The perspective of tensor geometry is even more helpful in acquiring insight into the T2 model. The constraint implied by (25) or (26) is that μ lie in $\mathfrak{A} \otimes \mathfrak{B} \otimes R^K$. Other examples of the value of the geometric perspective could be cited, but perhaps these will suffice to make the point.

There is, I admit, some validity to the authors' stated reason for avoiding tensor terminology. Much of the extant literature on tensors is written either for physicists or for mathematicians, and neither slant is particularly well suited for the statistical applications at hand. There is a tendency for the treatments to be on the one hand too abstract and on the other too specialized because of the focus on tensor powers of R^N. The concepts of covariance and contravariance, for example, arise when a tensor is regarded as an abstraction whose numerical representation is the result of an arbitrarily chosen coordinate system. The chemometrician need not struggle to understand these concepts. The numbers in his or her arrays are real data and not just coordinates with respect to some arbitrary basis.

Contrary to the impression one might get from books, the theoretical underpinnings of basic tensor product geometry are not difficult to comprehend. A key idea is the tensor product mapping, of which the outer product is a canonical example. The feature which distinguishes a tensor product mapping from other bilinear functions with two vector arguments, for example, the inner product, is its preservation of linear independence. Thus, if the F_1 columns of A are linearly independent and the F_2 columns of B are linearly independent, the F_1F_2 outer products of columns of A with columns of B are linearly independent. The span of these F_1F_2 outer products is by definition the tensor product $\mathfrak{A} \otimes \mathfrak{B}$, where $\mathfrak{A}$ is the span of the columns of A and $\mathfrak{B}$ is the span of the columns of B.

This presentation omits some mopping up details that are required for a fully rigorous definition of the tensor product of vector spaces. The essential features of the concept, however, are there. Anyone who can follow the development in the previous paragraph has a handle on a set of conceptual tools that can provide a valuable framework for interpreting multilinear models. These conceptual tools require names. We can use the existing tensor terminology or invent new terms. Others may have added bells and whistles that we don't want or need, but the terminology of tensors exists for the tools we do need, and it has been around for years. I say we should use it.

Rejoinder

Sue Leurgans and Robert T. Ross

We thank the editors for securing the comments of Kroonenberg and of deLeeuw, both of whom have contributed to the development and application of multiway methods in psychometrics, and of Burdick, a statistician with chemistry collaborators. Much of the work on multilinear models is deeply embedded in subject matter, and many contributions have been made outside the single application (spectroscopy) we have emphasized here. We thank the discussants for adding their views of important contributions in a variety of areas.

Our reply consists of a section with comments on mathematical issues raised by all discussants and a section of specific responses to selected points raised by each discussant.

1. GENERAL COMMENTS

We think one reason for the increased interest in arrays in recent decades is that arrays are no longer theoretical abstractions, but can be defined and manipulated in many high-level languages. The statistical package S (Becker, Chambers and Wilks, 1988) is just one of the programs that permits array calculations. Besides changing the nature of the questions that are important, the ability to calculate easily enables one to check conjectures.

The mathematical theory of arrays that needs to be applied to these questions requires some features that are more general than most descriptions of tensors. The most basic requirements are that the number of levels in each way of the array be arbitrary and not necessarily equal and that the elements of the array not be required to satisfy any symmetry assumptions. The different ways of the array need to be treated symmetrically in mathematics, although different applications may treat the ways differently. The notation employed will vary with the purpose of the exposition; see our reply to Burdick below. Other examples of publications on array results are Knuth (1965) and Lickteig (1985) [from the Geladi (1989) paper cited by Kroonenberg] and Lickteig's references.

The basic question about arrays is how to approximate an array by a simpler array or how to decompose

an array into a sum of several simple arrays. For matrices, the singular value decomposition provides the answer. Therefore, some papers have focused on generalizations of singular value decompositions to 3-way arrays. Some of these papers explicitly state that there is no single extension that preserves all the properties of singular value decompositions of matrices. Indeed, several of the properties of the PARAFAC model demonstrate this: the fact that the nonlinear least-squares estimates of the PARAFAC model are not nested in F, the fact that the F vectors for each way are not orthogonal and the fact that a $2 \times 2 \times 2$ array can require 3 vectors to represent it. One corollary should be emphasized: the selection of a mathematical setting is more critical for optimization problems concerning 3-way arrays than for problems concerning matrices. For example, Okamoto and Kanazawa (1968) demonstrate that the "best" low-rank approximation to a single covariance matrix is the approximation provided by principal components whenever the quality of the approximation is measured by a function of the eigenvalues that is isotonic with respect to the usual partial order on $\Re^P$. [See Rao (1980) for related material.] Many norms have this characterization. Okamota and Kanazawa's proof relies on eigenvalue decompositions for symmetric matrices, but extends easily to general matrices if singular value decompositions replace eigenvalue decompositions. However, the different norms will provide different simultaneous low-rank approximations to two or more matrices.

2. SPECIFIC COMMENTS

2.1 deLeeuw

Jan deLeeuw raises the question of stability. There are many types of stability; we describe progress on leave-out-one calculations and on volumes of confidence regions, and describe when we have seen instability due to nonidentifiability.

Arboleda (1991) obtained formulas for the curvature arrays of multilinear models and applied them to obtain formulas to approximate changes in the parameters when any single data point is omitted. Since the corresponding formulas for linear models provide the exact change in the parameters, it is natural to hope that the approximation is good for multilinear models. In a $10 \times 12 \times 5$ example, she demonstrated that the (much more easily computed) 600 approximations are very close to the 600 corresponding estimates obtained by applying nonlinear least-squares to arrays missing one element. Her work suggests that some stability can be anticipated in practice.

One measure of statistical stability is the volume of the joint confidence regions for the parameters. The standard hope in nonlinear models is that the volume of the joint confidence region is well approximated by the reciprocal of the determinant of the information matrix. Abel (1991) has shown that the determinant of the $F(I + J + K - 2) \times F(I + J + K - 2)$ information matrix can be simplified to a function of determinants of three $F \times F$ matrices and one $F(3F - 2) \times F(3F - 2)$ matrix. He also evaluated determinants when some of the elements of the array are not measured. He compared a class of balanced designs with a class of unbalanced designs. The balanced designs ("checkerboard designs") are derived from unions of Latin squares. The unbalanced designs ("crosshatch designs") emphasize certain preselected levels for each way. For linear models and for most design criteria, balanced designs dominate unbalanced designs. Abel's examples imply that this heuristic principle fails for multilinear models. Neither the balanced class nor the unbalanced class dominates (in the determinant criterion). Moreover, some very incomplete designs have nearly the same information as the complete array.

We can also report some results from simulation studies in which we generated arrays with $F = 1$, 2 or 3 and fit (22) with $F' = 1$, 2 or 3. When the number of parameters fit is no larger than the number of parameters present ($F' \leq F$), parameter vectors from independent simulations cannot be distinguished on an overhead transparency. However, when more parameters were estimated than were present in the mean of the simulated array ($F' > F$), one or more vectors for each way were so unstable that the parameter vectors from independent simulations have no discernable similarity. This numerical observation reflects the nonidentifiability of α_f and β_f whenever $\gamma_f = 0$.

Knowledge of the mathematical requirements for identifiability makes it easier to monitor near-nonidentifiability in practice. For example, in Figure 8, the curves depicting the dependence of the tryptophan-15 and impurity components on quencher concentration are nearly collinear. Because the impurity is 100 times weaker than the tryptophans, this near indeterminancy probably has little effect on the parameters estimated for the two tryptophans. The near collinearity of the excitation spectra of the two tryptophans is more problematic, because this characteristic of the protein under study cannot be changed by removing the impurity.

2.2 Kroonenberg

Kroonenberg's survey of other literature is much appreciated. We add comments in only four areas: the speed of data collection in spectroscopy, the role of preprocessing in spectroscopy, the nature of constraints on the parameter matrices and an alternative hierarchy of models.

In his Section 3.1, Kroonenberg quotes Sanchez and Kowalski (1990), noting that there is "an abundance

of instruments that can automatically collect precise third-order data arrays in a short-time." While the spectroscopist can collect data far faster than the social scientist, it can still take a long time to fill all the elements of a 3- or 4-way array: too long for a biological specimen to remain intact. For example, at our typical data collection rate of 1 point each 5 seconds, it would take over 12 hours to collect the data to fill a $30 \times 30 \times 10$ array.

Certain kinds of preprocessing are so common that the spectroscopist hardly thinks about them. Like many chemical determinations, most spectroscopic determinations require two measurements, one of the specimen of interest and one of a blank. The blank is as much like the specimen of interest as possible, but it lacks the key ingredients that give the spectroscopic signals being studied. It is the specimen data array minus the blank data array to which the multilinear model is actually fitted. With fluorescence measurements, the most important reason for doing the analysis in this way is that a phenomenon known as Raman scattering gives a strong signal that does not follow a multilinear model.

In his Section 3.5, Kroonenberg alludes to the linearly constrained CANDECOMP of Carroll, Pruzansky and Kruskal (1980). A CANDELINC model for an N-way expected value array satisfying (27) is the assumption that the N parameter matrices $\boldsymbol{A}_n$, $n = 1, \ldots, N$ are not arbitrary $I_n \times F$ matrices, but that each column is a linear function of $S_n < I_n$ parameters. Written in equations, the assumption is that there are $I_n \times S_n$ known matrices $\boldsymbol{X}_n$, $n = 1, \ldots, N$ such that

$$\boldsymbol{A}_n = \boldsymbol{X}_n \tau_n, \qquad n = 1, \ldots, N.$$

[This equation is equation (3′) of Carroll, Pruzansky and Kruskal (1980).] The new parameter matrices $\tau_1, \ldots, \tau_N$ can be defined so that $\boldsymbol{X}_n$ has orthonormal columns, forcing $\boldsymbol{X}_N^{\dagger} = \boldsymbol{X}_n^T$. Carroll, Pruzansky and Kruskal (1980) prove that the nonlinear least-squares estimates of $\tau_1, \ldots, \tau_N$ (with suitable identifiability conventions) based on $\boldsymbol{Y}$ are exactly the same as the estimates based on $\boldsymbol{Y}^*$, where

$$\boldsymbol{Y}^* = \{\times_{n=1}^{N} \boldsymbol{X}_n^T\} *_N \boldsymbol{Y}.$$

Restating the proof in the terminology and notation of this paper, one of the key steps is recognition that the Associativity Proposition and the orthonormality of the columns of $\boldsymbol{X}_n$ imply that the expectation of $\boldsymbol{Y}^*$ is

$$\{\times_{n=1}^{N} \tau_n\} *_N \boldsymbol{I},$$

where $\boldsymbol{I}$ is the "diagonal" $F \times \cdots \times F$ identity array. The equation above is an unconstrained PARAFAC model for the smaller $S_1 \times \cdots \times S_N$ array $\boldsymbol{Y}^*$ with unconstrained parameter matrices $\tau_1, \ldots, \tau_n$.

We have not seen any examples in spectroscopy with linear constraints on the parameter matrices of an identifiable PARAFAC model, although interest is building in models in which vectors for one or more ways are a function of a small number of other parameters. For example, Lee et al. (1992) fit trilinear models in which the levels of the third way corresponded to different concentrations $(L_1, \ldots, L_K)$ of a ligand, a chemical that can bind to the fluorophore. Separate trilinear models were fit for 5 different ligands and 3 different temperatures. Functional forms for the chemistry dependence of two of the fluorophores present are hypothesized to be

$$\gamma_1[k] = \frac{\phi_1 \kappa_1}{L_k + \kappa_1}, \qquad \gamma_2[k] = \frac{\phi_2 L_k}{L_k + \kappa_2}.$$

The ϕ's and the κ's depend on the ligand and the temperature. For some ligands, κ_1 and κ_2 are not significantly different. The overall mean array μ is conditionally linear in the ϕ's given the $\boldsymbol{A}$, $\boldsymbol{B}$ and the κ's, so a CANDELINC step could be applied in an ALS estimation procedure. Given the unsatisfactory behavior of ALS algorithms and the ease with which the parametric submodels above can be fitted to the estimates from a model in which no structure is forced for the dependence on chemistry concentration, Lee et al. (1992) fit the equations above to the nonlinear least-squares estimates $\hat{\Gamma}$.

The use of parametric submodels is popular in time-resolved fluorescence spectroscopy, where it is often called *global analysis*. [See Beechem et al. (1991).] In the typical application, fluorescence intensity is measured using two independent variables, most commonly emission wavelength and time. A bilinear model is then fit to the data, with the dependence on one variable, usually time, fit with a specific parametric model for each component. The indeterminacy of a general bilinear model is usually avoided by using a sufficiently specific model, such as a sum of two negative exponentials.

Kroonenberg alludes to various hierarchies among multilinear models. These include PARAFAC as a T3 model with a diagonal core array. Implicit in such hierarchies are fixed choices of F or of $F_1, \ldots, F_N$. If the F's are changed, then T2 and T3 models can be viewed as constrained PARAFAC models. For example, the T2 model (25) can be written as a PARAFAC model (22) with F set equal to $F_1 F_2$ and special patterns for $\boldsymbol{A}$ and $\boldsymbol{B}$. The $I \times F_1 F_2$ matrix $\boldsymbol{A}$ will have only F_1 distinct columns; a similar condition will hold for $\boldsymbol{B}$. If energy transfer may occur, the T2 model (7) may be more appropriate than an unconstrained PARAFAC model. Thus, in some spectroscopy settings it is important to distinguish a T2 model from a less restrictive PARAFAC model.

2.3 Burdick

We thank Burdick for emphasizing the importance of tensors and for prompting us to focus our objections more sharply on tensor notation, not tensor theory.

We agree that a T3 model can be thought of as a tensor model for μ and that the geometric formulations clarifies the inevitability of the nonidentifiability of the parameters. A PARAFAC model is less usefully formulated as a tensor, as is evidenced by the fact that the conventional Einstein summation convention does not apply. This convention avoids summation signs by specifying that if an index occurs twice in an expression that index is to be summed over. This convention is exactly the convention of skipping integral signs, but leaving differentials:

$$\int_{x_1}\int_{x_2}\int_{x_3} f(x_1, x_1, x_3)\, dx_1 dx_2 dx_3$$

becomes

$$f(x_1, x_2, x_3)\, dx_1 dx_2 dx_3$$

instead of

$$\int f \text{ or } \int f d\lambda.$$

Indices never occur more than twice in conventional expressions (the need for care in the use of these conventions is obvious), so a PARAFAC model can only be written conventionally by using the T3 representation, complete with artificial "diagonal" core array.

3. CONCLUSION

We thank the discussants for augmenting our discussion of applications of multilinear models and of the mathematical and computational underpinnings. The effectiveness of these models in spectroscopy and in other setting suggests that multilinear models merit further statistical attention.

ADDITIONAL REFERENCES

Abel, R. B. (1991). Experimental design for multilinear models in spectrofluoroscopy. Ph.D. dissertation, Ohio State Univ.

Arabie, P., Carroll, J. D. and DeSarbo, W. S. (1987). *Three-Way Scaling and Clustering.* Sage, Beverly Hills.

Arboleda, M. L. (1991). Case deletion for multilinear models. Ph.D. dissertation, Ohio State Univ.

Becker, R. A., Chambers, J. M. and Wilks, A. R. (1988). *The New S Language.* Wadsworth & Brooks/Cole, Pacific Grove, Calif.

Beechem, J. M., Gratton, E., Ameloot, M., Knutson, J. R. and Brand, L. (1991). The global analysis of fluorescence intensity and anisotropy decay data: Second-generation theory and programs. In *Topics in Fluorescence Spectroscopy, Volume 2: Principles* (J. R. Lakowicz, ed.) 241-305. Plenum, New York.

Carroll, J. D. and Arabie, P. (1980). Multidimensional scaling. *Annual Review of Psychology* **31** 607-649.

Carroll, J. D. and Pruzansky, S. (1984). The CANDECOMP-CANDELINC family of models and methods for multidimensional data analysis. In *Research Methods for Multimode Data Analysis* (H. G. Law, C. W. Snyder Jr., J. A. Hattie and R. P. McDonald, eds.) 372-402. Praeger, New York.

Carroll, J. D., Pruzansky, S. and Kruskal, J. B. (1980). CANDELINC: A general approach to multidimensional analysis of many-way arrays with linear constraints on parameters. *Psychometrika* **43** 3-24.

Carroll, J. D., De Soete, G. and Pruzansky, S. (1989). Fitting of the latent class model via iteratively reweighted least squares CANDECOMP with nonnegativity constraints. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.) 123-130. North-Holland, Amsterdam.

d'Aubigny, G. and Polit, E. (1989). Some optimality properties of the generalization of n-way tables with specified matrices. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.) 39-52. North-Holland, Amsterdam.

deLeeuw, J. and Pruzansky, S. (1978). A new computational method to fit the weighted Euclidean distance model. *Psychometrika* **43** 479-490.

Denis, J. B. and Dhorne, T. (1989). Orthogonal tensor decomposition of 3-way tables. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.) 31-38. North-Holland, Amsterdam.

Faith, D. P., Minchin, P. and Belbin, L. (1987). Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* **69** 57-68.

Franc, A. (1989). Multiway matrices: Some algebraic results. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.) 19-30. North-Holland, Amsterdam.

Geladi, P. (1989). Analysis of multi-way (multi-mode) data. *Chemometrics and Intelligent Laboratory Systems* **7** 11-30.

Gifi, A. (1990). *Nonlinear Multivariate Analysis.* Wiley, Chicester.

Harshman, R. A. (1972). Determination and proof of minimum uniqueness conditions for PARAFAC1. *UCLA Working Papers in Phonetics* **22** 111-117.

Harshman, R. A. and DeSarbo, W. S. (1984). An application of PARAFAC to a small sample problem, demonstrating preprocessing, orthogonality constraints, and split-half diagnostic techniques. In *Research Methods for Multimode Data Analysis* (H. G. Law, C. W. Snyder Jr., J. A. Hattie and R. P. McDonald, eds.) 602-642. Praeger, New York.

Harshman, R. A. and Lundy, M. E. (1984a). The PARAFAC model for three-way factor analysis and multidimensional scaling. In *Research Methods for Multimode Data Analysis* (H. G. Law, C. W. Snyder Jr., J. A. Hattie and R. P. McDonald eds.) 122-215. Praeger, New York.

Harshman, R. A. and Lundy, M. E. (1984b). Data preprocessing and the extended PARAFAC model. In *Research Methods for Multimode Data Analysis* (H. G. Law, C. W. Snyder Jr., J. A. Hattie and R. P. McDonald, eds.) 216-284. Praeger, New York.

Harshman, R. A., Lundy, M. E. and Kruskal, J. B. (1985). Comparison of trilinear and quadrilinear methods: Strengths, weaknesses, and degeneracies. Paper presented at the annual meeting of the Classification Society, St. John, Newfoundland, Canada.

Hayashi, C. and Hayashi, F. (1987). A new algorithm to solve PARAFAC-model. *Behaviormetrika* **11** 49-60.

Hayashi, F., Yamaoka, K. and Terao, H. (1982). A comparative study between two algorithms of PARAFAC model [in Japanese]. *Japanese Journal of Behaviormetrika* **1** 47-62.

Hitchcock, F. L. (1927). The expression of a tensor or a polyadic as a sum of products. *Journal of Mathematics and Physics* **6** 164-189.

Kiers, H. A. L. (1988). Comparison of "Anglo-Saxon" and "French" three-mode methods. *Statist. Anal. Données* **13** 14-32.

KIERS, H. A. L. (1991). Hierarchical relations among three-way methods. *Psychometrika* **56** 449–470.

KIERS, H. A. L. and KRIJNEN, W. P. (1991). An efficient algorithm for PARAFAC of three-way data with large numbers of observation units. *Psychometrika* **56** 147–152.

KNUTH, D. E. (1965). Finite semifields and projective planes. *J. Algebra* **2** 182–217.

KRIJNEN, W. P. and TEN BERGE, J. M. F. (1991). Contrast-vrije oplossingen van het CANDECOMP/PARAFAC-model. [Contrast-free solutions of the CANDECOMP/PARAFAC-model.] *Kwantitatieve Methoden* **12** 87–96.

KROONENBERG, P. M. (1983). *Three-Mode Principal Component Analysis*. DSWO Press, Leiden, The Netherlands.

KROONENBERG, P. M. (1988). Three-mode analysis. *Encyclopedia of Statistical Sciences* **9** 231–236. Wiley, New York.

KROONENBERG, P. M. (1989). Singular value decompositions of interactions in three-way contingency tables. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.) 169–184. North-Holland, Amsterdam.

KRUSKAL, J. B. (1976). More factors than subjects, tests and treatments: An indeterminacy theorem for canonical decomposition and individual differences scaling. *Psychometrika* **41** 281–293.

KRUSKAL, J. B. (1984). Multilinear methods. In *Research Methods for Multimode Data Analysis* (H. G. Law, C. W. Snyder Jr., J. A. Hattie and R. P. McDonald, eds.) 36–72. Praeger, New York.

KRUSKAL, J. B. (1989). Rank, decomposition, and uniqueness for 3-way and N-way arrays. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.) 7–18. North-Holland, Amsterdam.

KRUSKAL, J. B., HARSHMAN, R. A. and LUNDY, M. E. (1989). How 3-MFA data can cause degenerate PARAFAC solutions, among other relationships. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.) 123–130. North-Holland, Amsterdam.

LAW, H. G., SNYDER, C. W., JR., HATTIE, J. A. and MCDONALD, R. P. (1984). *Research Methods for Multimode Data Analysis*. Praeger, New York.

LAWSON, C. L. and HANSON, R. J. (1974). *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs, NJ.

LEE, J., ROSS, R. T., THAMPI, S. and LEURGANS, S. E. (1992). Resolution of the properties of hydrogen-bonded tyrosine using trilinear analysis of fluorescence. *J. Physical Chemistry*. To appear.

LICKTEIG, T. (1985). Typical tensorial rank. *Linear Algebra Appl.* **69** 95–120.

LIPOVETSKII, S. S. (1984). Component analysis of tables with many inputs. *Industrial Laboratory* **50** 474–481. [Translated from the Russian, in *Zavodskaya Laboratoriya* (1984) **50** 52–57.]

LUNDY, M. E., HARSHMAN, R. A. and KRUSKAL, J. B. (1989). A two-stage procedure incorporating good features of both trilinear and quadrilinear models. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.) 123–130. North-Holland, Amsterdam.

MCDONALD, R. P. (1984). The invariant factors model for multimode data. In *Research Methods for Multimode Data Analysis* (H. G. Law, C. W. Snyder Jr., J. A. Hattie and R. P. McDonald, eds.) 285–307. Praeger, New York.

MORE, T. (1973). Axioms and theorems for a theory of arrays. *IBM J. Res. Develop.* **17** 135–165.

OKAMOTO, M. and KANAZAWA, M. (1968). Minimization of eigenvalues of a matrix and optimality of principal components. *Ann. Math. Statist.* **39** 859–863.

OLDENBURGER, R. (1934). Composition and rank of n-way matrices and multilinear forms. *Ann. of Math.* **35** 622–657.

RAO, C. R. (1980). Matrix approximations and reduction of dimensionality in multivariate statistical analysis. In *Multivariate Analysis V* (P. R. Krishnaiah, ed.) 3–22. North-Holland, Amsterdam.

RAY, W. J. and COLE, H. W. (1985). EEG Alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive process. *Science* **228** 750–752.

SCHÖNEMANN, P. H. (1972). An algebraic solution for a class of subjective metric models. *Psychometrika* **37** 441–451.

SMILDE, A. K. (1992). Three-way analyses: Problems and prospects. *Chemometrics and Intelligent Laboratory Systems* 10.

SMILDE, A. K., VAN DER GRAAF, P. H., DOORNBOS, D. A., STEERNEMAN, T. and SLEURINK, A. (1990). Multivariate calibration of reversed-phase chromatographic systems, some designs based on three-way data analysis. *Analytica Chimica Acta* **235** 41–51.

SNYDER JR., C. W., LAW, H. G. and HATTIE, J. A. (1984). Overview of multimode analytic methods. In *Research Methods for Multimode Data Analysis* (H. G. Law, C. W. Snyder Jr., J. A. Hattie and R. P. McDonald, eds.) 2–35. Praeger, New York.

STRASSEN, V. (1968). Gaussian elimination is not optimal. *Numer. Math.* **13** 354–356.

TAKANE, Y., YOUNG, F. W. and DE LEEUW, J. (1977). Nonmetric individual difference multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika* **42** 7–67.

TEN BERGE, J. M. F. (1989). Convergence of PARAFAC preprocessing procedures and the Deming-Stephan method of iterative proportional fitting. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.) 53–64. North-Holland, Amsterdam.

TEN BERGE, J. M. F. (1991). Kruskal's polynomial for $2 \times 2 \times 2$ arrays and a generalization to $2 \times n \times n$ arrays. *Psychometrika* **56** 631–636.

TEN BERGE, J. M. F. and KIERS, H. A. L. (1989). Convergence properties of an iterative procedure of ipsatizing and standardizing a data matrix, with applications to PARAFAC/CANDECOMP preprocessing. *Psychometrika* **54** 231–236.

TEN BERGE, J. M. F., KIERS, H. A. L. and DE LEEUW, J. (1988). Explicit CANDECOMP/PARAFAC solutions for a contrived $2 \times 2 \times 2$ array of rank three. *Psychometrika* **53** 579–584.

TUCKER, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika* **31** 279–311.

YOSHIZAWA, T. (1987). Singular value decomposition of multi-array data and its applications. Technical Report CS-87-038, Dept. Computer Science, Yamanashi Univ.