbf{A})$ , are any orthogonal unit vectors with the correct normal.

*Proof*

We may take  $j = 5, k = 6$  by an index permutation. We may put  $\mathbf{y}_5 = \mathbf{y}_6 = 0$  by translation-invariance. The Laplace expansion of a determinant implies

$$\begin{aligned} \det(\mathbf{A}) &= \det(\mathbf{u}_5, \mathbf{u}_6, \mathbf{u}_1) \det(\mathbf{y}_2 \times \mathbf{u}_2, \mathbf{y}_3 \times \mathbf{u}_3, \mathbf{y}_4 \times \mathbf{u}_4) \\ &\quad - \det(\mathbf{u}_5, \mathbf{u}_6, \mathbf{u}_2) \det(\mathbf{y}_1 \times \mathbf{u}_1, \mathbf{y}_3 \times \mathbf{u}_3, \mathbf{y}_4 \times \mathbf{u}_4) \\ &\quad + \det(\mathbf{u}_5, \mathbf{u}_6, \mathbf{u}_3) \det(\mathbf{y}_1 \times \mathbf{u}_1, \mathbf{y}_2 \times \mathbf{u}_2, \mathbf{y}_4 \times \mathbf{u}_4) \\ &\quad - \det(\mathbf{u}_5, \mathbf{u}_6, \mathbf{u}_4) \det(\mathbf{y}_1 \times \mathbf{u}_1, \mathbf{y}_2 \times \mathbf{u}_2, \mathbf{y}_3 \times \mathbf{u}_3) \\ &= (\mathbf{u}_5 \times \mathbf{u}_6) \cdot \mathbf{f}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4) \end{aligned} \quad (33)$$

for a vector  $\mathbf{f}$  [19, p. 110]. Thus  $\hat{\mathbf{u}}_5 \times \hat{\mathbf{u}}_6$ , by Lemma 2, must parallel  $\mathbf{f}(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3, \hat{\mathbf{u}}_4)$  for nontrivial  $\delta(\mathbf{A})$ ; and  $|\hat{\mathbf{u}}_5 \times \hat{\mathbf{u}}_6|$ , by inspection, must equal unity for maximal  $|\det(\mathbf{A})|$ .  $\square$

• **Lemma 4**

If  $\mathbf{y}_i = \mathbf{y}_j = \mathbf{y}_k$  for any distinct  $i, j, k$ , then  $\delta(\mathbf{A})$  involves  $\mathbf{u}_i, \mathbf{u}_j, \mathbf{u}_k$  only through  $\det(\mathbf{u}_i, \mathbf{u}_j, \mathbf{u}_k)$ . Hence the vectors  $\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \hat{\mathbf{u}}_k$  for a nonconstant  $\delta(\mathbf{A})$  are any orthonormal basis for the space  $E^3$ .

*Proof*

We may take  $i = 4, j = 5, k = 6$ . We may put  $\mathbf{y}_4 = \mathbf{y}_5 = \mathbf{y}_6 = 0$ . The Laplace expansion [19, p. 110] implies

$$\begin{aligned} \det(\mathbf{A}) &= -\det(\mathbf{y}_1 \times \mathbf{u}_1, \mathbf{y}_2 \times \mathbf{u}_2, \mathbf{y}_3 \times \mathbf{u}_3) \det(\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6). \end{aligned} \quad (34)$$

However  $|\det(\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6)|$  takes its maximum value when  $(\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6)$  is any orthonormal basis.  $\square$

• **Lemma 5**

If  $\mathbf{B}$  is any real  $3 \times 3$  matrix and  $\mathbf{v}_1, \mathbf{v}_2$  are any real 3-vectors, then  $\max \{(\mathbf{v}_1^T \mathbf{B} \mathbf{v}_2)^2: |\mathbf{v}_1| = |\mathbf{v}_2| = 1\}$  is the largest eigenvalue of  $\mathbf{B}^T \mathbf{B}$ , and this eigenvalue is the largest root of  $\det(\lambda - \mathbf{B}^T \mathbf{B})$ . If  $\text{rank}(\mathbf{B}) < 3$ , then  $\lambda$  divides  $\det(\lambda - \mathbf{B}^T \mathbf{B})$ , and the maximum of this lemma is the larger root of the quotient.

*Proof*

Lemma 2 yields these known facts; indeed

$$\max \{(\mathbf{v}_1^T \mathbf{B} \mathbf{v}_2)^2: |\mathbf{v}_1| = 1\} = |\mathbf{B} \mathbf{v}_2|^2 = \mathbf{v}_2^T \mathbf{B}^T \mathbf{B} \mathbf{v}_2. \quad \square \quad (35)$$

The subspace  $\langle \mathbf{a}_1, \dots, \mathbf{a}_m \rangle$ , for any vectors  $\mathbf{a}_i$ , will be the subset  $\{\sum_i \rho_i \mathbf{a}_i: \rho_i \in \mathbb{R}\}$  of the space  $E^3$ . If  $\mathbf{a}, \mathbf{b}$  are any nonzero vectors in  $E^3$ , then  $\langle \mathbf{a} \rangle, \langle \mathbf{b} \rangle$  are undirected straight lines through the origin. Also if  $\mathbf{a}_0 + \mathbf{b}_0 \neq 0$ , where

$$\mathbf{a}_0 = \mathbf{a}/|\mathbf{a}|, \quad \mathbf{b}_0 = \mathbf{b}/|\mathbf{b}|, \quad (36)$$

then clearly  $\langle \mathbf{a}_0 + \mathbf{b}_0 \rangle$  bisects  $\angle(\mathbf{a}, \mathbf{b})$ . Therefore, the bisectors of the pair  $\{\mathbf{a}, \mathbf{b}\}$  will be the unit vectors  $\pm \mathbf{e}$  in the line  $\langle \mathbf{a}_0 + \mathbf{b}_0 \rangle$ . However, the intersection of the lines  $\langle \mathbf{a} \rangle, \langle \mathbf{b} \rangle$  yields two pairs of opposite angles. Accordingly, the principal bisectors for this pair of undirected lines will be the unit vectors along all bisectors of nonobtuse angles: either two pairs of opposite vectors for perpendicular lines or one pair of opposite vectors in other cases, hence one pair, in the obvious directions, for coincident lines. Thus, if  $\text{sgn}(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} = 0$ , then all four bisectors of  $\{\mathbf{a}, \pm \mathbf{b}\}$  are principal bisectors of  $\{\langle \mathbf{a} \rangle, \langle \mathbf{b} \rangle\}$ , while, if  $\text{sgn}(\mathbf{a} \cdot \mathbf{b}) = \varepsilon \neq 0$ , then the two bisectors of  $\{\mathbf{a}, \varepsilon \mathbf{b}\}$  are the principal bisectors of  $\{\langle \mathbf{a} \rangle, \langle \mathbf{b} \rangle\}$ .

• **Lemma 6**

If  $\mathbf{a}, \mathbf{b}$  are arbitrary 3-vectors and  $\omega(\mathbf{v}) = |(\mathbf{a} \cdot \mathbf{v})(\mathbf{b} \cdot \mathbf{v})|$ , then

$$\begin{aligned} \{2(\mathbf{a} \cdot \mathbf{v})(\mathbf{b} \cdot \mathbf{v}): |\mathbf{v}| = 1\} \\ = [\mathbf{a} \cdot \mathbf{b} - |\mathbf{a}| \cdot |\mathbf{b}|, \mathbf{a} \cdot \mathbf{b} + |\mathbf{a}| \cdot |\mathbf{b}|] \end{aligned} \quad (37)$$

and  $\max \{\omega(\mathbf{v}): |\mathbf{v}| = 1\} = (|\mathbf{a} \cdot \mathbf{b}| + |\mathbf{a}| \cdot |\mathbf{b}|)/2$ . If either  $\mathbf{a}$  or  $\mathbf{b}$  is zero, then  $\hat{\mathbf{v}}$  is any unit vector. If both  $\mathbf{a}$  and  $\mathbf{b}$  are nonzero, then  $\hat{\mathbf{v}}$  is any principal bisector of  $\langle \mathbf{a} \rangle, \langle \mathbf{b} \rangle$ .

*Proof*

If  $\mathbf{a}, \mathbf{b}$  are linearly dependent, then these results are obvious. If  $\mathbf{a}, \mathbf{b}$  are any independent vectors and  $\mathbf{a}_0, \mathbf{b}_0$  are the unit vectors (36), then  $(\mathbf{a} \cdot \mathbf{v})(\mathbf{b} \cdot \mathbf{v}) = \mathbf{v}^T \mathbf{C} \mathbf{v}$ , where

$$2\mathbf{C} = \mathbf{a}\mathbf{b}^T + \mathbf{b}\mathbf{a}^T = |\mathbf{a}| |\mathbf{b}| \{\mathbf{a}_0 \mathbf{b}_0^T + \mathbf{b}_0 \mathbf{a}_0^T\}. \quad (38)$$

The matrix  $\mathbf{C}$  has eigenvectors  $\mathbf{a}_0 \times \mathbf{b}_0, \mathbf{a}_0 \pm \mathbf{b}_0$ ; an extremizing  $\mathbf{v}$  must parallel some eigenvector.  $\square$

**5. Accelerometer layouts with three positions**

The assumption  $n = 6$  yields the simplification  $\delta(\mathbf{A}) = |\det(\mathbf{A})|$ . Now, further plausible constraints will provide analytical solutions, though unrestricted maxima may demand numerical techniques. Indeed, Section 4 so limits admissible locations that several  $\hat{\mathbf{y}}_i$  probably take coincident values. The unknown  $\hat{\mathbf{y}}_i$ , by Theorem 2, Part 1, must include three noncollinear points; no four  $\hat{\mathbf{y}}_i$ , by Theorem 2, Part 3, may occupy the same point. Hence any optimal six instruments admit the following location multiplicities:

[1, 2, 3], [2, 2, 2], [1, 1, 1, 3],

[1, 1, 2, 2], [1, 1, 1, 1, 2], [1, 1, 1, 1, 1, 1].

Here, Theorem 3 covers all [1, 2, 3] configurations, and Theorem 4 treats all [2, 2, 2] configurations.

#### • Theorem 3

For a six-instrument layout, assume  $\mathbf{y}_4 = \mathbf{y}_5 = \mathbf{y}_6$ , and, by translation-invariance, let this common point be the origin. In these coordinates, suppose  $\mathbf{y}_2 \times \mathbf{y}_3 = 0$ , but, in other respects, let  $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$  be arbitrary vectors. Then

$$\max \{\delta(\mathbf{A}) : \text{all } |\mathbf{u}_i| = 1\} = |\mathbf{y}_1 \times \mathbf{y}_2| \cdot |\mathbf{y}_3|. \quad (39)$$

If either  $\mathbf{y}_1 \times \mathbf{y}_2 = 0$  or  $\mathbf{y}_3 = 0$ , then the  $\hat{\mathbf{u}}_i$  are any unit vectors. Otherwise  $\hat{\mathbf{u}}_1 = \pm \mathbf{y}_1 \times \mathbf{y}_2 / |\mathbf{y}_1 \times \mathbf{y}_2|$ , and  $\hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3$  are any orthogonal unit vectors such that  $\hat{\mathbf{u}}_2 \times \hat{\mathbf{u}}_3 = \pm \mathbf{y}_3 / |\mathbf{y}_3|$ ; while  $(\hat{\mathbf{u}}_4, \hat{\mathbf{u}}_5, \hat{\mathbf{u}}_6)$  is any orthonormal basis.

#### Proof

Relations (34) and (4) imply

$$\det(\mathbf{A}) = (\mathbf{u}_1 \cdot \mathbf{y}_1 \times \mathbf{y}_2)(\mathbf{u}_2 \times \mathbf{u}_3 \cdot \mathbf{y}_3)(\mathbf{u}_4 \cdot \mathbf{u}_5 \times \mathbf{u}_6). \quad (40)$$

Hence the results are trivial when either  $\mathbf{y}_1 \times \mathbf{y}_2$  or  $\mathbf{y}_3$  is zero. Otherwise, Lemma 2 yields the stated  $\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3$ ; and Lemma 4 yields the stated  $\hat{\mathbf{u}}_4, \hat{\mathbf{u}}_5, \hat{\mathbf{u}}_6$ . Finally, substitution produces (39).  $\square$

#### • Theorem 4

Let the instrument locations form a [2, 2, 2] configuration. Specifically, let

$$\mathbf{y}_1 = \mathbf{y}_2 = \mathbf{a}_1, \quad \mathbf{y}_3 = \mathbf{y}_4 = \mathbf{a}_2, \quad \mathbf{y}_5 = \mathbf{y}_6 = \mathbf{a}_3; \quad (41)$$

and let the vertices  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$  define a triangle  $\Delta$ . Then

$$\max \{\delta(\mathbf{A}) : \text{all } |\mathbf{u}_i| = 1\} = (\text{area of } \Delta)(\text{perimeter of } \Delta). \quad (42)$$

Also, the  $\hat{\mathbf{u}}_i$  are any unit vectors when the triangle area is zero. Otherwise  $(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2), (\hat{\mathbf{u}}_3, \hat{\mathbf{u}}_4), (\hat{\mathbf{u}}_5, \hat{\mathbf{u}}_6)$  are pairs of orthogonal unit vectors, and if

$$\mathbf{v}_1 = \mathbf{u}_1 \times \mathbf{u}_2, \quad \mathbf{v}_2 = \mathbf{u}_3 \times \mathbf{u}_4, \quad \mathbf{v}_3 = \mathbf{u}_5 \times \mathbf{u}_6, \quad (43)$$

then each  $\hat{\mathbf{v}}_i$  parallels the angle bisector at the corresponding  $\mathbf{a}_i$ .

#### Proof

If the area is zero, then the  $\mathbf{y}_i$  are necessarily collinear, and the result is trivial since  $\delta(\mathbf{A})$  is identically zero. Otherwise, by translation-invariance, suppose  $\mathbf{a}_3 = 0$ , and, in this coordinate system, define

$$\mathbf{b}_1 = \mathbf{a}_1/|\mathbf{a}_1|, \quad \mathbf{b}_2 = \mathbf{a}_2/|\mathbf{a}_2|, \quad \mathbf{n} = \mathbf{a}_1 \times \mathbf{a}_2/|\mathbf{a}_1 \times \mathbf{a}_2|, \quad (44)$$

$$\gamma = (|\mathbf{a}_2|/|\mathbf{a}_1|)^{1/2}.$$

Now  $\langle \mathbf{a}_1, \mathbf{a}_2 \rangle$  is the plane of the triangle, and  $\mathbf{n}$  is the normal to this plane. Also Lemma 3 implies that  $(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2), (\hat{\mathbf{u}}_3,$

$\hat{\mathbf{u}}_4), (\hat{\mathbf{u}}_5, \hat{\mathbf{u}}_6)$  are orthogonal pairs, whence maximal  $\delta(\mathbf{A})$  requires that  $\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \hat{\mathbf{v}}_3$  be unit vectors. Moreover, we can obtain directly

$$\det(\mathbf{A}) = |\mathbf{a}_1 \times \mathbf{a}_2| (|\mathbf{a}_1| \cdot |\mathbf{a}_2|)^{1/2} \varepsilon(\mathbf{A}),$$

$$\varepsilon(\mathbf{A}) = (1/\gamma)(\mathbf{b}_1 \cdot \mathbf{v}_1)(\mathbf{n} \times \mathbf{v}_2 \cdot \mathbf{v}_3) - \gamma(\mathbf{b}_2 \cdot \mathbf{v}_2)(\mathbf{n} \times \mathbf{v}_1 \cdot \mathbf{v}_3) \quad (45)$$

by relations (33) and (4). Therefore, we need only maximize  $|\varepsilon(\mathbf{A})|$  for unit vectors  $\mathbf{v}_1, \mathbf{v}_2$ , and  $\mathbf{v}_3$ .

But  $\hat{\mathbf{v}}_3$ , by Lemma 2, is orthogonal to  $\mathbf{n}$ ; so  $\hat{\mathbf{v}}_3$ , by (44), is contained in  $\langle \mathbf{a}_1, \mathbf{a}_2 \rangle$ . Thus we may restrict the variable vector  $\mathbf{v}_3$  to this plane, and we may consider the rotated vector  $\mathbf{v}_0 = \mathbf{v}_3 \times \mathbf{n}$  in the same plane. If we introduce the bisector  $\mathbf{b}_0$  of  $\{\mathbf{b}_1, \mathbf{b}_2\}$  and we define the unit vector  $\mathbf{f}(\theta) = \mathbf{b}_0 \cos \theta + (\mathbf{n} \times \mathbf{b}_0) \sin \theta$  in  $\langle \mathbf{a}_1, \mathbf{a}_2 \rangle$ , then  $\mathbf{b}_1 = \mathbf{f}(-\phi_0)$ ,  $\mathbf{b}_2 = \mathbf{f}(\phi_0)$  for some given  $\phi_0$ , while  $\mathbf{v}_0 = \mathbf{f}(\phi)$  for some unknown  $\phi$ . Actually, the permutation symmetry of  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  permits assuming all three variable vectors in  $\langle \mathbf{a}_1, \mathbf{a}_2 \rangle$ . Moreover, we can eliminate the unknown  $\mathbf{v}_1, \mathbf{v}_2$  by Lemma 5, because

$$\varepsilon(\mathbf{A}) = \mathbf{v}_1^T \mathbf{B} \mathbf{v}_2, \quad \mathbf{B} = (1/\gamma) \mathbf{b}_1 \mathbf{v}_0^T - \gamma \mathbf{v}_0 \mathbf{b}_2^T; \quad (46)$$

indeed, we need only calculate the largest eigenvalue of  $\mathbf{B}^T \mathbf{B}$ , where

$$\mathbf{B}^T \mathbf{B} = \gamma^{-2} \mathbf{v}_0 \mathbf{v}_0^T + \gamma^2 \mathbf{b}_2 \mathbf{b}_2^T - \cos(\phi + \phi_0)[\mathbf{v}_0 \mathbf{b}_2^T + \mathbf{b}_2 \mathbf{v}_0^T]. \quad (47)$$

If a vector in the plane  $\langle \mathbf{a}_1, \mathbf{a}_2 \rangle$  has the form  $\xi \mathbf{v}_0 + \eta \mathbf{b}_2$ , then its image under the mapping  $\mathbf{B}^T \mathbf{B}$  has the form  $\xi_* \mathbf{v}_0 + \eta_* \mathbf{b}_2$ , where  $(\xi_*, \eta_*)^T = \mathbf{M}(\xi, \eta)^T$ , and

$$\mathbf{M} = \begin{pmatrix} \gamma^{-2} - \cos(\phi + \phi_0) \cos(\phi - \phi_0) & \gamma^{-2} \cos(\phi - \phi_0) - \cos(\phi + \phi_0) \\ \gamma^2 \cos(\phi - \phi_0) - \cos(\phi + \phi_0) & \gamma^2 - \cos(\phi + \phi_0) \cos(\phi - \phi_0) \end{pmatrix}. \quad (48)$$

Hence the nonzero eigenvalues satisfy the equation  $0 = \det(\lambda - \mathbf{M})$ , and this quadratic equation has the roots

$$\lambda_{\pm} = (\rho + \sigma_0) + (\rho + \sigma) \pm 2(\rho + \sigma_0)^{1/2}(\rho + \sigma)^{1/2},$$

$$\rho = (\gamma^{-1} - \gamma)^2/4, \quad \sigma_0 = \sin^2 \phi_0, \quad \sigma = \sin^2 \phi. \quad (49)$$

To obtain the maximum of the largest eigenvalue  $\lambda_+$  clearly demands the largest value, unity, of the variable  $\sigma$ . Thus  $\hat{\mathbf{v}}_0$  is orthogonal to  $\mathbf{b}_0$ , or  $\hat{\mathbf{v}}_3$  is parallel to  $\mathbf{b}_0$ . Geometrically, each  $\hat{\mathbf{v}}_i$ , by permutation symmetry, is therefore a bisector of the corresponding angle.

Now define the triangle  $\Delta$  by its vertices  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ , and locate the intersection of its angle bisectors by the vector  $\mathbf{a}_0$ . Also assume the outward sense for the unit bisectors  $\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \hat{\mathbf{v}}_3$ , and denote the interior angles at the same points by the quantities  $\theta_i$ . The auxiliary triangle  $\Delta_1$  with vertices  $\mathbf{a}_0, \mathbf{a}_2, \mathbf{a}_3$  has interior angles  $\theta_2/2, \theta_3/2$  at the last

two listed points. But the remaining angle is  $\angle(\hat{\mathbf{v}}_2, \hat{\mathbf{v}}_3)$ , whence this angle is  $\pi - (\theta_2 + \theta_3)/2$ ; and the  $\theta_i$  have sum  $\pi$ , whence

$$\mathbf{n} \cdot \hat{\mathbf{v}}_2 \times \hat{\mathbf{v}}_3 = \sin((\pi + \theta_1)/2) = \cos(\theta_1/2) = \mathbf{b}_1 \cdot \hat{\mathbf{v}}_1. \quad (50)$$

A similar argument yields the cycled identity

$$\mathbf{n} \cdot \hat{\mathbf{v}}_3 \times \hat{\mathbf{v}}_1 = \sin((\pi + \theta_2)/2) = \cos(\theta_2/2) = \mathbf{b}_2 \cdot \hat{\mathbf{v}}_2, \quad (51)$$

while the triangle  $\Delta$  has area  $|\mathbf{a}_1 \times \mathbf{a}_2|/2$ . These identities imply (42), because (45) becomes

$$\begin{aligned} \max|\det(\mathbf{A})| &= |\mathbf{a}_1 \times \mathbf{a}_2| \{|\mathbf{a}_1|(\mathbf{b}_1 \cdot \hat{\mathbf{v}}_1)(\mathbf{n} \cdot \hat{\mathbf{v}}_2 \times \hat{\mathbf{v}}_3) \\ &\quad + |\mathbf{a}_2|(\mathbf{b}_2 \cdot \hat{\mathbf{v}}_2)(\mathbf{n} \cdot \hat{\mathbf{v}}_3 \times \hat{\mathbf{v}}_1)\} \\ &= (\text{area of } \Delta) \{|\mathbf{a}_1|(1 + \cos \theta_1) + |\mathbf{a}_2|(1 + \cos \theta_2)\} \\ &= (\text{area of } \Delta) \{|\mathbf{a}_1| + |\mathbf{a}_2| + |\mathbf{a}_1 - \mathbf{a}_2|\}. \quad \square \end{aligned} \quad (52)$$

## 6. General results for [1, 1, 1, 3] configurations

The next two sections optimize four-position layouts. Specifically, Theorem 5 treats nondegenerate [1, 1, 1, 3] configurations, whereas Theorem 3 includes the remaining possibilities. The forthcoming theorem requires a preliminary normalization. Any unit vectors  $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$  define relative angles  $\theta_{12}, \theta_{23}, \theta_{31}$  by

$$\mathbf{z}_j \cdot \mathbf{z}_k = \cos \theta_{jk} \text{ with } 0 \leq \theta_{jk} = \theta_{kj} \leq \pi. \quad (53)$$

Any nonzero vectors  $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$  define unit vectors  $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$ , unique up to signs, with each  $\mathbf{z}_i$  in  $\langle \mathbf{y}_i \rangle$ . However, we can choose these signs, when we have any such  $\mathbf{y}_i$ , so that

$$\theta_{12} + \theta_{23} \leq \pi, \theta_{23} + \theta_{31} \leq \pi, \theta_{31} + \theta_{12} \leq \pi. \quad (54)$$

Indeed, if we change  $\mathbf{z}_i \rightarrow -\mathbf{z}_i$ , where  $i, j, k$  are distinct indices, then clearly we map  $\theta_{jk} \rightarrow \theta_{jk}$  but  $\theta_{ij} \rightarrow \pi - \theta_{ij}$ ,  $\theta_{ik} \rightarrow \pi - \theta_{ik}$ . Either we can eliminate all obtuse relative angles by such reversals, or, if this reduction is impossible, then we can eliminate all acute relative angles in this way. But we can immediately satisfy (54) in the first case; while, if angle  $\theta_{jk}$  is minimal, then we need only reverse  $\mathbf{z}_i$  in the second case.

### • Theorem 5

Given any pairwise independent vectors  $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ , but taking  $\mathbf{y}_4 = \mathbf{y}_5 = \mathbf{y}_6 = 0$ , choose the normalized unit vectors  $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$  to make the relative angles satisfy (54). Now, if  $\sigma = \theta_{12} + \theta_{23} + \theta_{31}$ , then

$$\max \{ \delta(\mathbf{A}) : \text{all } |\mathbf{u}_i| = 1 \} / |\mathbf{y}_1| \cdot |\mathbf{y}_2| \cdot |\mathbf{y}_3| = \sin(\sigma/2) \text{ or } 1, \quad (55)$$

according as  $0 \leq \sigma \leq \pi$  or  $\pi \leq \sigma \leq 2\pi$ . Also,  $(\hat{\mathbf{u}}_4, \hat{\mathbf{u}}_5, \hat{\mathbf{u}}_6)$ , in either case, is any orthonormal basis for  $E^3$ . If  $\sigma \leq \pi$ , and  $(i, j, k)$  is any permutation of  $(1, 2, 3)$ , then  $\hat{\mathbf{u}}_i$  bisects

$\{\mathbf{z}_{ij}, \mathbf{z}_{ik}\}$ , where  $\mathbf{z}_{ij} = \mathbf{z}_i \times \mathbf{z}_j$  for short. If  $\sigma > \pi$ , then  $\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3$  do not have unique alignments.

### Proof

Lemma 4 already implies an arbitrary orthonormal basis  $(\hat{\mathbf{u}}_4, \hat{\mathbf{u}}_5, \hat{\mathbf{u}}_6)$ . Relation (34) therefore entails

$$\begin{aligned} \max |\det(\mathbf{A})| &= |\mathbf{y}_1| \cdot |\mathbf{y}_2| \cdot |\mathbf{y}_3| \\ &\quad \cdot \max \{ |\det(\mathbf{z}_1 \times \mathbf{u}_1, \mathbf{z}_2 \times \mathbf{u}_2, \mathbf{z}_3 \times \mathbf{u}_3)| : \text{all } |\mathbf{u}_i| = 1 \}. \end{aligned} \quad (56)$$

But the last determinant equals  $\mathbf{u}_3 \cdot \mathbf{r}(\mathbf{u}_1, \mathbf{u}_2)$ , where

$$\mathbf{r}(\mathbf{u}_1, \mathbf{u}_2) = [(\mathbf{z}_1 \times \mathbf{u}_1) \times (\mathbf{z}_2 \times \mathbf{u}_2)] \times \mathbf{z}_3. \quad (57)$$

Thus  $\hat{\mathbf{u}}_3$  parallels  $\mathbf{r}(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2)$  by Lemma 2, and  $\hat{\mathbf{u}}_3 \cdot \mathbf{z}_3 = 0$  by (57), whence all  $\hat{\mathbf{u}}_i \cdot \mathbf{z}_i = 0$  by symmetry, and each  $\mathbf{z}_i \times \hat{\mathbf{u}}_i$  has unit length as a result. Also  $\mathbf{z}_{jk} \neq 0$  for distinct  $j, k$ , and  $|\mathbf{z}_{jk}| = |\sin \theta_{jk}|$  by definition (53). However, these facts suggest the orthonormal basis

$$\mathbf{f}_3 = \mathbf{z}_2, \quad \mathbf{f}_1 = \mathbf{z}_{32}/|\mathbf{z}_{32}|, \quad \mathbf{f}_2 = \mathbf{f}_3 \times \mathbf{f}_1; \quad (58)$$

and Lemma 2 offers the additional reduction

$$\begin{aligned} [\max \delta(\mathbf{A})]^2 / |\mathbf{y}_1|^2 |\mathbf{y}_2|^2 |\mathbf{y}_3|^2 \\ &= \max \{ |\mathbf{r}(\mathbf{u}_1, \mathbf{u}_2)|^2 : |\mathbf{u}_1| = |\mathbf{u}_2| = 1 \} \\ &= \max \{ (\mathbf{z}_{31} \cdot \mathbf{u}_1)^2 + (\mathbf{z}_{32} \cdot \mathbf{u}_2)^2 \\ &\quad - 2(\mathbf{z}_{31} \cdot \mathbf{u}_1)(\mathbf{z}_{32} \cdot \mathbf{u}_2)(\mathbf{z}_{02} \cdot \mathbf{u}_2) : |\mathbf{u}_1| = |\mathbf{u}_2| = 1 \} \\ &= \max \{ (\mathbf{z}_{31} \cdot \mathbf{u}_1)^2 + \mathbf{u}_2^T \mathbf{C} \mathbf{u}_2 : |\mathbf{u}_1| = |\mathbf{u}_2| = 1 \}. \end{aligned} \quad (59)$$

Here we have abbreviated  $\mathbf{z}_{02} = (\mathbf{z}_1 \times \mathbf{u}_1) \times \mathbf{z}_2$ ; and we have introduced

$$\begin{aligned} \mathbf{C} &= [\sin^2 \theta_{23} - 2(\mathbf{z}_{31} \cdot \mathbf{u}_1)(\mathbf{z}_{32} \cdot \mathbf{z}_{02})] \mathbf{f}_1 \mathbf{f}_1^T \\ &\quad - (\mathbf{z}_{31} \cdot \mathbf{u}_1)(\mathbf{z}_{32} \times \mathbf{z}_{02} \cdot \mathbf{f}_3) [\mathbf{f}_1 \mathbf{f}_2^T + \mathbf{f}_2 \mathbf{f}_1^T]. \end{aligned} \quad (60)$$

Now, the maximum (59) over the unit vector  $\mathbf{u}_2$  yields the largest eigenvalue of  $\mathbf{C}$ , and the representation (60) in the basis  $\{\mathbf{f}_i\}$  gives a  $3 \times 3$  matrix for  $\mathbf{C}$ . The third row and column of this matrix have no nonzero entries, by definition (60). Thus the nonzero eigenvalues  $\lambda_{\pm}$  satisfy the characteristic equation for the nontrivial  $2 \times 2$  block, and the maximal eigenvalue  $\lambda_+$  produces the further simplification

$$\begin{aligned} 2|\mathbf{r}(\mathbf{u}_1, \hat{\mathbf{u}}_2)|^2 &= 1 + \cos \theta_{23} \cdot (2\omega - \cos \theta_{23}) \\ &\quad + |\sin \theta_{23}| \cdot [1 - (2\omega - \cos \theta_{23})^2]^{1/2}, \end{aligned} \quad (61)$$

with  $\omega(\mathbf{u}_1) = (\mathbf{z}_{21} \cdot \mathbf{u}_1)(\mathbf{z}_{31} \cdot \mathbf{u}_1)$ . The derivation of (61) entails the reality of the square root. Hence some real angle  $\gamma(\mathbf{u}_1)$  satisfies  $\cos \gamma(\mathbf{u}_1) = \cos \theta_{23} - 2\omega(\mathbf{u}_1)$ , and the relation (61) becomes

$$2|\mathbf{r}(\mathbf{u}_1, \hat{\mathbf{u}}_2)|^2 = 1 - \cos \gamma \cos \theta_{23} + |\sin \gamma \sin \theta_{23}|. \quad (62)$$

However, Lemma 6 proves that

Range  $[2\omega(\mathbf{u}_1)]$

$$= [(\mathbf{z}_{21} \cdot \mathbf{z}_{31}) - |\mathbf{z}_{21}| \cdot |\mathbf{z}_{31}|, (\mathbf{z}_{21} \cdot \mathbf{z}_{31}) + |\mathbf{z}_{21}| \cdot |\mathbf{z}_{31}|], \quad (63)$$

while definitions imply  $\mathbf{z}_{21} \cdot \mathbf{z}_{31} = \cos \theta_{23} - \cos \theta_{12} \cos \theta_{31}$ ; and relations (54) show that  $|\theta_{12} - \theta_{31}| \leq \theta_{12} + \theta_{31} \leq \pi$ , whence substitution produces

$$\text{Range} [\cos \gamma(\mathbf{u}_1)] = [\cos (\theta_{12} + \theta_{31}), \cos |\theta_{12} - \theta_{31}|]. \quad (64)$$

Clearly (62) can become 2, and (55) be unity, precisely when  $\cos \gamma$  can become  $\cos (\pi - \theta_{23}) = -\cos \theta_{23}$ , or when  $|\theta_{12} - \theta_{31}| \leq \pi - \theta_{23} \leq \theta_{12} + \theta_{31}$ . Moreover (54) already insures the first inequality; therefore the conclusion requires only  $\pi \leq \sigma$ . If we let

$$\mathbf{u}_1 = [\xi_2(\mathbf{z}_1 \times \mathbf{z}_{21}) + \xi_3(\mathbf{z}_1 \times \mathbf{z}_{31})]/\det(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) \quad (65)$$

for some  $\xi_2, \xi_3$ , and we use  $|\hat{\mathbf{u}}_1|^2 = 1, \omega(\hat{\mathbf{u}}_1) = \cos \theta_{23}$ , then we find

$$\begin{aligned} [\xi_2 \sin \theta_{12} + \xi_3 \sin \theta_{13}]^2 &= \sin^2 \theta_{12} + \sin^2 \theta_{13} - \sin^2 \theta_{23} \\ &\quad - 2\xi \sin \theta_{12} \sin \theta_{13} \cos \theta_{23} \end{aligned} \quad (66)$$

with  $\xi = \pm$ . This quadratic pair of equations has two distinct pairs of solutions.

If  $\sigma < \pi$ , then  $|\mathbf{r}(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2)| < 1$ . Therefore,  $\cos \gamma(\hat{\mathbf{u}}_1)$  must be an endpoint of the range (64), because (62) can have no smooth maximum in the variable  $\cos \gamma(\mathbf{u}_1)$ . Indeed, the smallest  $\cos \gamma$ , by trial, yields the largest  $|\mathbf{r}|$ ; and the corresponding  $\hat{\mathbf{u}}_1$ , by Lemma 6, bisects  $\{\mathbf{z}_{21}, \mathbf{z}_{31}\}$ ; while the resulting (55), by calculation, is the stated  $\sin(\sigma/2)$ .  $\square$

#### Remarks

Each line  $\langle y_i \rangle$  determines an orthogonal plane through the origin, and these three planes define four opposite pairs of solid angles. Moreover the points  $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$ , on the unit sphere, specify the vertices of a spherical triangle, and pairwise intersections of the orthogonal planes contain the vertices of its *polar triangle* [28, p. 112]. Both the polar triangle and its opposite have vertex angles  $\pi - \theta_{jk}$  and area  $2\pi - \sigma$  [28, pp. 112-114]. Hence the remaining solid angles have total measure  $2\sigma$ , and these traditional geometric concepts yield a dual condition:  $\sigma \leq \pi$  precisely when some solid angle exceeds  $\pi$ . If we take the cosine of the equivalent statement  $\theta_{12} + \theta_{31} \leq \pi - \theta_{23}$ , and we square the result, in the modified form  $\sin \theta_{12} \sin \theta_{31} \leq \cos \theta_{12} \cos \theta_{31} + \cos \theta_{23}$ , then we find an algebraic version of this requirement:

$$\begin{aligned} 1 &\leq (\mathbf{z}_1 \cdot \mathbf{z}_2)^2 + (\mathbf{z}_2 \cdot \mathbf{z}_3)^2 + (\mathbf{z}_3 \cdot \mathbf{z}_1)^2 \\ &\quad + 2(\mathbf{z}_1 \cdot \mathbf{z}_2)(\mathbf{z}_2 \cdot \mathbf{z}_3)(\mathbf{z}_3 \cdot \mathbf{z}_1) \\ &= \cos^2 \theta_{12} + \cos^2 \theta_{23} + \cos^2 \theta_{31} \\ &\quad + 2 \cos \theta_{12} \cos \theta_{23} \cos \theta_{31}. \end{aligned} \quad (67)$$

The homogeneity of relation (67) obviates our normalization of the  $\mathbf{z}_i$ . If we reverse any  $\mathbf{z}_i$ , then we change nothing.

#### 7. Special results for [1, 1, 2, 2] configurations

The one theorem of this section optimizes a restricted family of [1, 1, 2, 2] configurations. This special result, with Theorem 5, covers the relevant layouts with four distinct positions. Some remarks at the end explore the problem of further extensions.

##### Theorem 6

If  $\mathbf{y}_1 = -\mathbf{y}_2 = \mathbf{b}, \mathbf{y}_3 = \mathbf{y}_4 = \mathbf{a}, \mathbf{y}_5 = \mathbf{y}_6 = 0$ , where  $\mathbf{a}, \mathbf{b}$  are arbitrary vectors, then

$$\max \{\delta(\mathbf{A}) : \text{all } |\mathbf{u}_i| = 1\} = 2|\mathbf{a}| \cdot |\mathbf{a} \times \mathbf{b}| \quad (68)$$

wherever  $|\mathbf{b}| \leq |\mathbf{a}|$ . If  $\mathbf{a} \times \mathbf{b} = 0$ , then the  $\hat{\mathbf{u}}_i$  are arbitrary unit vectors. If  $\mathbf{a} \times \mathbf{b} \neq 0$  and  $\mathbf{f}_0 = \mathbf{a} \times \mathbf{b}/|\mathbf{a} \times \mathbf{b}|$ , then  $\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2$  are unit vectors both parallel to  $\mathbf{f}_0$ , and  $(\hat{\mathbf{u}}_3, \hat{\mathbf{u}}_4), (\hat{\mathbf{u}}_5, \hat{\mathbf{u}}_6)$  are any orthonormal pairs respectively perpendicular to  $\mathbf{a}, \mathbf{f}_0$ .

##### Proof

All these statements are trivial when  $\mathbf{a} \times \mathbf{b} = 0$ , because the  $\mathbf{y}_i$  are collinear, and  $\delta(\mathbf{A}) \equiv 0$ . Otherwise  $(\hat{\mathbf{u}}_3, \hat{\mathbf{u}}_4), (\hat{\mathbf{u}}_5, \hat{\mathbf{u}}_6)$  are orthonormal pairs by Lemma 3; and if  $\mathbf{v}_1 = \mathbf{u}_5 \times \mathbf{u}_6, \mathbf{v}_2 = \mathbf{u}_3 \times \mathbf{u}_4$  in (33), then

$$\begin{aligned} -\det(\mathbf{A})/|\mathbf{a} \times \mathbf{b}| &= \det(\mathbf{b}, \mathbf{u}_1, \mathbf{u}_2) \det(\mathbf{f}_0, \mathbf{v}_1, \mathbf{v}_2) \\ &\quad + (\mathbf{a} \cdot \mathbf{v}_2)[(\mathbf{v}_1 \cdot \mathbf{u}_1)(\mathbf{f}_0 \cdot \mathbf{u}_2) + (\mathbf{f}_0 \cdot \mathbf{u}_1)(\mathbf{v}_1 \cdot \mathbf{u}_2)] \end{aligned} \quad (69)$$

by (4). If we define  $\kappa = |\mathbf{b}| \det(\mathbf{f}_0, \mathbf{v}_1, \mathbf{v}_2)/(\mathbf{a} \cdot \mathbf{v}_2)$  and introduce

$$\mathbf{f}_{-1} = \mathbf{a}/|\mathbf{a}|, \quad \mathbf{f}_{-2} = \mathbf{f}_0 \times \mathbf{f}_{-1}, \quad \mathbf{f}_1 = \mathbf{b}/|\mathbf{b}|, \quad \mathbf{f}_2 = \mathbf{f}_0 \times \mathbf{f}_1, \quad (70)$$

then we rewrite (69), and obtain

$$\begin{aligned} -\det(\mathbf{A})/|\mathbf{a} \times \mathbf{b}| &= (\mathbf{a} \cdot \mathbf{v}_2) \mathbf{u}_1^T \mathbf{B} \mathbf{u}_2, \\ \mathbf{B} &= \mathbf{f}_0 \mathbf{v}_1^T + \mathbf{v}_1 \mathbf{f}_0^T + \kappa[\mathbf{f}_2 \mathbf{f}_0^T - \mathbf{f}_0 \mathbf{f}_2^T]. \end{aligned} \quad (71)$$

Thus  $\max \{\delta(\mathbf{A})^2/|\mathbf{a} \times \mathbf{b}|^2 : |\mathbf{u}_1| = |\mathbf{u}_2| = 1\}$ , by Lemma 5, is precisely the largest eigenvalue of  $\mathbf{B}^T \mathbf{B}$ . But (71), in the orthonormal basis  $(\mathbf{f}_0, \mathbf{f}_1, \mathbf{f}_2)$ , gives a matrix for  $\mathbf{B}$ , and the roots of the characteristic polynomial  $\det(\lambda - \mathbf{B}^T \mathbf{B})$  are the eigenvalues of  $\mathbf{B}^T \mathbf{B}$ . Also  $\text{rank}(\mathbf{B}) = 2$ , whence one root is zero. Now the other roots, by direct calculation, are

$$\lambda_{\pm} = \kappa^2 + (\mathbf{f}_0 \cdot \mathbf{v}_1)^2 \pm 2[(\mathbf{f}_0 \cdot \mathbf{v}_1)^2 + \kappa^2 - \kappa^2(\mathbf{f}_1 \cdot \mathbf{v}_1)^2]^{1/2}, \quad (72)$$

and relation (71), via this reduction, yields

$$\begin{aligned}
& \max \{ \delta(A)^2 / |a \times b|^2 : |u_1| = |u_2| = 1 \} \\
&= (a \cdot v_2)^2 [1 + (f_0 \cdot v_1)^2] + |b|^2 (v_1 \cdot v_2 \times f_0)^2 \\
&\quad + 2|a \cdot v_2| \{ (a \cdot v_2)^2 (f_0 \cdot v_1)^2 \\
&\quad + |b|^2 (v_1 \cdot v_2 \times f_0)^2 [1 - (f_1 \cdot v_1)^2] \}^{1/2} \\
&\leq \{ |a \cdot v_2| + [(a \cdot v_2)^2 (f_0 \cdot v_1)^2 + |b|^2 (v_1 \cdot v_2 \times f_0)^2]^{1/2} \}^2. \quad (73)
\end{aligned}$$

Here our manipulations, for convenience, have assumed nonzero  $a \cdot v_2$ , but the result (73), by continuity, does not require this.

The expression  $(a \cdot v_2)^2 (f_0 \cdot v_1)^2 + |b|^2 (v_1 \cdot v_2 \times f_0)^2$  in the last bracket is a quadratic form in  $v_1$ . Indeed,  $\max \{ (a \cdot v_2)^2, |b|^2 |v_2 \times f_0|^2 \}$  is its maximum value, because  $f_0$  and  $v_2 \times f_0$  are orthogonal vectors. But (73) becomes  $4|a|^2$  when  $v_1 = \pm f_0$ ,  $v_2 = \pm f_{-1}$ . Thus the square root of (73) implies

$$\begin{aligned}
2|a| &\leq \max \{ \delta(A) : \text{all } |u_i| = 1 \} \\
&\leq |a \cdot \hat{v}_2| + \max \{ |a \cdot \hat{v}_2|, |b| |\hat{v}_2 \times f_0| \} \\
&\leq \max \{ 2|a \cdot \hat{v}_2|, |a \cdot \hat{v}_2| + |b| \} \quad (74)
\end{aligned}$$

by definition. Therefore,  $\hat{v}_2 = \pm f_{-1}$  whenever  $|b| \leq |a|$ ; and if we insert this value, then we obtain

$$\begin{aligned}
4 &\leq \max \{ \delta(A)^2 / |a|^2 |a \times b|^2 : \text{all } |u_i| = 1 \} \\
&\leq 1 + (f_0 \cdot \hat{v}_1)^2 + (|b|/|a|)^2 (f_{-2} \cdot \hat{v}_1)^2 \\
&\quad + 2\{ (f_0 \cdot \hat{v}_1)^2 + (|b|/|a|)^2 (f_{-2} \cdot \hat{v}_1)^2 [1 - (f_1 \cdot \hat{v}_1)^2] \}^{1/2} \\
&\leq \{ 1 + [(f_0 \cdot \hat{v}_1)^2 + (|b|/|a|)^2 (f_{-2} \cdot \hat{v}_1)^2]^{1/2} \}^2. \quad (75)
\end{aligned}$$

Therefore,  $\hat{v}_1 = \pm f_0$  whenever  $|b| < |a|$ . However, if  $|b| = |a|$  and  $E = f_0 f_0^T + f_{-2} f_{-2}^T$ , then

$$\begin{aligned}
4 &\leq \max \{ \delta(A)^2 / |a|^2 |a \times b|^2 : \text{all } |u_i| = 1 \} \\
&= 1 + |E \hat{v}_1|^2 + 2\{ |E \hat{v}_1|^2 - (f_1 \cdot \hat{v}_1)^2 (f_{-2} \cdot \hat{v}_1)^2 \}^{1/2}; \quad (76)
\end{aligned}$$

whence  $E \hat{v}_1 = \hat{v}_1$ , and either  $f_{-2} \cdot \hat{v}_1 = 0$  or  $f_1 \cdot \hat{v}_1 = 0$ . Clearly,  $\hat{v}_1$ , in the first case, is orthogonal to  $\langle f_{-1}, f_{-2} \rangle$ , and  $\hat{v}_1$ , in the second case, is orthogonal to  $\langle f_{-1}, f_1 \rangle = \langle a, b \rangle$ , whence  $\hat{v}_1$ , in either case, is parallel to  $f_0$ . If we use these values in (71), then we find

$$\begin{aligned}
2 &\leq \max \{ \delta(A) / |a| \cdot |a \times b| : \text{all } |u_i| = 1 \} \\
&= 2|f_0 \cdot \hat{u}_1| \cdot |f_0 \cdot \hat{u}_2|, \quad (77)
\end{aligned}$$

by (74). Hence,  $\delta(A)$  attains its maximum when  $\hat{u}_1$  and  $\hat{u}_2$  parallel  $f_0$ .  $\square$

#### • Remarks

All stated conclusions fail when  $|b| > |a|$ ; some optimal directions change, and  $\max \delta(A) > 2|a| \cdot |a \times b|$ . Indeed, let  $v_1 = f_0 \cos \theta + f_{-2} \sin \theta$ , rather than  $f_0$ ; but put  $v_2 = \pm f_{-1}$ , as in Theorem 6. The value of  $|b|/|a|$  does not affect the truth of (73). Hence, from this specialization we obtain

$$\begin{aligned}
& \max \{ \delta(A)^2 / |a|^2 |a \times b|^2 : |u_1| = |u_2| = 1, v_2 = \pm f_{-1} \} \\
&= 2 + \mu\sigma + 2[1 + \mu\sigma - \nu\sigma^2]^{1/2}, \\
&\mu = (|b|/|a|)^2 - 1, \nu = (|b|/|a|)^2 (f_1 \cdot f_{-1})^2, \sigma = \sin^2 \theta; \quad (78)
\end{aligned}$$

and, for small  $\theta$ , we expand

$$\max \{ \delta(A)^2 / |a|^2 |a \times b|^2 : \dots \} = 4 + 2\mu\sigma + O(\sigma^2). \quad (79)$$

If  $|b| > |a|$ , then  $\mu > 0$ . Thus small perturbations of the foregoing  $\hat{v}_1$  yield a larger quantity than the stated  $\max \delta(A)$ , and new values of some  $\hat{u}_i$  give the global maximum under this condition. The required analysis seems to involve further special cases.

## 8. Optimal layouts with natural "feel"

Here the foregoing theorems, combined with certain obvious remarks, yield the optimal layouts permitted by all relevant considerations. Section 4 requires some off-axial accelerometers. Thus our proposed configurations [12] situate some instruments on the attached ring, where this suggested ring encircles the pen axis at its nonwriting end. Indeed, the possible locations, by Section 4, can only be the pen point, the far end, and this ring. Hence  $\alpha$  and  $\beta$ , throughout this section, will, respectively, be the effective pen length and the ring radius. However, a recording pen with natural "feel" demands relatively unchanged moments of inertia, since an applied torque on any rigid body affects its angular motion through these moments [20, Chap. 5]. This does not preclude off-axial accelerometers, but it urges relatively few such instruments. It also suggests a small ratio  $\beta/\alpha$ , and this further motivates optimal measurement directions. (Natural "feel" still leaves considerable freedom to the designer; the local Radio Shack uses a special pen with triangular cross-section.) Our successive paragraphs will allow first one, then two, then three off-axial instruments; the ultimate choice must reflect some further practical circumstance.

Moreover, optimal solutions will plausibly exhibit high symmetry, and symmetric locations will obviously facilitate normal writing. Therefore, the analysis of this section assumes another constraint, with no further justification, namely, that any accelerometers on the ring occupy positions with equiangular spacing. But a desire for the fewest instruments entails a layout with precisely six accelerometers, and Theorem 2, Part 3, for nonzero  $\delta(A)$ , permits at most three instruments in any one spot. Let the symbol  $[\ell, m]$ , for any integers  $\ell$  and  $m$ , denote a layout with  $\ell$  end, and  $m$  point, accelerometers. Our proposed three cases, and the last few remarks, imply

$$0 \leq \ell \leq 3, 0 \leq m \leq 3, 3 \leq \ell + m \leq 5. \quad (80)$$

Our discussion employs the pen coordinate system,

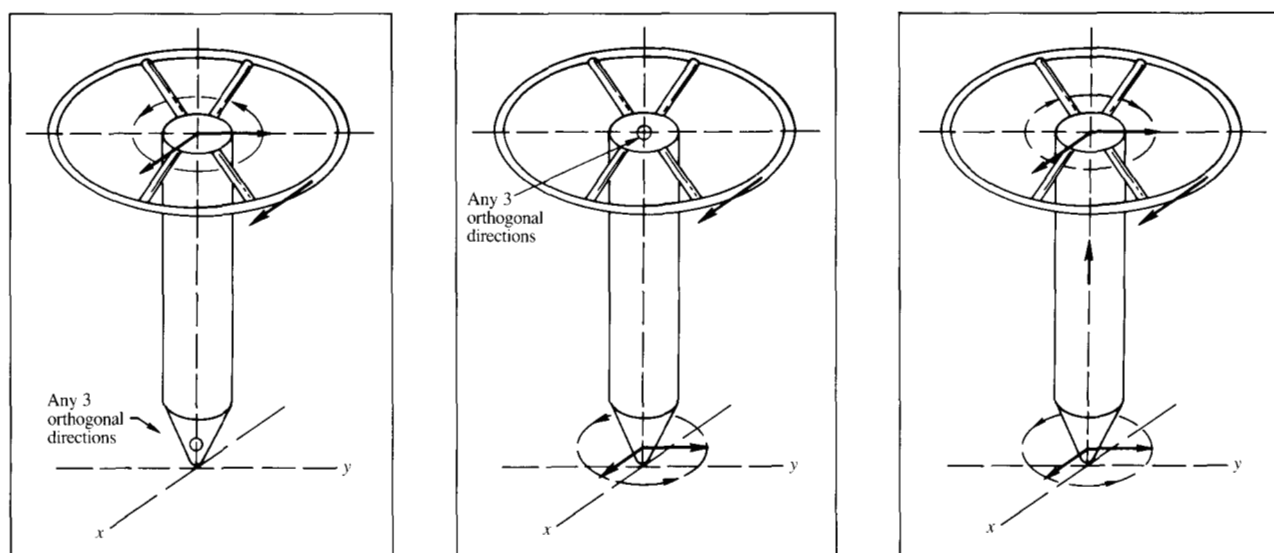


Figure 1 Optimal layouts with one ring accelerometer.

which places the origin at the pen point; it recalls the imbedded orthonormal basis  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ , which aligns  $\mathbf{e}_3$  up the pen axis.

A single off-axis instrument admits two possibilities:  $[3, 2>$  and  $[2, 3>$ . Theorem 3, in each subcase, yields the same maximum for  $\delta(\mathbf{A})$ :

$$\max \delta(\mathbf{A}_{[3,2>}) = \max \delta(\mathbf{A}_{[2,3>}) = \alpha^2 \beta. \quad (81)$$

The one accelerometer off the pen axis has a direction along the ring tangent. The coincident pair have any orthogonal directions whose common normal is the pen axis. The remaining three parallel any orthonormal basis. Figure 1 depicts these alternatives.

Any two ring instruments must take opposite positions, and the resulting layouts will have three types:  $[3, 1>$ ,  $[2, 2>$ ,  $[1, 3>$ . Theorems 5 (twice) and 6 yield

$$\begin{aligned} \max \delta(\mathbf{A}_{[3,1>}) &= \alpha \beta^2, \\ \max \delta(\mathbf{A}_{[2,2>}) &= \max \delta(\mathbf{A}_{[1,3>}) = 2\alpha^2 \beta. \end{aligned} \quad (82)$$

Hence the last two provide optimal configurations, and their ring accelerometers measure tangential components. The end direction in the  $[1, 3>$  subcase parallels the straight line between the ring instruments; any orthonormal basis, by Lemma 4, offers the best directions at the pen point. The end directions in the  $[2, 2>$  subcase are any orthonormal vectors in the accelerometer plane; any orthonormal pair both perpendicular to the axis give best point directions for the arrangement. Figure 2 depicts these alternatives.

Three ring instruments with equally spaced positions admit all layouts of four types:  $[3, 0>$ ,  $[2, 1>$ ,  $[1, 2>$ ,  $[0, 3>$ . Taking the origin at the far end, we obtain  $\max \delta(\mathbf{A}_{[3,0>}) = \beta^3$  by Theorem 5; expanding  $\det(\mathbf{A})$  on one row, we conclude  $\max \delta(\mathbf{A}_{[2,1>}) = \text{const.} \alpha \beta^2$  for some constant. To evaluate the last subcase, we specify the ring locations:

$$\mathbf{y}_1 = \alpha \mathbf{e}_3 + \beta \mathbf{e}_1, \quad \mathbf{y}_{2,3} = \alpha \mathbf{e}_3 - (\beta/2) \mathbf{e}_1 \pm (\beta\sqrt{3}/2) \mathbf{e}_2. \quad (83)$$

The angles  $\theta_{12}, \theta_{23}, \theta_{31}$  of Theorem 5 have a common value  $\theta$ , by symmetry. However,

$$(\alpha^2 + \beta^2) \cos \theta = \mathbf{y}_1 \cdot \mathbf{y}_2 = \alpha^2 - (\beta^2/2) \quad (84)$$

by (83), and  $\sigma = 3\theta$  by definition, whence

$$\begin{aligned} 2 \sin^2(\sigma/2) &= 1 - \cos 3\theta = 1 + 3 \cos \theta - 4 \cos^3 \theta \\ &= (1 - \cos \theta)(1 + 2 \cos \theta)^2 \end{aligned} \quad (85)$$

by standard formulas, and

$$\max \delta(\mathbf{A}_{[0,3>}) = (3\sqrt{3}/2) \alpha^2 \beta \approx 2.59808 \alpha^2 \beta \quad (86)$$

by Theorem 5. Again the ring accelerometers measure tangential components, and any orthonormal basis provides point directions. If this orthonormal basis includes the vector  $\pm \mathbf{e}_3$ , and the corresponding instrument assumes an end location, then the resulting  $\delta(\mathbf{A}_{[1,2>})$  reproduces the value (86). Rotational symmetry suggests the conjecture that optimal  $[1, 2>$  layouts have axial end directions. If this is true, then the preceding construction gives another optimal layout. Figure 3 depicts these alternatives.

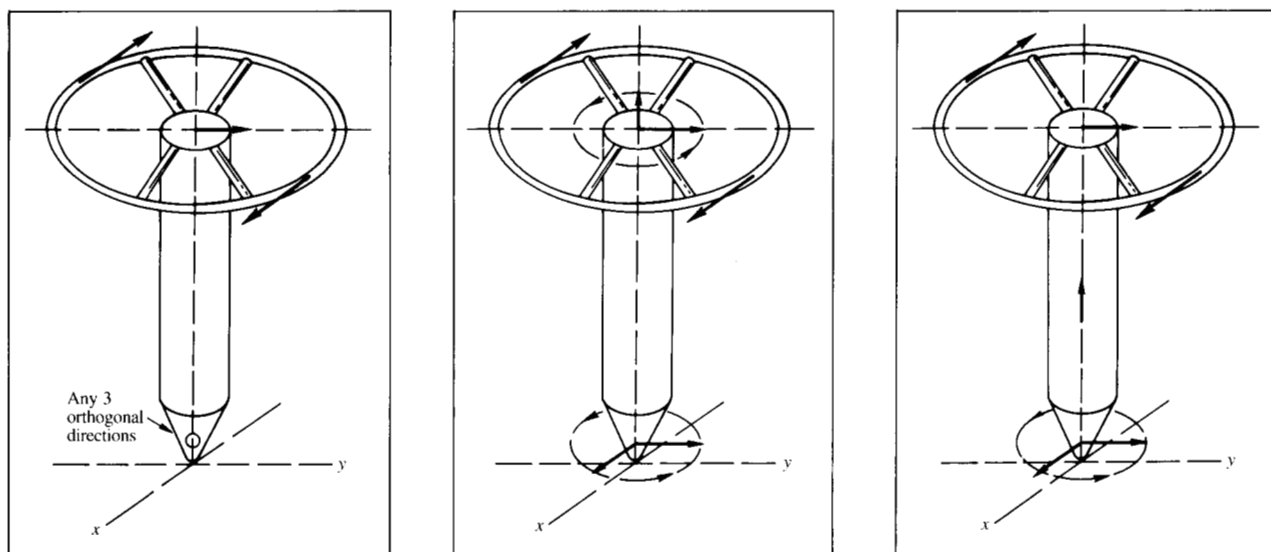


Figure 2 Optimal layouts with two ring accelerometers.

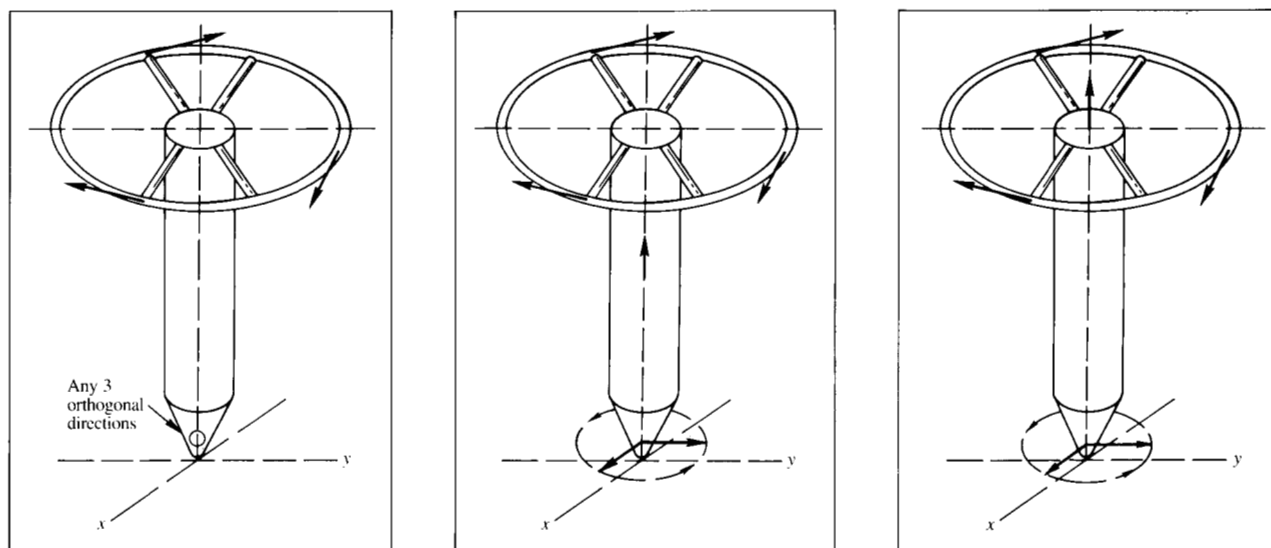


Figure 3 Optimal layouts with three ring accelerometers.

### 9. General properties of tridirectional arrangements

Some designs by previous investigators [13, 14] employ just two of six accelerometers to measure each of three vector components. Hence this section generalizes such layouts, because coincident  $\mathbf{u}_i$  generate useful simplifications. Indeed, subtraction within (8) immediately yields equations without  $\mathbf{U}(\tau)$ . If  $\mathbf{u}_1 = \mathbf{u}_2 = \mathbf{v}$ , for example, while  $\mathbf{y}_1$  and  $\mathbf{y}_2$  are arbitrary locations in the pen, then

$$\mathbf{v} \times (\mathbf{y}_2 - \mathbf{y}_1) \cdot \mathbf{w}' + (\mathbf{v} \times \mathbf{w}) \cdot [(\mathbf{y}_2 - \mathbf{y}_1) \times \mathbf{w}] + \psi_2(\tau) - \psi_1(\tau) = 0 \quad (87)$$

for this configuration. Clearly, three independent relations like (87) allow the recovery of  $\mathbf{w}(\tau)$ , and subsequent integration of (6) provides the values of  $\mathbf{U}(\tau)$ . Any non-singular  $\mathbf{A}$ , of course, permits such a block-triangularization [10], but accelerometers with coincident  $\mathbf{u}_i$  offer a particularly direct reduction. This simple reduction demands at least three independent equalities within the set  $\{\mathbf{u}_1, \dots, \mathbf{u}_6\}$ ; thus it allows at most three distinct vectors among the  $\mathbf{u}_i$ . The rank condition (10) requires at least three independent vectors within this same set; thus it implies precisely three linearly independent vectors

$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  among the chosen  $\mathbf{u}_i$ . An accelerometer layout with this property will be called a *tridirectional* arrangement in this discussion.

If we consider the matrix  $\mathbf{A}^T$  for any such arrangement, then we can relabel the six accelerometers in some order for which  $\mathbf{u}_1 = \mathbf{v}_1, \mathbf{u}_2 = \mathbf{v}_2, \mathbf{u}_3 = \mathbf{v}_3$ . If we subtract the first three columns from suitable others, then we can reduce  $\mathbf{A}^T$  to a block-triangular matrix with the form

$$(\mathbf{A}_{\text{red}})^T = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & 0 \\ \mathbf{y}_1 \times \mathbf{v}_1 & \mathbf{y}_2 \times \mathbf{v}_2 & \mathbf{y}_3 \times \mathbf{v}_3 & \mathbf{A}_0 \end{pmatrix}. \quad (88)$$

Here  $\mathbf{A}_0$  is a  $3 \times 3$  matrix, and

$$\pm \delta(\mathbf{A}) = \det(\mathbf{A}) = \det(\mathbf{A}_{\text{red}}) = \det(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \det(\mathbf{A}_0). \quad (89)$$

Indeed, we can label the six accelerometers so that the vectors  $\mathbf{u}_i$  include these  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  with nonincreasing frequency. Therefore, we can classify all tridirectional arrangements, in that the resulting multiplicities admit just three possible sequences, by direct enumeration:

$$(4, 1, 1); (3, 2, 1); (2, 2, 2). \quad (90)$$

However, this analysis may ignore (4, 1, 1) arrangements, because Theorem 2, Part 4, implies vanishing  $\delta(\mathbf{A}_{(4,1,1)})$ . Also, for a (3, 2, 1) arrangement

$$\mathbf{A}_0 = (\mathbf{z}_1 \times \mathbf{v}_1, \mathbf{z}_2 \times \mathbf{v}_1, \mathbf{z}_3 \times \mathbf{v}_2), \quad (91)$$

where  $\mathbf{z}_1 = \mathbf{y}_4 - \mathbf{y}_1, \mathbf{z}_2 = \mathbf{y}_5 - \mathbf{y}_1, \mathbf{z}_3 = \mathbf{y}_6 - \mathbf{y}_2$ ; and thus, by (4),

$$\begin{aligned} \pm \delta(\mathbf{A}_{(3,2,1)}) &= \det(\mathbf{A}) \\ &= -(\mathbf{v}_1 \times \mathbf{v}_2 \cdot \mathbf{v}_3) (\mathbf{v}_1 \times \mathbf{v}_2 \cdot \mathbf{z}_3) (\mathbf{v}_1 \cdot \mathbf{z}_1 \times \mathbf{z}_2). \end{aligned} \quad (92)$$

Finally, for a (2, 2, 2) arrangement,

$$\mathbf{A}_0 = (\mathbf{z}_1 \times \mathbf{v}_1, \mathbf{z}_2 \times \mathbf{v}_2, \mathbf{z}_3 \times \mathbf{v}_3); \quad (93)$$

where  $\mathbf{z}_1 = \mathbf{y}_4 - \mathbf{y}_1, \mathbf{z}_2 = \mathbf{y}_5 - \mathbf{y}_2, \mathbf{z}_3 = \mathbf{y}_6 - \mathbf{y}_3$ ; and thus, by (89),

$$\begin{aligned} \pm \delta(\mathbf{A}_{(2,2,2)}) &= \det(\mathbf{A}) \\ &= \det(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \det(\mathbf{z}_1 \times \mathbf{v}_1, \mathbf{z}_2 \times \mathbf{v}_2, \mathbf{z}_3 \times \mathbf{v}_3). \end{aligned} \quad (94)$$

We fix the differences  $\mathbf{z}_j$  and seek optimal directions  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ .

#### • Theorem 7

If  $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$  are any given vectors, then

$$\begin{aligned} \max \{ \delta(\mathbf{A}_{(3,2,1)}) : |\mathbf{v}_1| = |\mathbf{v}_2| = |\mathbf{v}_3| = 1 \} \\ = (1/2) \{ |\mathbf{z}_1 \times \mathbf{z}_2| \cdot |\mathbf{z}_3| + |(\mathbf{z}_1 \times \mathbf{z}_2) \times \mathbf{z}_3| \}. \end{aligned} \quad (95)$$

If either  $\mathbf{z}_1 \times \mathbf{z}_2$  or  $\mathbf{z}_3$  is zero, then all  $\hat{\mathbf{v}}_i$  are arbitrary unit

vectors. If both  $\mathbf{z}_1 \times \mathbf{z}_2$  and  $\mathbf{z}_3$  are nonzero, then  $\hat{\mathbf{v}}_2 = \mathbf{n}$ , where  $\mathbf{n}$  is any unit vector normal to  $\langle \mathbf{z}_1 \times \mathbf{z}_2, \mathbf{z}_3 \rangle$ , and  $\hat{\mathbf{v}}_3 = \pm \hat{\mathbf{v}}_1 \times \hat{\mathbf{v}}_2$ , where  $\hat{\mathbf{v}}_1$  is any principal bisector of  $\langle \mathbf{z}_1 \times \mathbf{z}_2 \rangle$  and  $\langle \mathbf{n} \times \mathbf{z}_3 \rangle$ .

#### Proof

We need treat only nonzero  $\mathbf{z}_1 \times \mathbf{z}_2$  and  $\mathbf{z}_3$ ; otherwise these results are all trivial. We must then have nonzero  $\hat{\mathbf{v}}_1 \times \hat{\mathbf{v}}_2$ ; otherwise  $\delta(\mathbf{A}_{(3,2,1)})$  is identically zero. Hence, given  $\hat{\mathbf{v}}_1$  and  $\hat{\mathbf{v}}_2$ , we find  $\hat{\mathbf{v}}_3 = \pm \hat{\mathbf{v}}_1 \times \hat{\mathbf{v}}_2 / |\hat{\mathbf{v}}_1 \times \hat{\mathbf{v}}_2|$  by Lemma 2; and, given  $\hat{\mathbf{v}}_1$  and  $\hat{\mathbf{v}}_3$ , we note

$$\max \delta(\mathbf{A}_{(3,2,1)}) = |\hat{\mathbf{v}}_1 \times \hat{\mathbf{v}}_2|^2 |(\hat{\mathbf{v}}_3 \cdot \mathbf{z}_3)(\hat{\mathbf{v}}_1 \cdot \mathbf{z}_1 \times \mathbf{z}_2)| \quad (96)$$

by computation. Thus  $|\hat{\mathbf{v}}_1 \times \hat{\mathbf{v}}_2|$  must be unity, and the  $\hat{\mathbf{v}}_i$  must be orthonormal. If  $\mathbf{E}$  is the orthogonal projection onto  $\langle \hat{\mathbf{v}}_1, \hat{\mathbf{v}}_3 \rangle$ , then

$$\max \delta(\mathbf{A}_{(3,2,1)}) = |(\hat{\mathbf{v}}_3 \cdot \mathbf{E} \mathbf{z}_3)(\hat{\mathbf{v}}_1 \cdot \mathbf{E}(\mathbf{z}_1 \times \mathbf{z}_2))| \quad (97)$$

by (96). However, maximality of (97) implies  $\mathbf{z}_1 \times \mathbf{z}_2, \mathbf{z}_3$  contained in  $\langle \hat{\mathbf{v}}_1, \hat{\mathbf{v}}_3 \rangle$ , whence orthogonality of the  $\hat{\mathbf{v}}_i$  requires  $\hat{\mathbf{v}}_2$  perpendicular to  $\langle \mathbf{z}_1 \times \mathbf{z}_2, \mathbf{z}_3 \rangle$ . Thus  $\hat{\mathbf{v}}_2$  is the defined  $\mathbf{n}$ , and

$$\max \delta(\mathbf{A}_{(3,2,1)}) = |(\hat{\mathbf{v}}_1 \cdot \mathbf{n} \times \mathbf{z}_3)(\hat{\mathbf{v}}_1 \cdot \mathbf{z}_1 \times \mathbf{z}_2)|. \quad (98)$$

Now Lemma 6 determines  $\hat{\mathbf{v}}_1$ , and (37) produces (95).  $\square$

#### 10. Partial analysis of (2, 2, 2) arrangements

The present section, on (2, 2, 2) arrangements, finds an analytical solution only for two parallel  $\mathbf{z}_i$ . Later, this partial conclusion offers sufficient design guidance, although the complete problem poses unexpectedly greater difficulties. Indeed, the best (3, 2, 1) arrangements, by Theorem 7, require orthonormal  $\hat{\mathbf{v}}_i$  for any nontrivial  $\mathbf{z}_i$ , while the best (2, 2, 2) arrangements, by unpublished results, admit such  $\hat{\mathbf{v}}_i$  only for orthogonal  $\mathbf{z}_i$ . A preliminary result bounds the optimal  $\delta(\mathbf{A})$ .

#### • Theorem 8

Let  $\theta_{12}, \theta_{23}, \theta_{31}$ , as in Theorem 5, be the normalized angles between nonzero vectors  $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$ . If either  $0 \leq \sigma \leq \pi$  or  $\pi \leq \sigma \leq 2\pi$ , where  $\sigma = \theta_{12} + \theta_{23} + \theta_{31}$ , then, respectively,

$$\begin{aligned} &2 \sin(\sigma/2) \text{ or } 2 \\ &\geq 2 \cdot \max \{ \delta(\mathbf{A}_{(2,2,2)}) / |\mathbf{z}_1| |\mathbf{z}_2| |\mathbf{z}_3| : \text{all } |\mathbf{v}_i| = 1 \} \\ &\geq \max(\sin \theta_{12}, \sin \theta_{23}, \sin \theta_{31}) \\ &\quad + \det(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) / |\mathbf{z}_1| |\mathbf{z}_2| |\mathbf{z}_3|. \end{aligned} \quad (99)$$

#### Proof

If we weaken our requirements, then we obtain the upper bound (55). The relations

$$\mathbf{z}_2 = |\mathbf{z}_3| \mathbf{f}_1, \quad \mathbf{f}_1 \times \mathbf{z}_2 = |\mathbf{f}_1 \times \mathbf{z}_2| \mathbf{f}_2, \quad \mathbf{f}_1 \times \mathbf{f}_2 = \mathbf{f}_3, \quad (100)$$

in the given order, define an orthonormal basis from any

vectors  $\mathbf{z}_i$ . If  $\mathbf{v}_2 = \mathbf{f}_2$  and  $\mathbf{v}_3 = \mathbf{f}_3$ , then vector identities produce

$$\delta(\mathbf{A}_{(2,2,2)}) = |(\mathbf{v}_1 \cdot \mathbf{z}_3)(\mathbf{v}_1 \cdot \mathbf{z}_1 \times \mathbf{z}_2)| \quad (101)$$

and Lemma 6 asserts

$$\begin{aligned} 2 \cdot \max \delta(\mathbf{A}_{(2,2,2)}) \\ \geq 2 \cdot \max \{ |(\mathbf{v}_1 \cdot \mathbf{z}_3)(\mathbf{v}_1 \cdot \mathbf{z}_1 \times \mathbf{z}_2)| : |\mathbf{v}_1| = 1 \} \\ = |\mathbf{z}_1 \times \mathbf{z}_2 \cdot \mathbf{z}_3| + |\mathbf{z}_1| \cdot |\mathbf{z}_2| \cdot |\mathbf{z}_3| \sin \theta_{12}. \quad \square \end{aligned} \quad (102)$$

• **Theorem 9**

If  $\mathbf{z}_1, \mathbf{z}_2$  are arbitrary vectors but  $\mathbf{z}_2 \times \mathbf{z}_3 = 0$ , then

$$\begin{aligned} \max \{ \delta(\mathbf{A}_{(2,2,2)}) : |\mathbf{v}_1| = |\mathbf{v}_2| = |\mathbf{v}_3| = 1 \} \\ = (3\sqrt{3}/8) |\mathbf{z}_1 \times \mathbf{z}_2| \cdot |\mathbf{z}_3|. \end{aligned} \quad (103)$$

If either  $\mathbf{z}_1, \mathbf{z}_2$  are parallel vectors or some  $\mathbf{z}_i = 0$ , then the  $\hat{\mathbf{v}}_i$  are arbitrary unit vectors. Otherwise, if  $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$  are nonzero noncollinear vectors, and  $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$  is the right-handed orthonormal basis satisfying

$$\mathbf{f}_3 = \mathbf{z}_3/|\mathbf{z}_3|, \mathbf{f}_1 = \mathbf{z}_1 \times \mathbf{z}_2/|\mathbf{z}_1 \times \mathbf{z}_2|, \mathbf{f}_2 = \mathbf{f}_3 \times \mathbf{f}_1, \quad (104)$$

then  $\hat{\mathbf{v}}_1 = (\mathbf{f}_1 \sin 2\psi + \mathbf{f}_3 \cos 2\psi)$  and  $\hat{\mathbf{v}}_2, \hat{\mathbf{v}}_3$  are any orthogonal unit vectors satisfying

$$\hat{\mathbf{v}}_2 \times \hat{\mathbf{v}}_3 = \pm(\mathbf{f}_1 \sin \psi + \mathbf{f}_3 \cos \psi), \quad (105)$$

where  $\psi = \pm\pi/6$ .

*Proof*

The relation (94), for parallel  $\mathbf{z}_2, \mathbf{z}_3$ , implies

$$\delta(\mathbf{A}_{(2,2,2)}) = |(\mathbf{v}_1 \cdot \mathbf{v}_2 \times \mathbf{z}_3)(\mathbf{v}_1 \cdot \mathbf{z}_1 \times \mathbf{z}_2)(\mathbf{v}_2 \times \mathbf{v}_3 \cdot \mathbf{z}_3)| \quad (106)$$

by relation (4). Clearly  $\delta(\mathbf{A}_{(2,2,2)}) \equiv 0$  and the  $\hat{\mathbf{v}}_i$  are arbitrary unless  $\mathbf{z}_1 \times \mathbf{z}_2 \neq 0$  and  $\mathbf{z}_3 \neq 0$ . If  $\mathbf{v}_2 \times \mathbf{v}_3 = |\mathbf{v}_2 \times \mathbf{v}_3| \mathbf{v}_0$ , then  $|\hat{\mathbf{v}}_2 \times \hat{\mathbf{v}}_3|$  must be unity. Thus substitution yields

$$\max \delta(\mathbf{A}_{(2,2,2)}) = |\mathbf{z}_1 \times \mathbf{z}_2| \cdot |\mathbf{z}_3| \cdot |(\hat{\mathbf{v}}_0 \cdot \hat{\mathbf{v}}_1)(\hat{\mathbf{v}}_0 \cdot \mathbf{f}_3)(\hat{\mathbf{v}}_1 \cdot \mathbf{f}_1)| \quad (107)$$

and Lemma 6 shows

$$\begin{aligned} \max \delta(\mathbf{A}_{(2,2,2)}) \\ = (1/2) |\mathbf{z}_1 \times \mathbf{z}_2| \cdot |\mathbf{z}_3| \cdot |\hat{\mathbf{v}}_1 \cdot \mathbf{f}_1| \{1 + |\hat{\mathbf{v}}_1 \cdot \mathbf{f}_3|\}, \end{aligned} \quad (108)$$

where  $\hat{\mathbf{v}}_0$  is any principal bisector of  $\langle \mathbf{f}_3 \rangle, \langle \hat{\mathbf{v}}_1 \rangle$ . Clearly,  $\hat{\mathbf{v}}_1$  cannot involve  $\mathbf{f}_2$ , so that  $\hat{\mathbf{v}}_1$  may be assumed  $\mathbf{f}_1 \sin \theta + \mathbf{f}_3 \cos \theta$ ; and if  $\omega(\theta) = |\sin \theta| \cdot \{1 + |\cos \theta|\}$ , then

$$\max \delta(\mathbf{A}_{(2,2,2)}) = (1/2) |\mathbf{z}_1 \times \mathbf{z}_2| \cdot |\mathbf{z}_3| \cdot \omega(\theta). \quad (109)$$

However,  $\omega(\theta)$  is an even function with period  $\pi$ , whence its maxima on  $[-\pi, +\pi]$  have locations  $\pm 2\psi, \pm(\pi - 2\psi)$ . Moreover, we find

$$d\omega/d\theta = (\cos \theta + 1)(2 \cos \theta - 1) \text{ on } [0, \pi/2], \quad (110)$$

and we obtain  $2\psi = \pi/3$  from (110). Substitution produces (103).  $\square$

• **Remark**

An extra constraint in Theorem 9 provides some particular examples of solutions. Indeed, the identity  $\hat{\mathbf{v}}_2 \cdot \hat{\mathbf{v}}_0 = \hat{\mathbf{v}}_3 \cdot \hat{\mathbf{v}}_0 = 0$  and the requirement

$$\hat{\mathbf{v}}_2 \cdot \mathbf{f}_3 = \hat{\mathbf{v}}_3 \cdot \mathbf{f}_3 = \xi_3 \quad (111)$$

yield the consequence  $\hat{\mathbf{v}}_2 \cdot \mathbf{f}_1 = \hat{\mathbf{v}}_3 \cdot \mathbf{f}_1 = \xi_1$ ; while the normalization  $|\hat{\mathbf{v}}_2| = |\hat{\mathbf{v}}_3| = 1$  and the fact  $\hat{\mathbf{v}}_2 \neq \hat{\mathbf{v}}_3$  add the property  $\hat{\mathbf{v}}_2 \cdot \mathbf{f}_2 = -\hat{\mathbf{v}}_3 \cdot \mathbf{f}_2 = \xi_2$ . Also,  $0 = \hat{\mathbf{v}}_2 \cdot \hat{\mathbf{v}}_3$ , so that

$$0 = \xi_1^2 - \xi_2^2 + \xi_3^2, \quad (112)$$

and  $\hat{\mathbf{v}}_0 = \hat{\mathbf{v}}_2 \times \hat{\mathbf{v}}_3$ , so that

$$\begin{aligned} (\varepsilon_1/2)(\varepsilon_2 \mathbf{f}_1 - \sqrt{3} \mathbf{f}_3) = 2\xi_2(\xi_3 \mathbf{f}_1 - \xi_1 \mathbf{f}_3), \\ \varepsilon_1 = \pm 1, \varepsilon_2 = \pm 1. \end{aligned} \quad (113)$$

But the last two relations produce  $1 = |\hat{\mathbf{v}}_0|^2 = 4\xi_2^2(\xi_1^2 + \xi_3^2) = 4\xi_2^4$ , whence the individual (113) components imply

$$\begin{aligned} \hat{\mathbf{v}}_2 = (\varepsilon_3/2\sqrt{2})(\varepsilon_1\sqrt{3}\mathbf{f}_1 + 2\mathbf{f}_2 + \varepsilon_1\varepsilon_2\mathbf{f}_3), \\ \hat{\mathbf{v}}_3 = (\varepsilon_3/2\sqrt{2})(\varepsilon_1\sqrt{3}\mathbf{f}_1 - 2\mathbf{f}_2 + \varepsilon_1\varepsilon_2\mathbf{f}_3), \quad \varepsilon_3 = \pm 1. \end{aligned} \quad (114)$$

**11. Optimal layouts of tridirectional type**

Here we reconsider layouts with few off-axial instruments, and we determine optima of our tridirectional types. The less familiar (3, 2, 1) arrangements achieve the previous unconstrained maxima. The more intuitive (2, 2, 2) arrangements produce consistently inferior results.

Any (2, 2, 2) arrangement, by Section 9, partitions the six accelerometers into three pairs, where each pair have noncoincident locations, but their measurements have identical directions. If such a layout contains one off-axial instrument, then one pair couples this location with either the pen point or the far end, while two pairs contain accelerometers at both the pen point and the far end. The optimal directions, by Theorem 9, yield

$$\max \delta(\mathbf{A}_{(2,2,2)}) = (3\sqrt{3}/8)\alpha^2\beta \quad (115)$$

in either case. Moreover, each of the preceding (2, 2, 2) arrangements becomes a layout with two off-axial instruments when the end accelerometer of either axial pair assumes the ring position opposite the given off-axial instrument, and the hypothesized smallness of the ratio  $\beta/\alpha$  permits only a slight increase in the maximal  $\delta(\mathbf{A})$ .

An alternative (2, 2, 2) arrangement with two opposed ring locations would pair each ring accelerometer with a separate end instrument, and pair another end accelerometer with one point instrument. But then  $(3\sqrt{3}/8)\alpha\beta^2$  is the maximal  $\delta(\mathbf{A})$ , and even (115) exceeds this value. The final (2, 2, 2) arrangement with two opposed ring locations associates the off-axial accelerometers with each other, and proposes two distinct pairs along the pen axis. The optimal directions, by Theorem 9, yield

$$\max \delta(A_{(2,2,2)}) = (3\sqrt{3}/4)\alpha^2\beta \quad (116)$$

in this last case. Therefore the best (2, 2, 2) arrangements with one or two off-axial instruments are, respectively, the layouts with maximal values (115) or (116). But the accelerometers specify a plane, for all such layouts, and Theorem 9 prescribes the same optimal directions in all cases. If  $\langle e_2, e_3 \rangle$  is this plane, and  $\psi = \pm\pi/6$ , then the direction for the nonaxial pair is either unit vector of form  $\pm(e_1 \sin 2\psi + e_3 \cos 2\psi)$ , while the directions for the two axial pairs are any orthonormal vectors with cross-product  $\pm(e_1 \sin \psi + e_3 \cos \psi)$ . Our previous disclosure [12] exhibits the symmetric specialization (114).

Any (3, 2, 1) arrangement, by Section 9, includes three noncoincident accelerometers with a common direction; also, two other distinct instruments have the same direction, and a final accelerometer has its own direction. If such a tridirectional arrangement contains just one off-axial instrument, then the threefold group must occupy all three locations, while the distinct pair covers both axial extremes, and the singleton takes any remaining position. The optimal directions, by Theorem 7, yield

$$\max \delta(A_{(3,2,1)}) = \alpha^2\beta \quad (117)$$

in each case. Again we obtain corresponding layouts with two off-axial locations when we shift the end accelerometer of the threefold group to assume the ring position opposite the given off-axial instrument. The optimal directions, by Theorem 7, yield

$$\max \delta(A_{(3,2,1)}) = 2\alpha^2\beta \quad (118)$$

in these cases. Any other (3, 2, 1) arrangement with two opposed ring accelerometers can have a threefold group with only one off-axial instrument. However, the largest  $\delta(A)$  among such layouts is  $\alpha\beta(\alpha^2 + \beta^2)^{1/2}$  by Theorem 7, because the most distant pair, under these conditions, links the other ring location with the pen point. Therefore, the best (3, 2, 1) arrangements with one or two off-axial instruments are, respectively, the layouts with maximal values (117) or (118).

Previously, Section 8 achieved the same maxima with no tridirectional assumption; hence a design loses no efficiency through this constraint. If we compare arrangements with the same positions, then we observe

$$\max \delta(A_{(2,2,2)}) / \max \delta(A_{(3,2,1)}) = 3\sqrt{3}/8 \approx 0.64952 \quad (119)$$

in every case. Also the (3, 2, 1) locations specify a unique plane, and Theorem 7 prescribes the best measurements. If this plane is  $\langle e_2, e_3 \rangle$ , then the threefold direction is  $\pm e_1$ , while the twofold direction is  $\pm e_2$ , and the singleton is  $\pm e_3$ . Our previous disclosure [12] exhibits one such layout. Our present analysis justifies this prior announcement.

## Acknowledgments

I wish to thank Alan J. Hoffman for debating algebraic assumptions in this work, Robert H. Risch for directing my attention to Risler's paper, Ralph A. Willoughby for locating various results of linear algebra, and Philip S. Wolfe for providing several theorems from convex analysis.

## References

1. J. J. Denier van der Gon and J. Ph. Thuring, "The Guiding of Human Writing Movements," *Kybernetik* **2**, 145-148 (1965).
2. J. Vredenburg and W. G. Koster, "Analysis and Synthesis of Handwriting," *Philips Tech. Rev.* **32**, 73-78 (1971).
3. M. Yasuhara, "Steps Toward Handwriting Analysis and Recognition," *Visible Language* **5**, 229-248 (1971).
4. J. Sternberg, "Automated Signature Verification Using Handwriting Pressure," 1975 *WESCON Technical Papers*, No. 31/4, Los Angeles, CA, Sept. 1975.
5. M. Yasuhara and M. Oka, "Signature Verification Experiment Based on Nonlinear Time Alignment: A Feasibility Study," *IEEE Trans. Syst. Man Cybernetics SMC-7*, 212-216 (1977).
6. N. M. Herbst and J. F. Morrissey, "Signature Verification Method and Apparatus," U.S. Patent 3,983,535, 1976.
7. N. M. Herbst and C. N. Liu, "Automatic Signature Verification Based on Accelerometry," *IBM J. Res. Develop.* **21**, 245-253 (1977).
8. N. M. Herbst and C. N. Liu, "Automatic Verification of Signatures by Means of Acceleration Patterns," *Proceedings, IEEE Computer Society Conference on Pattern Recognition and Image Processing*, June 6-8, 1977, Troy, NY, IEEE Computer Society, Long Beach, CA, 1977, pp. 332-337.
9. C. N. Liu, N. M. Herbst, and N. J. Anthony, "Automatic Signature Verification: System Description and Field Test Results," *IEEE Trans. Syst. Man Cybernetics SMC-9*, 35-38 (1979).
10. J. S. Lew, "Kinematic Theory of Signature Verification Measurements," *Math. Biosciences* **48**, 25-51 (1980).
11. J. S. Lew, "Signature Motion Recovery with Incomplete Initial Data," *IBM Tech. Disclosure Bull.* **20**, 5441-5442 (1978).
12. J. S. Lew, "Designs of Instrumented Pens for Signature Verification," *IBM Tech. Disclosure Bull.* **20**, 819-822 (1977).
13. A. R. Schuler, A. Grammatikos, and K. A. Fegley, "Measuring Rotational Motion with Linear Accelerometers," *IEEE Trans. Aerospace Electron. Syst. AES-3*, 465-472 (1967).
14. J. R. W. Morris, "Accelerometry—A Technique for the Measurement of Human Body Movements," *J. Biomechanics* **6**, 729-736 (1973).
15. J. S. Lew, "Accelerometer Layouts for Signature Verification," *Research Report RC 7154*, IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 1978.
16. J. S. Lew, "Optimal Designs of Instrumented Pens for Signature Verification," *IBM Tech. Disclosure Bull.* **21**, 3415-3419 (1979).
17. *Inertial Guidance*, G. R. Pitman, Jr., Ed., John Wiley & Sons, Inc., New York, 1962.
18. J. J. Risler, "Le Théorème des Zéros en Géométries Algébrique Réelles," *Bull. Soc. Math. France* **104**, 729-736 (1973).
19. W. H. Greub, *Linear Algebra* (3rd ed.), Springer-Verlag, New York, 1967.
20. H. G. Goldstein, *Classical Mechanics*, Addison-Wesley Publishing Co., Inc., Reading, MA, 1953.
21. A. S. Wineman and A. C. Pipkin, "Material Symmetry Restrictions on Constitutive Equations," *Arch. Rational Mech. Anal.* **17**, 184-214 (1964).

22. J. S. Lew, "Material Symmetry Restrictions for Certain Locally Compact Symmetry Groups," *Ill. J. Math* **11**, 202-212 (1967).
23. H. Weyl, *The Classical Groups*, Princeton University Press, Princeton, NJ, 1946.
24. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, London, 1965.
25. B. L. van der Waerden, *Modern Algebra*, Vol. 1, Frederick Ungar Publishing Co., New York, 1949.
26. F. R. Gantmacher, *The Theory of Matrices*, Vol. 1, Chelsea Publishing Co., New York, 1959.
27. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
28. K. L. Nielsen and J. H. Vanlonkhuyzen, *Plane and Spherical Trigonometry*, Barnes & Noble, Inc., New York, 1944.

*Received August 3, 1979*

*The author is located at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.*