with operator algebras and measured equivalence relations, but it is valuable to have them spelled out for a larger audience.

3. The study of multidimensional Markov shifts, while still at an early stage, is bound to grow in importance, partly because of links with other active areas (sample buzzwords: tiling, percolation). The examples in the last three chapters, especially in chapter 5, give a wonderfully interesting first view of the subject.

My one complaint about the book is that the system of cross-referencing is unnecessarily confusing; for example, "(1.5)" has more than one meaning.

J. FELDMAN
UNIVERSITY OF CALIFORNIA, BERKELEY

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 27, Number 1, July 1992

0273-0979/92 $1.00 + $.25 per page

Topics in matrix analysis, by Roger A. Horn and Charles R. Johnson. Cambridge Univ. Press, 1991, viii + 607 pp. $59.50. ISBN 0-521-30587-X

In his comprehensive historical study [10] Kline wrote:

> Though determinants and matrices received a great deal of attention in the nineteenth century and thousands of papers were written on these subjects, they do not constitute great innovations in mathematics... Neither determinants nor matrices have influenced deeply the course of mathematics despite their utility as compact expressions and despite the suggestiveness of matrices as concrete groups for the discernment of general theorems of group theory...

Despite these sentiments, doubtlessly shared by a substantial part of the mathematical public, interesting, difficult, and important work on matrix theory continues to appear at an accelerating pace. As evidence of this, within the last year or so, Brualdi and Ryser published *Combinatorial matrix theory* [5], *Abstract linear algebra* by Curtis was posthumously published [6], and the second volume of the Horn and Johnson work (H & J), the subject of this review, made its long awaited appearance.

The book literature in matrix theory exploded in the sixties and early seventies with literally dozens of rather pedestrian efforts. This deluge occurred partly in response to NSF educational initiatives that dictated a new undergraduate curriculum in which "linear algebra" and "finite math" were to be introduced at the earliest possible moment. As might have been predicted, the logical outcome in the eighties was the inclusion of elementary matrix theory in ponderous calculus books already too heavy to lift unaided. Nonetheless, significant books on matrices and their mathematical applications have appeared irregularly over the last fifty years. Bourbaki's *Algebra* [4] commits 476 pages to linear and multilinear algebra, albeit at a predictably rarefied level. Even so, Bourbaki devotes space to some very old fashioned (and hard) matrix/determinant problems; e.g., the evaluation of the Cauchy determinant $\det(a_i + b_j)^{-1}$; the Sylvester-Franke theorem, conditions for a p-vector to be pure (i.e., decomposable) in $\bigwedge^p V$, the pth exterior space over V.

Bourbaki does not share the views expressed in the Kline quotation (from [4, p. 655]):

> Linear algebra is both one of the oldest and one of the newest branches of mathematics. On the other hand, at the origins of mathematics are the problems which are solved by a single multiplication or division, that is by calculating a value of a function $f(x) = ax$, or by solving an equation $ax = b$: these are typical problems of linear algebra and it is impossible to deal with them, indeed even to pose them correctly, without "thinking linearly."
>
> On the other hand, not only these questions but almost everything concerning equations of the first degree had long been relegated to elementary teaching, when the modern development of the notions of field, ring, topological vector space, etc. came to isolate and emphasize the essential notions of linear algebra (for example duality); then the essentially linear character of almost the whole of modern mathematics was perceived, of which "linearization" is itself one of the distinguishing traits, and linear algebra was given the place it merits.

Thirty one years ago Richard Bellman wrote *Introduction to Matrix Analysis* [3], which more or less defined the term "matrix analysis." It is interesting to compare the contents of [3] and H & J. The two books have chapters on functions of matrices, stability and the Lyapunov theory, eigenvalue inequalities and Kronecker products. If the H & J predecessor volume [9] is included then [3] and H & J share additional chapters on hermitian matrices and matrices with nonnegative entries. Both Bellman and H & J define matrix analysis as those parts of linear algebra that either use or address problems in mathematical analysis. Of course, this definition is a bit fuzzy and does not take into account items such as the proof, using compound matrices, of the famous Weyl majorization inequalities relating eigenvalues λ_j and singular values α_j of an n-square nonsingular A [11]

$$\sum_{j=1}^{k} \log|\lambda_j| \le \sum_{j=1}^{k} \log|\alpha_j|, \quad k = 1, \ldots, n \tag{1}$$

with equality for $k = n$, $|\lambda_j| \ge |\lambda_{j+1}|$, and $\alpha_j \ge \alpha_{j+1}$. In fact, the equivalent

$$\prod_{j=1}^{k} |\lambda_j| \le \prod_{j=1}^{k} \alpha_j, \qquad k = 1, \ldots, n$$

appears as an exercise in [1] where it is credited to E. T. Browne in a paper published in 1928.

The first chapter in H & J is entitled "The field of values," also called the "numerical range," and is denoted by $W(A)$. This set is the image in $\mathbb{C}$ of the surface of the unit sphere in $\mathbb{C}^n$ under the mapping $x \to x^*Ax$. It is obvious that $W(A)$ is compact. What is considerably less obvious is the fact that $W(A)$

is also convex, the content of the famous Toeplitz-Hausdorff theorem. The standard proof shows that $W(A)$ is convex for any 2×2 matrix A and then reduces the general $n \times n$ case to $n = 2$. In a book important to the matrix crowd, Halmos [7] wrote of the Toeplitz-Hausdorff theorem:

> Every known proof is computational... A conceptual proof would be desirable even (or especially?) if the concepts it uses are less elementary than the ones in a computational proof.

Since 1967, a number of proofs have appeared, some of which purport to satisfy Halmos's request. Whether they do or do not is disputable. In any event, the numerical range has been generalized in several ways, e.g.,

$$W_m(A) = \{\det(X^*AX) | \det(X^*X) = 1\}$$

where A is $n \times n$, X is $n \times m$, and $m \le n$. Clearly $W_1(A) = W(A)$ is convex, but $W_2(A)$ is not convex (try $A = \mathrm{diag}(1, 1, i, i)$). Nevertheless, $W_m(A)$ does have some interesting properties. For example: it contains all possible products taken m at a time of the eigenvalues of A; the variable matrix X can be restricted to partial isometries ($X^*X = I_m$) without altering $W_m(A)$. In the 88 pages that H & J devote to the field of values, some of these generalizations and their present status are discussed.

Every engineering student learns early in his training that the behavior of solutions to the linear system of ordinary differential equations $\dot{x} = Ax + b(t)$ depends on the distribution of the eigenvalues of A. Nonlinear systems $\dot{x} = f(x, t)$ can frequently be studied by expanding f in a Taylor series about a singular point x_0, discarding the nonlinear terms in x, and examining the resulting linear system in which the coefficient matrix A is a Jacobian matrix. The entire apparatus of inertia and Lyapunov stability was created in order to find computationally feasible methods for deciding about the location of the eigenvalues of a general complex matrix. The original Lyapunov theorem is given a careful proof in Chapter 2 of H & J. It is an important result because it reduces the unpleasant issue of estimating eigenvalues to the more tractable problem of solving a linear matrix equation. There are interesting treatments in this chapter of classes of stable matrices, the M-matrices and P-matrices, and the extent to which such classical inequalities as the Fischer inequality apply to them.

The third chapter of H & J is entitled "Singular value inequalities" and it is filled with interesting results relating norms, majorization, unitary invariance, eigenvalues, and singular values. The authors quite rightly identify K. Fan, A. Horn, G. Pólya, and H. Weyl as originators of the principal results in this highly important area of matrix analysis. To give some of the flavor of this work, consider a normal matrix A with eigenvalues $\lambda_1, \ldots, \lambda_n$ and corresponding orthonormal (o.n.) eigenvectors $u_1, \ldots, u_n$. Let $x_1, \ldots, x_k$ be any k o.n. vectors with $x_{k+1}, \ldots, x_n$ a completion to an o.n. basis. It is simple to confirm that the matrix $S_{j,t} = [|u_t^* x_j|^2]$ is doubly stochastic and that

$$x_j^* A x_j = S_{(j)}\lambda, \qquad j = 1, \ldots, k,$$

where $S_{(j)}$ is the jth row of S and λ is the column n-tuple of eigenvalues of

A. Thus

$$\sum_{j=1}^{k} x_j^* A x_j = \sum_{j=1}^{k} S_j \lambda. \tag{2}$$

Birkhoff's theorem in [11] implies that S is in the convex hull of the set of n-square permutation matrices P_σ and hence $\sum_{j=1}^{k} S_{(j)}\lambda$ is in the convex hull of the $n!$ numbers $\lambda_{\sigma(1)} + \cdots + \lambda_{\sigma(k)}$, in which σ runs over S_n, the symmetric group of degree n. In Fan's 1949 result [11], A is hermitian and thus from (2) one can conclude that

$$\min_{\sigma \in S_n} \lambda_{\sigma(1)} + \cdots + \lambda_{\sigma(k)} \leq \sum_{j=1}^{k} x_j^* A x_j \leq \max_{\sigma \in S_n} \lambda_{\sigma(1)} + \cdots + \lambda_{\sigma(k)}. \tag{3}$$

Fan's inequality (3) was the starting point of subsequent variational characterizations of singular values that has resulted in a vast literature on this subject over the last forty years. An excellent exposition of this theory may also be found in a recent book by Amir-Moéz [2].

If U_1, U_2, V_1, and V_2 are vector spaces over a common field and $T_i\colon U_i \to V_i$, $i = 1, 2$, are linear then $T_1 \otimes T_2\colon U_1 \otimes U_2 \to V_1 \otimes V_2$, the tensor product of T_1 and T_2, is the linear map that satisfies

$$T_1 \otimes T_2 u_1 \otimes u_2 = T_1 u_1 \otimes T_2 u_2$$

for all $u_i \in U_i$, $i = 1, 2$. With natural choices of bases, the matrix representation of $T_1 \otimes T_2$ is a Kronecker product of matrices $A \otimes B$. Since the space of $m \times n$ matrices can be regarded as a tensor product of two spaces of tuples, it easily follows that the linear map $X \to AXB$ has a matrix representation $B^{\mathrm{T}} \otimes A$. In Chapter 4 of H & J, "Matrix equations and the Kronecker product," it is shown how these simple observations have been profitably applied to study the Lyapunov mapping $X \to XA + A^* X$, the commutator mapping $X \to AX - XA$, and general matrix equations $AX + XB = C$. As H & J observe, there has been much work done to characterize linear mappings on matrix spaces that preserve a stipulated property. This "linear preserver"problem frequently has the following general form: Let T be a linear transformation defined on a vector space V. Let U be some subset of V. What are necessary and sufficient conditions (n.a.s.c.) on T such that U is preserved, i.e., $T(U) \subset U$? Sometimes this problem is stated in terms of an invariant f defined on U; that is, what are n.a.s.c. on T such that for each $u \in U$, $f(Tu) = f(u)$?

This invariance problem has been of interest to mathematicians since the last century, and there is vast literature on the subject. In fact, it is difficult to limit sensibly the extent of a comprehensive bibliography. Stephen Pierce at San Diego State University is assembling such a bibliography in connection with the forthcoming publication of an extensive survey on linear preservers in the journal *Linear and Multilinear Algebra.* In many cases the invariance problem specializes to $V = M_{m,n}(F)$. In other words, it is required to determine n.a.s.c. on a linear T,

$$T\colon M_{m,n}(F) \to M_{m,n}(F),$$

such that $T(U) \subset U$ for some appropriate set U, or possibly $f(T(A)) = f(A)$ for all $A \in U$. For example, take $V = M_{m,n}(\mathbb{C})$ and U to be the set of partial isometries, i.e., determine T such that whenever $A^*A = I_n$ it follows that $T(A)^*T(A) = I_n$. Frequently such problems can be reduced to the study of those T that preserve rank, i.e., $U = R_k$ is the totality of matrices of rank k:

$$\operatorname{rank} A = k \quad \text{implies} \quad \operatorname{rank} T(A) = k.$$

In turn, this question leads to a study of the structure of subspaces of $M_{m,n}(F)$ whose nonzero elements are all of rank k. R. Westwick and H. Flanders, among many others, have done work on this question. There are several connections of the invariance problem with Kronecker products but probably the simplest can be described as follows. As noted earlier, the space of matrices $M_{m,n}(F)$ is a tensor product

$$M_{m,n}(F) = M_{1,m}(F) \otimes M_{1,n}(F)$$

in which the tensor map is the dyad map $x \otimes y = x^{\mathrm{T}}y$. It is not difficult to confirm that A has rank k iff

$$A = \sum_{t=1}^{k} x_t \otimes y_t \tag{4}$$

in which $x_1, \dots, x_k$ as well as $y_1, \dots, y_k$ are linearly independent. Thus a linear T preserves rank k iff T sends every tensor $\sum_{t=1}^{k} x_t \otimes y_t$ of irreducible length k into another tensor of irreducible length k. A representation such as (4) is said to have irreducible length k if k is minimal over all possible such representations.

Finally, it is important to mention that there is a considerable amount of current activity related to the isometry-preserver problem mentioned above. This issue is: characterize those linear T,

$$T\colon M_{m,n}(\mathbb{C}) \to M_{m,n}(\mathbb{C}),$$

that preserve various functions of the singular values.

The concept of the Hadamard product of two conformal matrices has been in the literature beginning with the memorable 1911 paper by Schur [12]. In that paper Schur proved that the entrywise product (Hadamard product) $A \cdot B$ of two positive definite n-square hermitian matrices is also positive definite. Actually, $A \cdot B$ is a principal submatrix of the Kronecker product $A \otimes B$, from which the positive definite property immediately follows. Apparently, the name "Hadamard product" is due to J. von Neumann who used it in his lectures at the Institute for Advanced Study in the early 1940s. In a recent paper [8], Horn wrote

> von Neumann's usage may have been the result of the long-lasting influence of a famous 1899 paper in which Hadamard studied two Maclaurin series $f(z) = \sum a_n z^n$ and $g(z) = \sum b_n z^n$ with positive radii of convergence and their composition $h(z) = \sum a_n b_n z^n$, obtained as a coefficient-wise product. He showed that the function $h(\cdot)$ can be obtained from $f(\cdot)$ and $g(\cdot)$ by an integral convolution, and he proved that any singularity z_1 of $h(\cdot)$ must be of the form $z_1 = z_2 z_3$, where z_2 and z_3 are singularities of $f(\cdot)$ and $g(\cdot)$, respectively. Even though Hadamard never mentioned entry-wise products of matrices in

> this paper, the importance of his product theorem for power series and his mathematical eminence (he had proved the Prime Number Theorem in 1896) gave analysts a reason to associate Hadamard's name with term-by-term products of all kinds.

Probably the most recent expository material on the Hadamard product appears in Chapter 5 of H & J and in [9]. The authors exhibit various analytical contexts in which the Hadamard product arises. For example, consider the integral operator

$$K(f) = \int_a^b K(x, y) f(y)\, dy. \tag{5}$$

If the kernel $K(x, y)$ is replaced by the pointwise product of two kernels, $K(x, y)H(x, y)$, then (5) becomes

$$L(f) = \int_a^b K(x, y) H(x, y) f(y)\, dy. \tag{6}$$

In (6) the linear map $f \to L(f)$ can be viewed as a limit of matrix-vector multiplications. If $K(x, y)$ and $H(x, y)$ are continuous positive semidefinite kernels in the sense that

$$(K(f), f) = \int_a^b \int_a^b K(x, y) f(x) \overline{f(y)}\, dx\, dy \geq 0,$$

then Schur's result implies that L is also positive semidefinite.

There are some quite unexpected consequences of a few simple (and nonobvious) observations about the Hadamard product. For example, if

$$B = A \operatorname{diag}(\lambda_1, \ldots, \lambda_n) A^{-1}$$

then

$$d_B = A \cdot (A^{-1})^T \lambda \tag{7}$$

in which d_B is the n-tuple of main diagonal entries of B and $\lambda = (\lambda_1, \ldots, \lambda_n)^{\mathrm{T}}$. If B is normal, so that A can be taken to be unitary, then (7) shows that $d_B = S\lambda$ where $S = A \cdot \overline{A}$ is doubly stochastic. As we saw before, this relation leads immediately to the important majorization inequalities of K. Fan, G. Pólya, A. Horn, and H. Weyl.

The final chapter, "Matrices and functions," is 178 pages in length and is frequently very heavy going indeed. Things start out innocuously enough with the standard results about scalar polynomials in a square matrix. The Cayley-Hamilton theorem implies that any scalar polynomial $f(A)$ in an n-square matrix A can be expressed in terms of a polynomial p of degree less than n. The problem of finding p is, as the authors observe,

> ... a special case of an interpolation problem that arises again so it is worthwhile to examine it carefully.

They then launch into a full scale presentation of the Lagrange-Hermite interpolation problem, the Newton divided difference formulas, and their representation using the Cauchy integral theorem. Under certain circumstances, a nonpolynomial function f, defined on a domain $D \subset \mathbb{C}$ that includes the eigenvalues of A, can be used to define a matrix function $f(A)$. If $(\lambda-\alpha)^{\mu}$ is the highest degree elementary divisor of $\lambda I - A$ involving α then the derivatives $f^{(j)}(\alpha)$, $j = 0, \ldots, \mu - 1$ must exist. If these conditions obtain for all eigenvalues α then a coherent definition of $f(A)$ can be made that is independent of the similarity that brings A to Jordan normal form. The latter part of this chapter is devoted to the chain rule for differentiating a matrix function. For example, if A and B are positive definite Hermitian, how is $(d/dt)(tA + (1-t)B)^{1/2}$ effectively evaluated for $0 < t < 1$? The book ends with an outline of the Loewner theory of monotone matrix functions.

There is no doubt that this two volume work, *Matrix Analysis* and *Topics in Matrix Analysis*, is an important and unique contribution to the contemporary book literature on matrix theory. The authors unify and organize an extensive research literature that has developed since the end of World War II. The mathematics in H & J is typical of this field in being difficult and very specific. As such, it requires considerable concentration from the reader. It has been observed elsewhere that matrix analysis is not a spectator sport. Certainly the reader of this book will not go very far without tackling some of the nonroutine problems posed at the end of every section. Mercifully, the authors have included a (somewhat skimpy) "Hints" section at the end of the book.

In his preface to [3], Bellman listed what he called "the many fundamental aspects of matrix theory" that were not discussed in his book. These were: a computational treatment of matrices; the combinatorial theory of matrices, topological aspects of matrix theory, group representations, ideal theory by way of matrices due to Poincaré, and integer matrices. Finally, Bellman wrote:

> On the distant horizon, we foresee a volume on the advanced theory of matrix analysis. This would contain, among other results, various aspects of the theory of functions of matrices, the Loewner theory, the Siegel theory of modular functions of matrices, and the R-matrices of Wigner. In the more general theory of functionals of matrices, the Baker-Campbell-Hausdorff theory leads to the study of product integrals. These theories have assumed dominant roles in many parts of mathematical physics.

H & J meets some, but by no means all, of Bellman's specifications. There is plenty of opportunity for the next generation to fill in the gaps.

References

1. A. C. Aitken, *Determinants and matrices*, 9th edition, Oliver and Boyd, Edinburgh, 1956.
2. Ali R. Amir-Moéz, *Extreme properties of linear transformations*, Polygonal, Washington, NJ, 1990.
3. Richard Bellman, *Introduction to matrix analysis*, (1st and 2nd editions), McGraw-Hill, New York, 1960 and 1970.
4. Nicholas Bourbaki, *Elements of mathematics, Algebra* I, chapters 1–3, Addison-Wesley, Reading, MA, 1974.
5. Richard A. Brualdi and Herbert J. Ryser, *Combinatorial matrix theory*, Encyclopedia Math. Appl., Cambridge Univ. Press, New York, 1991.
6. Morton L. Curtis, *Abstract linear algebra*, Springer-Verlag, New York, 1990.
7. Paul R. Halmos, *A Hilbert space problem book*, American Book, van Nostrand, Reinhold, NY, 1967.

8. Roger A. Horn, *The Hadamard product*, Proc. Sympos. Appl. Math. (Charles R. Johnson, ed.), vol. 40, Amer. Math. Soc., Providence, RI, 1990.
9. Roger A. Horn and Charles R. Johnson, *Matrix analysis*, Cambridge Univ. Press, New York, 1985.
10. Morris Kline, *Mathematical thought from ancient to modern times*, Oxford Univ. Press, New York, 1972.
11. Marvin Marcus and Henryk Minc, *A survey of matrix theory and matrix inequalities*, Allyn and Bacon, Boston, MA, 1964.
12. Issai Schur, *Bemerkungen zur Theorie der beshränkten Bilinearformen mit unendlich vielen Veränderlichen*, J. Reine Angew. Math. **140** (1911), 1–28.

Marvin Marcus
University of California at Santa Barbara

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 27, Number 1, July 1992

0273-0979/92 $1.00 + $.25 per page

Discrete subgroups of semisimple Lie groups, by G. A. Margulis. Springer-Verlag, New York, 1991, 387 pp., $85.00. ISBN 3-540-17179-X

The subject of this remarkable book is, as its title indicates, discrete subgroups of semisimple Lie groups. Before describing these objects, it will be useful to recall some basic ideas about Lie groups themselves.

Lie groups arise in a wide variety of situations in geometry and algebra. Being by definition those groups that admit a compatible manifold structure, they arise in geometry as, roughly speaking, the finite-dimensional transformation groups of manifolds or, somewhat more precisely, as the transformation groups of manifolds that can be given locally by finitely many real parameters. While the full diffeomorphism group of a manifold is too large to be finite dimensional with respect to natural topologies, there are many situations where one encounters subgroups that are finite-dimensional Lie groups. One of the most important such geometric situations is that of the isometry group of a Riemannian manifold. In this case it is a classical result of Myers and Steenrod that the isometry group is always a Lie group. The same is true for the symmetry group of certain other classes of geometric structures, e.g., pseudo-Riemannian manifolds, and conformal structures in dimensions at least 3. While there are many natural geometric structures for which the full symmetry group is not necessarily finite dimensional (e.g., volume forms, symplectic structures, complex structures) it is of interest in these cases to understand the finite-dimensional symmetry groups and to understand conditions under which the full symmetry group will be finite dimensional. In an algebraic setting Lie groups arise in a similar manner. The general linear group of a real or complex finite-dimensional vector space is a Lie group. (Of course one can consider this as simply a further example of the symmetry group of a structure on a manifold, namely, a vector space structure.) Closed subgroups will also be Lie groups and, in particular, a subgroup that is the stablizer of any one of the natural objects associated to a vector space will be a Lie group. For example, this is the case for the stabilizer of a tensor, a subspace, a flag, etc. While it is not true that every Lie group (even a connected one) is isomorphic to a linear group (i.e., a subgroup of some