High, Low, and Quantitative Roads in Linear Algebra

Robert C. Thompson
Department of Mathematics
University of California
Santa Barbara, California 93106

Submitted by David H. Carlson

ABSTRACT

The future of core linear algebra is studied, with attention to advanced tools, elementary devices, and the computer.

1. INTRODUCTION

When I was asked to discuss the future of the pure aspects of linear algebra, I accepted without much thought. Only later did I realize that this assignment exceeded my mathematical and expository capabilities. However, I am willing to try anything once, so this will be my attempt to say something about the future of some of the pure aspects of our subject.

In the first SIAM linear algebra conference [1], in 1983, I was asked to lecture on *core* linear algebra. I think my instructions now are similar, to discuss the *core*, but with special attention to its future. I did not know then what the *core* was, and I still do not know. Occasionally I hear some linear algebraists, especially some with an applied focus, speak of the *core*, so presumably the term has a meaning. The meaning I adopt here is that the core of linear algebra is whatever is reviewed in the 15 ("Linear and Multilinear Algebra; Matrix Theory") section of *Mathematical Reviews* [2]. On this basis I will attempt a quantitative prediction of the future of core linear algebra.

However, the principal thrust here will be that the core is blended in a significant way with many of the other classifications in *Mathematical Reviews*, and this will lead to a qualitative prediction. I hope to establish it by adopting a technique sometimes used outside the core but seldom within it: proof by example.

655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$5.00

2. A QUANTITATIVE PREDICTION

An important trend is the increasing availability of good and easily used quantitative tools. This means good and inexpensive computers and good and inexpensive software to run on these wonderful machines. An example is Matlab, in its implementation on personal computers, and Derive, Macsyma, Maple, Mathematica, Milo, and Theorist are some other examples. Matlab provides a beautifully easy command structure for computations involving linear equations, eigenvalues, and singular values, over real and complex scalars, with the hazards from finite precision computer approximations for real numbers kept to a minimum. A future release will include sparse matrix algorithms. Exact integer and multivariable symbolic computations are possible in Derive, Macsyma, Mathematica, Maple, Milo, and Theorist, permitting the testing of conjectures with a ring theory structure. Another useful tool is Galois (for calculations over finite fields), and yet another, for complex variable computations, is appropriately named f(z). Similar software packages are certain to be increasingly available. Conjectures can now be numerically tested at a burden often trivial in time and money, a major advance. Any linear algebraist not using these powerful vehicles to uncover the secrets hidden in matrices is surely at a competitive disadvantage relative to his/her linear brothers/sisters.

Armed with this power, I decided to predict the future numerically. I counted the number of papers reviewed in the 15 (linear algebra) section of *Mathematical Reviews* year by year, back to its beginning in 1940. This data is shown in Figure 1 as the (solid) piecewise linear curve connecting the points $(y, N(y))$, where y is the year and $N(y)$ the number of linear algebra (classification 15) items in *Mathematical Reviews* in year y. There is substantial statistical irregularity, but an underlying pattern seems visible. The straight line fitting the data with minimum l_2 residual is shown using dots. Its equation is

$$N = 7.5694598(y - 1940) + 6.148235.$$

If a parabola is fitted to the data, the coefficient of y^2 is quite small, suggesting that not much curvature is present. However, N must go to 0 as y goes to $-\infty$, so a linear or parabolic fit is conceptually unhappy, and more natural curve to fit is an exponential. The one with minimum l_2 residual is

$$N = 63.1867356\, e^{0.0392730477(y-1940)}.$$

This curve is shown in Figure 1 with broken dots. It appears to be a quite

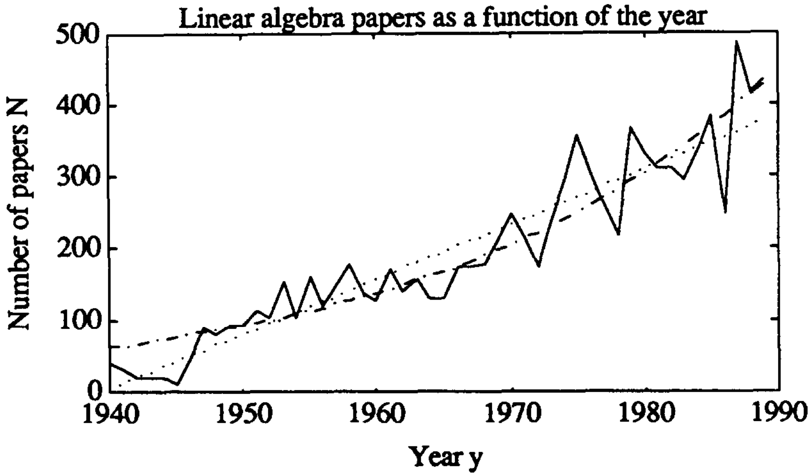

FIG. 1.

natural fit to the data, and on this basis one may infer that linear algebra is growing exponentially. Setting $y = 2000$ shows that

666 *is the estimated number of linear algebra papers in the year* 2000.

(It was 428 in 1989, and the straight line fit predicts only about 460 papers for $y = 2000$, surely unrealistically low.) This is our quantitative prediction of the future, and it explains why three linear algebra journals are thriving. Remember, for this prediction, that a paper belongs to linear algebra and year 2000 only if it is reviewed in the 15 section of Volume 100 of *Mathematical Reviews*. Of course, a confidence interval for this prediction should be stated.[1]

3. HIGH AND LOW ROADS

Some of the papers in the linear algebra data base achieve their objectives using powerful, advanced tools. Many, however, use only elementary techniques, relying instead on skill and strategy. These are the high and low roads

[1] The linear and exponential fits, and the graphics, were found using Matlab on a Macintosh Plus. Matlab is published by The Math Works, Inc., South Natick, MA 01760, U.S.A. The root mean square of the deviations from the fit (standard deviation) is about 41.6 for both linear and exponential fits.

of the title. Our prediction qualitatively describing the future is that

The high-low interaction will yield increasingly deep insights and powerful stimuli.

In the rest of this paper we support this prediction by citing examples showing an interplay between the high and low roads. The choice of examples is necessarily based, although loosely, on the author's own experience. The reader undoubtedly will see more examples.

The term *high* means the application of advanced tools to solve a matrix problem requiring only elementary concepts for its formulation, and perhaps for the description of its solution. Stated differently, it means the application of devices and concepts from an area of *Mathematical Reviews* other than the linear algebra classification. One of the most significant trends over the past several decades has been the multifaceted increasing use of high level tools to solve a wide range of linear algebra questions. These tools sometimes involve algebraic geometry—surely to be expected, since linear algebra problems frequently are many variable polynomial problems. But combinatorics and graph theory often underlie linear algebra questions, as does Lie theory, and sometimes functional analysis. Other areas often seen are projective geometry, control theory, and number theory. Matrix numerical analysis also has a major impact on pure linear algebra.

The high road may perhaps be described as "finding the right ideas" for the correct description of one's problem.[2] It really is quite accurate to state that the young future linear algebraist who hopes to find his own right ideas needs to be trained (at least) in graph theory, Lie theory, functional analysis, multilinear algebra, algebraic geometry, combinatorics, and numerical linear algebra.

The *low* road, the road using only concepts generally regarded as elementary, may lead to a worthwhile investigation at the high level. However, following the high road also creates an opportunity for a low road investigation, perhaps best stated as: simplify the proof!

Another distinction between *high* and *low* is this: Any part of mathematics that is familiar or understood is low, anything unfamiliar or not understood is high, and anything somewhat unfamiliar or partly understood is midlevel. Thus a topologist's assessment of the level of a particular topic may not agree with a linear algebraist's.

[2] The description *finding the right idea* arose in a conversation with S. Friedland.

4. THE NUMERICAL RANGE

Let A be a matrix, acting on n-space, and consider the set S of all complex numbers of the form

$$\sum_{j=1}^{k} c_j\left(Ax_j, x_j\right),$$

where (,) denotes the usual Hermitian positive definite inner product, the c_j are complex constants, and $x_1, \ldots, x_k$ are variable orthonormal sets of vectors, with k fixed. Thus S is the range of a map $U \to \operatorname{tr}(CUAU^*)$ from the unitary group to the complex plane, with fixed matrices A and C. [As usual, $*$ denotes the Hermitian adjoint, and we may take $C = \operatorname{diag}(c_1, \ldots, c_k, 0, \ldots, 0)$.] A question that R. Westwick investigated [11] is whether S is convex. A famous theorem from near the turn of the century, the Toeplitz-Hausdorff theorem, gives an affirmative answer when $k = 1$. When $c_1 = \cdots = c_k = 1$, around 1960 S was shown to be convex when A is normal in [8], and when A is arbitrary by Berger [4]. Using Morse theory (calculus on manifolds), Westwick proved (1973) the convexity when $c_1, \ldots, c_k$ are all real. The use of Morse theory puzzled those linear algebraists interested in the numerical range, and for some years an open question was to find an elementary proof of Westwick's theorem. This challenge was first met by Y. T. Poon [6], in 1980, who gave a nice elementary proof of the convexity. Here we have an excellent example of a high level approach creating an opportunity for an insightful contribution at a lower level.

Of course, the Westwick theorem immediately leads to the question of the convexity when the scalars $c_1, \ldots, c_k$ are complex, and Westwick had already shown that convexity is false. The next natural question, therefore, is whether the set S is star shaped. This was resolved by N. K. Tsing [10], who proved (1981) that S is indeed star-shaped, and provided the star center. Tsing's clever proof uses nothing advanced.

Since the Westwick theorem followed from manifold calculus in a geometrically insightful way, a return to the high level makes sense, to ask whether there is a geometrically insightful analog of the Tsing proof using manifold calculus. This surely exists, and it would be worthwhile to find it.

Here we have an example of an argument at an elementary level leading to a question at a higher level.

Bilinear inner products are as interesting as Hermitian ones, offering potentially easier theorems, since the difficulties induced by the law of inertia are lessened and sometimes removed. Thus consider the bilinear expression

$$\sum_{j=1}^{k} c_j[Ax_j, x_j],$$

where [,] is a bilinear form rather than a Hermitian one. However, in order to make the set T of values taken by this linear combination of forms into a bounded set, as the vectors x_i vary, these vectors must be restricted in some way. A natural restriction is to require that the x_i be orthonormal relative to a Hermitian definite inner product (,). So now two inner products are present. And now we ask whether T is convex. An equivalent description is whether the range T of the map $U \to \mathrm{tr}(CUAU^t)$ from the unitary group to the complex plane is convex, where A and C are fixed. (^{t} denotes transpose.)

It is easy to see that T has circular symmetry about the origin, and is connected. To prove convexity it is therefore only necessary to show that 0 is achieved, for some choice of unitary U. This was proved in [7] and [9] for special choices of A, and then in full by Choi, Laurie, Radjavi, and Rosenthal in a joint paper [3]. The key idea is to find a closed path $U(t)$ in the special unitary group such that its image under the map $U(t) \to \mathrm{tr}[CU(t)AU(t)^t]$ encloses the origin. Then one invokes the fact that the fundamental group of $\mathrm{SU}(n)$ is trivial, so the path is homotopic to a constant. Consequently the image path in the complex plane enclosing the origin must deform continuously to one not enclosing the origin, and therefore at some stage in the deformation must give a path in the complex plane passing through the origin. Consequently 0 is achievable, and hence our set T in the complex plane must be a disk centered at 0.

Here we have another example in which a more advanced point of view leads to an insightful solution of a low road matrix problem.

Westwick's and Tsing's theorems have recently been lifted to operators [5], and this is a different lift of a low road theorem to a higher level.

5. SIMILARITY INVARIANTS OF PRINCIPAL SUBMATRICES

The problem is this: Choose an $n \times n$ matrix M with entries in a field and given similarity invariants; then describe the possible similarity invariants of the leading $k \times k$ principal submatrix A of SMS^{-1}, as S ranges over all nonsingular matrices. This was solved by Sa [15] and Thompson [17], more or less simultaneously. Let the similarity invariants of M and A be

$$h_1(M) \mid \cdots \mid h_n(M) \quad \text{and} \quad h_1(A) \mid \cdots \mid h_k(A),$$

respectively. These are polynomials in a single variable, those for M forming a divisibility chain, and those for A another. (Here $\mid$ denotes divisibility.) Their product, for M, say, is the characteristic polynomial of M, and the last is the minimal polynomial of M. Here are the necessary and sufficient conditions that Sa and Thompson found:

$$h_i(M) \mid h_i(A) \mid h_{i+2(n-k)}(M), \qquad i = 1, \ldots, k,$$

under the convention that $h_j(M) = 0$ if $j > n$. Of course, we have degree constraints

$$\text{degree}\left[h_1(M) \cdots h_n(M)\right] = n, \qquad \text{degree}\left[h_1(A) \cdots h_k(A)\right] = k,$$

a seemingly trivial constraint.

Most of the Sa-Thompson proofs were not particularly hard. The special case when A has deficiency one in M admits a tidy formulation:

$$h_1(M) \mid h_1(A) \mid h_3(M), \ldots, \qquad h_{n-2}(M) \mid h_{n-2}(A) \mid h_n(M),$$
$$h_{n-1}(M) \mid h_{n-1}(A)$$

(interlacing with a step of two), together with the degree conditions.

Only for $k < n - 1$ did a genuine difficulty appear in the otherwise reasonably easy proofs, and it was a significant barrier. This difficult point appeared in the constructive phase; namely, if polynomials $h_i(M)$ and $h_j(A)$ are given satisfying the divisibility conditions, an argument has to be found producing the matrix M containing a principal submatrix A such that both have the prescribed similarity invariants. This of course is an induction argument, and in it the apparently trivial degree condition proved to be a major obstacle. The difficulty was that the constructive steps in the induction argument conserved polynomial divisibility very nicely, but not polynomial degree.

Thompson handled the obstacle by a long chain of intricate but elementary lemmas. Sa handled it by a short but unintuitive argument involving the approximation of a definite integral by an integral of step functions. Sa also obtained a convexity fact not found by Thompson. Here we have an example of two successful roads, one clearly higher than the other, with the higher road giving sharper results in a more economical manner. Later, P. Y. Cheng [12] reworked the obstacle with a rather combinatorial but still intricate argument. Unfortunately, none of these proofs is easy to follow, and it probably is true that each author understood only his/her own proof. According to G. N. de Oliveira [13], nobody else understood any of them. Thus these proofs could not be described as "natural" resolutions of the problem.

A good road to the theorem was found by Zaballa [19, 20] by invoking ideas from control theory. His not hard to describe proof is one that everybody can understand, albeit with effort, since the argument is still nontrivial (though it is clear). First, observe that, given a matrix A with prescribed similarity invariants, the problem to be solved is when matrices B, C, D can be found such that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

has prescribed similarity invariants. Zaballa's idea (which was based on a paper by Wimmer [18]), is that the polynomial matrix $[\lambda I - A, -B]$ is important in control theory, and the invariants describing it under a natural equivalence relation are invariant factors and control indices. So, given A, ask what invariant factors and control indices are achievable by choosing B suitably. Then, having a rectangular matrix $[\lambda I - A, -B]$ with given invariant factors and control indices, ask what invariant factors are achievable for

$$\begin{bmatrix} \lambda I - A & -B \\ -C & \lambda I - D \end{bmatrix}$$

under suitable choice of C and D. The two theorems so obtained intersect to prove the Sa-Thompson result. Bringing control indices into the picture is really quite natural, even they do not appear at the final theorem, because rectangular matrices of the form $[\lambda I - A, -B]$ are quite central to control theory, and therefore are well studied. So here we have a case in which finding the right idea led to the good solution. However, it is possible to eliminate all mention of control theory if another description of control indices is used: They are just the minimal indices, in the sense of matrix pencil theory, of the polynomial matrix $[\lambda I - A, -B]$. So if this idea had been noticed at the time of the original solutions, those solutions would have been the good ones.

This is an example in which the correct idea produced the best argument. We may say this is a high road solution because the required idea came from a field, control theory, normally regarded as external to linear algebra. But the theory of matrix pencils, well explained in Gantmacher's elementary textbook, is usually regarded as a part of linear algebras, so it equally accurate to say that the complete solution now lies in the low road.

Sa recently simplified his proof [16], bringing the convexity into clearer focus.

Nonprincipal submatrices are just as interesting as principal submatrices, and [14, 21] are two pertinent recent papers.

6. COMMUTATORS

A matrix or operator M is a commutator (or Lie commutator) if $M = AB - BA$ for some matrices or operators A and B. When is a matrix M a commutator? This is our low road, the finite dimensional situation. The high road is to ask the same question for operators, that is, the infinite dimensional case.

The finite dimensional result is easy to state [22]: a matrix M with entries in a field is a commutator $M = AB - BA$ of suitable matrices A, B over the same field if, and only if, $\operatorname{tr} M = 0$. Thus the presence of trace function yields a very clean result.

The commutator question for operators was taken up by Pearcy, Brown, and Halmos, and later by Brown and Pearcy [23], who produced a strong result: An operator on a separable Hilbert space is a commutator if and only if it is not of the form $\gamma I + C$, where γ is a nonzero scalar, I is the identity, and C is compact. Also very striking is that any operator on an infinite dimensional Hilbert space is a sum of two commutators (see [37].)

A matrix or operator M is a group commutator if $M = ABA^{-1}B^{-1}$ for suitable invertible matrices or operators A and B. For finite dimensional matrices over a field, $\det M = 1$ is a necessary condition for M to be a group commutator of suitable matrices A and B with entries in the same field, and this is sufficient [31–33] except for the two-square matrices over the field of two elements. The presence of a determinant function makes the problem a clean one. In infinitely many dimensions the facts appear not yet to be completely understood, but must be different [24, 37]. Illustration: any invertible operator on Hilbert space is a product of two group commutators. Second illustration: a scalar operator γI (with γ a complex scalar) is a group commutator if and only if $|\gamma| = 1$. The sharpest contrasting finite dimensional theorem is that γI_n, where γ is a primitive nth root of unity in the base field and $n \equiv 2 \pmod 4$, is a group commutator $ABA^{-1}B^{-1}$ with A and B both having entries in a base field and determinant one if and only if γ is a sum of two squares of elements from the base field.

The last remark pinpoints a significant difference between the finite and infinite dimensional problems: For the finite dimensional, rationality questions are important: over what coefficient field do the requisite matrices exist? The proofs over the complex numbers are relatively easy, and are harder only over fields with just a few elements. Indeed, over finite fields, the group commutator question is a finite group problem. For the infinite dimensional discussion, rationality is very little considered, complex scalars being assumed from the outset.

For self-commutators, those Hermitian matrices or operators M of the form $M = AA^* - A^*A$, rationality is unimportant, the scalars being the complex numbers. A Hermitian matrix M is a self-commutator if and only if it has zero trace [30], and a Hermitian operator M is a self-commutator if and only if it has nonpositive and nonnegative limit points in its spectrum (see [28] for the precise statement.)

Self-commutators for matrices or operators are unexpectedly linked with convexity in a collection of theorems and a conjecture due to S. Friedland [27]. Let A be a bounded linear operator on a Hilbert space. For matrices, a trace computation shows that $A^*A - AA^*$ is zero if it is nonnegative definite, but for operators $A^*A - AA^*$ can be nonnegative definite and nonzero. The shift operator $A:(x_1, x_2, \ldots) \to (0, x_1, x_2, \ldots)$ on l_2 is an example. If $A^*A - AA^*$ is nonnegative definite, Friedland proved that $\log\|e^{At}u\|$, as a real valued function of a real scalar t, is convex for every nonzero vector u, and therefore has a nonnegative second derivative at $t = 0$. The conjecture is the converse: if $\log\|e^{At}u\|$ has a nonnegative second derivative at $t = 0$ for every nonzero vector u, then $A^*A - AA^*$ is nonnegative semidefinite. The vanishing of $A^*A - AA^*$, that is, the normality of A, is proved in [27] to be equivalent to the convexity of both $\log\|e^{At}u\|$ and $\log\|e^{A^*t}u\|$ for every nonzero vector u. For matrices, only one of these conditions is needed.

A commutator problem may also lead in unexpected directions in the finite dimensional case, for we have the theorem that for even n the determinants of the factors A and B in $\gamma I_n = ABA^{-1}B^{-1}$ take prescribed values precisely when [34] a certain Hilbert symbol is 1. The Hilbert symbol belongs to number theory, and is a primary ingredient of the Hasse-Minkowski theory of quadratic forms over the p-adic numbers. This condition was restated in terms of the norm residue symbol of number theory by Estes [26], one of his methods being an application of the theory of cyclic algebras.

The finite dimensional facts thus for the most part are not specializations of the infinite dimensional ones, and actually the finite dimensional theory seems to have brought only a little impetus to the finite dimensional. Nor has the infinite dimensional study brought a lot back to the finite dimensional. The explanation, of course, is that the major difficulties are rather different: rationality in the finite dimensional case, and limits and convergence in the infinite dimensional. However, there is at least one theorem with a common thread: Any matrix or operator M has the form $M = PAQ - QAP$ for invertible matrices or operators P, A, Q, under the sole restriction that the underlying space must not be one dimensional [25]. Given M, how singular P, A, and Q may be is probably still an open question.

Nonetheless, the high road could continue to draw inspiration from the low. Example: A group commutator of two positive definite matrices, if it is normal, must be the identity [35]. To what extent is this valid for operators?

A recent paper of Vaserstein and Wheland [36] contains a nice historical survey of work on commutators of matrices and operators, including recent work involving commutators of matrices over rings.

The low and high roads for these problems thus have been highways with only a few crossings. This suggests that when searching for the right ideas for a finite dimensional problem, an upward tilting of the eyes to infinitely many dimensions may not be wise.

The study of commutators of matrices was begun by Shoda in his paper [29].

7. THE TRIANGLE INEQUALITY

If A is a matrix, define its matrix valued absolute value by

$$|A| = \left(A^*A\right)^{1/2}.$$

The matrix triangle inequality [40, 41] asserts that

$$|A+B| \leqslant U|A|U^* + V|B|V^*,$$

for unitary matrices U and V that depend on A and B. Without the presence of U and V the theorem is false. The inequality sign here means that the right side minus the left is positive semidefinite (the Loewner partial order.) Here we have a low road theorem, which to some extent was motivated by a paper by the theoretical physicists E. Seiler and B. Simon [39].

A question coming from theoretical physics will, we stipulate, belong to the high road, since it almost surely will be significant.

A natural high road question is to ask whether the same triangle inequality holds for operators. This was taken up by Akemann, Anderson, and Pederson [38]. Their investigation was made intricate by the issue of what the analogue of the unitary matrices U and V should be: they could become isometries (into maps), or they could be unitary in the fullest sense, namely isometries onto. Among Akemann, Anderson, and Pederson's results are two theorems exhibiting this dichotomy. The first is that for A and B in a von Neumann operator algebra acting on a separable Hilbert space, the triangle inequality does hold with U and V isometries. The second, for a C^* operator algebra, is that if the isometries are to be unitaries, then

$$|A+B| \leqslant U|A|U^* + V|B|V^* + \varepsilon I$$

for each positive ε. It seems to be still unknown whether the scalar εI term can be removed. In fact, it is not known whether this term can be removed when the operators are compact, and it would be worthwhile to investigate this question.

The matrix valued triangle inequality is still valid when the scalars are quaternions, with the same proof [42]. It is not been checked whether Akemann et al.'s results remain the same for operators on quaternionic spaces.

Thus the low road finite dimensional problem may fairly be said to have stimulated the high road infinite dimensional one, but the reverse seems not to be true.

8. THE FACIAL STRUCTURE OF THE UNIT BALL

The unit ball B in an algebra of operators on a Hilbert space is the collection of operators with norm at most 1. Of course, the structure of the unit ball depends on the norm being used. A question recently of interest in functional analysis circles is to describe the faces of the unit ball. This question was raised by Akemann [43]. Remember that a convex subset F of B is a face if whenever $tx + (1-t)y$ lies in F for x, y in B and $0 < t < 1$, one has x, y in F. This question, although asked in an infinite dimensional context, is still a worthwhile one in finitely many dimensions. The infinite dimensional version, of course, requires advanced techniques. Some parts of the finite dimensional question, however, can be handled by reasonably direct devices familiar to most linear algebraists. A recent paper by W. So [44] describes the faces for the Schatten p-norms.

This is an example of a high level question (faces of the unit ball in operator algebras) that is still a good one at the low level (matrices).

There are some matrix approximation problems that originate as special cases of approximation problems in infinitely many dimensions [43]. One of these is to approximate the J-matrix (all entries $1/n$) by a sparse matrix, using any convenient norm yielding a good theorem. This kind of problem, even though finite dimensional, may be quite difficult. P. Halmos has lectured on it in dramatic ways.

9. THE GERSCHGORIN CIRCLE THEOREM

There are few matrix theorems more familiar than Gerschgorin's, to be found in nearly every numerical analysis text. It provides circular disks in the complex plane, centered on the diagonal elements of a matrix, the union of which contains all the eigenvalues. The applications of this theorem, especially in numerical analysis, are very numerous. How could such a simple minded theorem be so important? Not quite as well known is the ovals of Cassini theorem [45]. This provides ovals in the complex plane, based on pairs of diagonal elements, the union of which encloses all the eigenvalues of the matrix. Long ago (in the 1950s) it was observed that there is no extension of this pair of theorems to one involving curves based on more than two diagonal elements. This is the elementary approach, and while it yields important facts, it seems to come to a halt too soon. These facts, and much more, are nicely summarized in [46] and [47, Chapter 6].

However, Richard Brualdi [46] understood that these two theorems possessed a formulation using graph theory ideas, and then a generalization did exist. This, then, is the high road. For uncertain reasons, graph theory is not always regarded as high, in the same sense that functional analysis is high. This surely is a matter of individual taste and perspective, and perhaps reflects the (false) idea that a theoretical framework yielding insights about how finitely many objects cannot be as sophisticated as one that studies infinitely many. Graph theory is regarded here as high, since it brings insight to an elementary problem that seemed to have no further life.

Let us describe part of Brualdi's results. First, a directed graph is associated with a matrix $A = [a_{ij}]_{1 \leqslant i, j \leqslant n}$. The vertices are the integers $1, \ldots, n$, and i is connected to j if a_{ij} is nonzero. A circuit is a sequence $i_1, i_2, \ldots, i_k$, i_1 of vertices such that each is connected to the next. Call a matrix *weakly irreducible* if every vertex of its graph belongs to at least one circuit. An irreducible matrix, one not permutation similar to the form

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix},$$

is always weakly irreducible. Now Brualdi's theorem states this: *The eigenvalues of a weakly irreducible matrix A lie in the union of the regions*

$$\prod_{i \in \gamma} | z - a_{ii} | \leqslant \prod_{i \in \gamma} R_i .$$

Here γ is a circuit, a region is all complex numbers z satisfying this inequality for a fixed circuit, and the union is taken over all circuits. The symbol R_i

means (as usual) the sum of the absolute values of the entries in row i of A excluding the diagonal entry.

Since the union extends over all circuits, at first sight this theorem appears to give a bigger union than the one provided by Gerschgorin's theorem, namely the union of the disks

$$|z - a_{ii}| \leqslant R_i.$$

However, this is not the case, since Brualdi showed that his theorem contains Gerschgorin's (and the ovals of Cassini theorem, too.)

It is quite easy to make the computer draw the Gerschgorin circles of the matrix, for example, using Matlab. Indeed, I did this for my undergraduate numerical analysis class. The program works very nicely. So I thought I would write a program to draw the inclusion regions provided by Brauldi's theorem. However, I found this unexpectedly troublesome, and I still do not know how to do it. A major difficulty is that all the circuits have to be found. Consequently it is much easier to have Matlab compute the eigenvalues than compute the inclusion regions containing them. So here we have an unexpected hazard (a pothole?) lurking in the high road: the barrier of combinatorial intransigence.

The intractibility of finding the circuits is well known to graph theory experts. This intractibility led C. R. Johnson [48] to ask whether a "part way" application of Brualdi's theorem is possible. Finding a good definition of "part way" is a key point in this question.

Can all eigenvalues of a matrix lie outside the union of regions defined by the naive (false) generalizations of the Gersgorin circle and Cassini oval theorems? Probably, but the published examples seem to have only two outside eigenvalues.

10. MATRICES, GRAPHS, INERTIA, NUMBER THEORY

A familiar type of theorem is that if a portion of a matrix A is fixed, the rest can be found such that the completed matrix has certain properties. Often the theorem states that the completion is possible if and only if certain conditions are satisfied. A standard example is this one: If a Hermitian matrix H is given with known eigenvalues, there exists a Hermitian matrix K with one more row (and column) having prescribed eigenvalues if, and only if, the eigenvalues of H and K interlaced.

In recent years, under the expert guidance of C. R. Johnson and coauthors, many theorems (especially theorems involving inertia) have been found in which the part of A that is given is described by a graph, often a chordal graph. See, for example, [49, 50, 54–61]. Here we have an example in which the presence of a new idea, the specification of a portion of a matrix by a graph, has led to an extremely rich store of questions and more than a few answers. The high road here is the reinvigoration, by the introduction of graph theory, of a part of matrix theory that seemed to be approaching completeness.

The number theoretical counterpart of these theorems is also interesting, and nontrivial: If the invariant factors of a portion of a matrix are given, when can the rest of the matrix be specified so that the whole has prescribed invariant factors? The Sa-Thompson-Zaballa theorem has this structure, and the linear algebra group lead by G. N. de Oliveira has obtained many similar results. The influence of graph theory in these results is not yet particularly strong and seems destined to grow.

The law of inertia and its p-adic analogue blend in the very important local-global principle of number theory. One problem that this part of number theory addresses is the embeddability of one rational symmetric matrix A in a congruence transform XBX^t of another, B, with X nonsingular. Thus a graph is present, but not mentioned. An obvious question is whether some of the ideas from the last two paragraphs may combine. This seems not much considered and may yield some theorems. An immediate question is: what properties must the graph specifying a submatrix of a rational matrix have if a local-global principle holds for its completion of a full matrix? The correct formulation of this question is needed before there will be a theorem. (The local-global principle is that if a completion is possible p-adically for every prime p, and for the real numbers, then it is possible over the rational numbers.)

Moving away from completion questions, an undirected graph defines a zero-one symmetric adjacency matrix A; viz. $a_{ij} = a_{ji}$ is 1 exactly when vertices i and j are connected. The isomorphism problem for these graphs is when two of them are the same apart from a relabeling of vertices. This is equivalent to asking when their adjacency matrices A, B are permutation equivalent: $B = PAP^t$ for some permutation matrix P. Thus if $B \neq XAX^t$ for any integral or rational matrix X, then the graphs are not isomorphic. The number theory of quadratic forms often furnishes answers to when $B = XAX^t$ for some X. A very active subject with a long history, for the graph isomorphism problem its scope has to be extended to inquire about the nature of the transforming matrix X. Friedland's paper [53] discusses these ideas and is attractive reading for those familiar with the number theory of quadratic forms.

Graph theory uses linear algebra in numerous ways, for example, by study of the spectrum of a matrix belonging to a graph [51, 52, 62]. And graph theory gives much back to linear algebra: one need only look at the increasing number of linear algebra papers involving a graph theoretical concept. The Gerschgorin circle discussion in the last section is an example. Thus the connection between linear algebra and graph theory is hardly new, yet its vigor and the force of the resulting ideas are increasing at a remarkable rate.

11. POWER EMBEDDINGS AND DILATIONS

The problem here is this: Given a matrix A, when can we embed A as a principal submatrix of a larger matrix M such that A^i is a principal submatrix of M^i for $i = 1, 2, \ldots, k$? Here k is fixed. This problem attracted attention as long ago [63] as the early 1950s, in the case that the larger matrix M was to be unitary. In this case A must be a contraction (every singular value at most one.) The problem became important in operator theory, with k infinite of course, and M acquired the name of a *unitary dilation* of A. The significant problem is to find the smallest M, in some sense of the term "smallest." Sz.-Nagy and Fois's book [64] devotes nearly 100 pages to this problem.

Going back to the finite dimensional case, a large class of related problems may be identified [65], all of which should possess a comparable dilation theory in the infinite dimensional case:

(1) If A has nonnegative entries, when can a doubly stochastic M be found such that $A^i \subset M^i$ for $i = 1, \ldots, k$? Here $A \subset B$ means that A is a principal submatrix of B.

(2) The same question when A is a contraction and M is to be unitary.

(3) The same question for integral matrices when A has prescribed invariant factors and M is to be unimodular.

(4) The same question when A is a general complex matrix and M is to be complex orthogonal.

(5) The same question when A is general and M is to be nonsingular.

In each case, the significant question is the size (= number of rows) in the smallest M. And the answer in each case takes this form: M exists if and only if it has at least $k\delta$ more rows than A, where δ measures the deficiency of A from the desired status. These deficiencies are, respectively, the least integer δ greater than or equal to:

(1) the sum of the entries of $I - A$;

(2) the number of singular values of A strictly less than one, equal to the rank of $I - AA^*$;

(3) the number of nonunit invariant factors of A;

(4) the rank of $I - AA^t$;

(5) the nullity of A.

In view of the existence of a significant infinite dimensional theory expanding the finite dimensional unitary case, does there exist an interesting infinite dimensional analogue of each of the other finite dimensional results? Here we have a collection of low road theorems looking for a high road.

12. THE SCHUBERT CALCULUS

A flag is a tower of subspaces,

$$V_1 \subset V_2 \subset \cdots \subset V_k,$$

contained within a finite dimensional space V. The Schubert calculus is concerned with triples of flags,

$$A_1 \subset A_2 \subset \cdots \subset A_k, \qquad B_1 \subset B_2 \subset \cdots \subset B_k, \qquad C_1 \subset C_2 \subset \cdots \subset C_k,$$

where $\dim A_i = a_i$, $\dim B_i = b_i$, $\dim C_i = c_i$, $i = 1, 2, \ldots, k$. The components of the vectors $a = (a_1, a_2, \ldots, a_k)$, $b = (b_1, b_2, \ldots, b_k)$, and $c = (c_1, c_2, \ldots, c_k)$ are strictly increasing positive integers. A specific question the Schubert calculus asks is this: under what conditions on the indices a, b, c is it guaranteed that a k-dimensional subspace L_k of V exist such that

$$\dim(L_k \cap A_i) \geqslant i, \qquad \dim(L_k \cap B_i) \geqslant i, \qquad \dim(L_k \cap C_i) \geqslant i$$

for $i = 1, 2, \ldots, k$. The three flags are said to intersect if the space L_k exists. Of course, whether it exists may well depend on the closure properties of the field of coefficients, so the question is asked over an algebraically closed field of scalars. There is an elaborate combinatorial apparatus available for investigating the question, involving Young tableaux, Schur symmetric polynomials, and the Littlewood-Richardson rule. (Chapter XIV of [68] is the classical presentation of the Schubert calculus, and tableaux, Schur symmetric polynomials, and the Littlewood-Richardson rule are explained beautifully in [72].) So the study of the Schubert calculus involves both geometry and combinatorics. Let us adopt the geometrical outlook here.

One reason the Schubert calculus is interesting in linear algebra is the extreme value characterization of eigenvalues found by J. Hersch and B. P. Zwahlen in the 1960s [67, 76]. To explain this, let H be a Hermitian matrix with eigenvalues $h_1 \geqslant h_2 \geqslant \cdots \geqslant h_n$ and associated orthonormal eigenvectors $e_1, e_2, \ldots, e_n$. The Hersch–Zwahlen extremal principle is that, for a scattering $i_1 < i_2 < \cdots < i_k$ of indices,

$$h_{i_1} + h_{i_2} + \cdots + h_{i_k} = \max_{L_k} \{ (Hx_1, x_1) + (Hx_2, x_2) + \cdots + (Hx_k, x_k) \}.$$

Here (,) denotes the usual positive definite inner product, and the vectors $x_1, x_2, \ldots, x_k$ are an orthonormal basis of a subspace L_k. It is easy to show that the sum in braces depends only on the subspace L_k and not on its orthonormal basis $x_1, x_2, \ldots, x_k$. The max is taken over all choices of subspaces L_k satisfying the following flag intersection property relative to the eigenspaces of H:

$$\dim\left[L_k \cap \operatorname{span}\left(e_{i_{k+1-s}}, \ldots, e_n\right)\right] \geqslant s \qquad \text{for} \quad s = 1, 2, \ldots, k.$$

This condition may look contrived, but be assured that it is a natural one.

Now let A, B, C be Hermitian matrices, with $C = A + B$. Take the eigenvalues of A, B, C to be $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_n$, $\beta_1 \geqslant \beta_2 \geqslant \cdots \geqslant \beta_n$, $\gamma_1 \geqslant \gamma_2 \geqslant \cdots \geqslant \gamma_n$, respectively. The application found by Hersch and Zwahlen [67, 76] of their extremal principal is this: In order to prove a spectral inequality

$$\gamma_{k_1} + \gamma_{k_2} + \cdots + \gamma_{k_r} \leqslant \alpha_{i_1} + \alpha_{i_2} + \cdots + \alpha_{i_r} + \beta_{j_1} + \beta_{j_2} + \cdots + \beta_{j_r},$$

it suffices to prove that the flags built from the eigenspaces of A, B, and C on the indices $i = (i_1, \ldots, i_r)$, $j = (j_1, \ldots, j_r)$, $k = (n - k_r + 1, \ldots, n - k_1 + 1)$ intersect.

This means that proving a spectral inequality is reduced to a geometric proposition, showing that triples of flags intersect. Some years ago S. Johnson [70] attempted to show that the geometry inherent in the Schubert calculus could explain the class of inequalities known as the (Alfred) Horn inequalities for the spectrum of a sum of Hermitian matrices [69]. This attempt made a lot of progress, but fell short of its objective.

Even if it could be completed (and it seems very likely that it can), the underlying philosophy is a deep one, that no one has really considered in its full generality.

To what extent are the inequalities of linear algebra explainable in geometric terms?

Linear algebra is full of inequalities, and inequalities perhaps are the biggest component of the subject. But the historical record throughout mathematics is that geometric insight penetrates underlying structure most deeply. Giving our question a precise meaning should therefore be unusually rewarding.

The Hersch-Zwahlen principle was discovered by R.C. Riddell [73], and lifted to the infinite dimensional context in a very substantial paper taking the highest of the high roads. Riddell's objective was to extend to operators the max-min extremal principle of Wieldandt [75] for sums of Hermitian matrix eigenvalues, and in the process of so doing he rediscovered the Hersch-Zwahlen principle and understood its greater efficiency. It seems clear that Riddell would have found major insights, were it not for his untimely death. His paper comes nearer than any other to answering our questions.

However, the geometry of the Schubert calculus ultimately depends on combinatorics, and specifically on the combinatorics belonging to the representation theory on the symmetric group (where the Littlewood-Richardson rule has its natural home.) This leads to another question:

To what extent are the inequalities of linear algebra explainable in combinatorial terms?

In some eyes, placing combinatorics on an equal footing with geometry is heresy. Those who have struggled through a linear algebra problem that unexpectedly turned combinatorial, and survived, will however appreciate the depth of the combinatorial viewpoint.

The study of the Schubert calculus was Hilbert's fifteenth problem; see [71] for an excellent discussion. The "calculus" part is that explicit computations are possible by virtue of the fact that an algebra is associated with the flags, the algebra of *Schubert cycles*. (A discussion is in [68]. The collection of spaces L_k meeting a flag forms a variety in the sense of algebraic geometry, and a Schubert cycle is a homology class of these varieties.) The algebra of Schubert cycles is a homomorphic image of the algebra of symmetric polynomials. The Littlewood-Richardson rule describes how the tensor product of two irreducible representations of a classical group decomposes as a direct sum of irreducible representations. It also describes how the product of Schur symmetric polynomials is expressed as a linear combination of Schur symmetric polynomials. Consequently, it is involved in the multiplication of Schubert cycles. The decomposition of the tensor product of irreducible representations is important in the field theories of modern physics, since an elementary particle is described by an irreducible representation, and particles interact by tensoring and then decomposing. A description of this circle of ideas is in [66]. Thus the low road attitude of just hoping to understand some spectral inequalities leads to a high road heading toward the *theories of everything* of modern physics.

The article [74] explains, among other things, why algebraic closure is needed.

13. THE SPECTRUM OF A SUM OF HERMITIAN MATRICES

Let A, B, C be Hermitian matrices, with $C = A + B$. A classical question is to describe the allowable spectrum of C assuming the spectra of A and B to be known. More precisely, given Hermitian matrices A and B, describe the allowable eigenvalues of $C = UAU^* + VBV^*$ as U and V run over the full unitary group. This question appears to have been formulated by I. M. Gel'fand, possibly jointly with M. A. Naimark, in the early 1950s. There is a paper by V. B. Lidskii [81] that established an inequality now known as Lidskii's inequality. Let A, B, and C have eigenvalues $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_n$, $\beta_1 \geqslant \beta_2 \geqslant \cdots \geqslant \beta_n$, and $\gamma_1 \geqslant \gamma_2 \geqslant \cdots \geqslant \gamma_n$, respectively. Then Lidskii's inequality is

$$\gamma_{i_1} + \gamma_{i_2} + \cdots + \gamma_{i_m} \leqslant \alpha_{i_1} + \alpha_{i_2} + \cdots + \alpha_{i_m} + \beta_1 + \beta_2 + \cdots + \beta_m$$

whenever the indices $i_1, \ldots, i_m$ are strictly increasing in the range $[1, n]$. These inequalities are nonsymmetric in the roles played by A and B. Of course, there is a similar inequality with the roles of A and B interchanged.

This problem, the spectrum of a sum of Hermitian matrices, apparently originated in joint work by Gel'fand and Naimark on the representation theory of the classical Lie groups, and the above inequality just displayed seems first to have been found using representation theory methods. The proof using representation theory was only published later, though [77]. It is easy to see that the problem belongs to a Lie algebra setting, since the skew Hermitian matrices are one of the classical Lie algebras over the real numbers. (Consider $iC = iA + iB$ in place of $C = A + B$.)

The elementary proof of the above inequality published by V. B. Lidskii [81] was too sparse in details to be understood by its readers. Wielandt [84] responded to this challenge, and provided a proof by creating a generalization of the Courant-Fischer max-min description of the eigenvalues of a Hermitian matrix. This description could be fitted to the situation at hand because of the nonsymmetric roles played by A and B. Unfortunately, this lack of symmetry obscured the fact that the problem really is a geometric one and actually lies in the geometry of finding triples of flags that intersect. The natural inequality that brings back the symmetry is this one [82]:

$$\gamma_{k_1} + \gamma_{k_2} + \cdots + \gamma_{k_m} \leqslant \alpha_{i_1} + \alpha_{i_2} + \cdots + \alpha_{i_m} + \beta_{j_1} + \beta_{j_2} + \beta_{j_m}$$

when $k_p = i_p + j_p - p$, $p = 1, \ldots, m$. This inequality has been found to hold in many other situations. Let us call it the *standard spectral inequality*. It is one of the simplest inequalities that can be described using a Young tableau, and experience has shown that it is valid in any situation in which a Lie or group representation theory formulation of the matrix question exists. See [83], for example.

To see spectral inequalities from a representation theory viewpoint is, of course, the high road, and it ultimately leads to the representation theory of the symmetric group, and thus to combinatorics. The low road is to give elementary proofs, in the spirit of V. B. Lidskii or (later) Alfred Horn [79] and others.

The representation theory proof of Lidskii's inequality was published in Berezin and Gel'fand's paper [77]. An interesting aspect of this paper is that it discusses the decomposition of the tensor product of representations of SU(3), obtaining values $\frac{2}{3}$ and $-\frac{1}{3}$ for certain quantities. These are the numbers that physicists determined to be the quark changes, when it was found that the decomposition of tensor products of representations of SU(3) describes quarks. Today, in physics, quarks are an ingredient of unified field theories, sometimes called theories of everything because their objective is the complete explanation of the structure of matter. Thus, as was noted in the last section, matrix spectral theory seems to have a high road leading to everything. Of course, enormously more than spectral inequalities are incorporated in theories of everything, yet this connection does tell us that when we play with eigenvalue inequalities, our game is a serious one.

Recently, B. V. Lidskii [80] announced a solution of the problem of describing the allowable spectrum of a sum of Hermitian matrices, but (as this is written) his proof is unpublished, J. Day, W. So, and I [78] have reconstructed a large part of it, relying heavily on the partial resolution of this problem published by A. Horn [79] in 1962. The proof contains an analytic first part that reduces the problem to a combinatorial second part. The analysis part follows Horn's method closely, which in turn was based on the brief description published by V. B. Lidskii. A complete discussion of the combinatorial part requires the insight that Young tableau combinatorics are involved, specifically the Young tableau combinatorics centering around the Littlewood-Richardson rule. It is in principle possible to avoid the tableaux in low dimensions by using a computer and an exhaustion of cases approach to verify the correctness of the needed combinatorial properties.

It is also possible to compute recursively the full family of spectral inequalities in low dimensional cases. A computer program written in BASIC has shown that there are 2063 linear inequalities describing the spectrum of a sum of two 7×7 Hermitian matrices.

C. R. Johnson has asked whether this vast collection can be reduced to a smaller essential set. One approach to this question is to note that, by virtue of the computational rules in the Schubert calculus known as Pieri's formula and

Giambelli's rule, generators for the algebra of Schubert cycles are known. This may mean that a basic set of spectral inequalities can be identified.

14. THE HADAMARD-SCHUR PRODUCT

The Hadamard or Schur product of matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ is $A \circ B = [a_{ij}b_{ij}]$. Many very attractive theorems have been found concerning it, and the recent survey article by Roger Horn [87] gives a delightful summary. See also the book [89]. The product is not artificial, Horn exhibiting natural occurrences of it in partial differential equations, monotone matrix functions, electrical engineering, and elsewhere.

Various computational (algebraic) tricks seem to be needed to prove theorems about the Hadamard-Schur product, and perhaps this is inevitable, since its definition is so coordinate dependent. The use of coordinates may be regarded as *a* low road not conforming to modern abstract style. Words of criticism, however, are permitted only to those offering a convincing high road.

The Hadamard-Schur product naturally occurs in majorization: Let x be a real column n-tuple and y another. Majorization gives necessary and sufficient conditions in order that $x = Sy$, where S is a doubly stochastic matrix. Given x and y, with x majorized by y, it has long been known that the doubly stochastic matrix S may be chosen to be orthostochastic, that is, $S = U \circ \bar{U}$ for U unitary, and $\bar{U}$ is complex conjugation applied to the entries of U. There are related theorems not as well known. Taking x and y now to be complex column n-tuples, we ask: when is $x = U \circ Vy$, or when is $x = U \circ Uy$, for unitary U and V? That is, when is y transformed to x by the Hadamard-Schur product of two unitary matrices, or by the Hadamard-Schur square of a unitary matrix? The answers to the questions are found by reformulating them as relations between the diagonal elements of a matrix and its singular values. This is a formulation known to have a Lie theory interpretation.

If the elements of x and y are taken in order of decreasing absolute value, the answer [95] to when $x = U \circ Vy$, is: if and only the vector of absolute values of entries of x is majorized by the vector of absolute values of the entries of y, together with the subtracted term inequality

$$|x_1| + \cdots + |x_{n-1}| - |x_n| \leqslant |y_1| + \cdots + |y_{n-1}| - |y_n|.$$

The conditions for $x = U \circ Uy$ are [96] that the vector of absolute values of the entries of x is majorized by the vector of the absolute values of the entries of

y, together with the following larger set of subtracted term inequalities:

$$|x_1| + \cdots + |x_{i-1}| - |x_i| - \cdots - |x_n| \leqslant |y_1| + \cdots$$

$$+ |y_{i-1}| - |y_i| + |y_{i+1}| + \cdots + |y_n|, \qquad i = n, n-1, \ldots, 1,$$

$$|x_1| + \cdots + |x_{n-3}| - |x_{n-2}| - |x_{n-1}| - |x_n|$$

$$\leqslant |y_1| + \cdots + |y_{n-2}| - |y_{n-1}| - |y_n|.$$

(The last condition is present for $n \geqslant 3$ only.)

The correct understanding of these conditions belongs to Lie theory, and more specifically belongs to the properties of the root systems associated with Coxeter-Dynkin diagrams. Indeed, the entire topic of majorization is profoundly generalized and much better understood by setting it in the context of finite reflection groups. Nice explanations of this are in [86] and [93]. This is our suggested high road for an interpretation of the Hadamard-Schur product.

The question of when $x = U \circ U^t y$ for some unitary U, that is, when y is transformed to x by the Hadamard-Schur product of a unitary matrix and its transpose, seems to be unresolved. It is equivalent to asking for the allowable vector of diagonal elements of a matrix UYU, where $Y = \mathrm{diag}(y_1, \ldots, y_n)$. Matrices of this form appear to have no natural description and no natural applications, however, so this may not be an interesting question.

To some extent analogies exist between spectral properties of a Hadamard-Schur product of matrices and spectral properties of general sums or products, especially those related to Hermitian matrices. At least one may frame conjectures based on a supposed analogy. Thanks to modern software, especially Matlab on a personal computer, it is very easy to test these conjectures. Here is one: Let the singular values of a matrix A be $s_1(A) \geqslant \cdots \geqslant s_n(A)$. Then for any matrices A and B,

$$s_1(A \circ B) \leqslant s_1(A) s_1(B).$$

An immediate conjecture is that

$$\prod_{t=1}^{m} s_t(A \circ B) \leqslant \prod_{t=1}^{m} s_t(A) \prod_{t=1}^{m} s_t(B),$$

$m = 1, \ldots, n$. Using Matlab, a counterexample is quickly obtained, even though the corresponding relation with the natural product AB on the left (in place of $A \circ B$) has long been known to be true. One reason this conjecture fails for the Hadamard-Schur product is that the right side may vanish and the left not, since $A \circ B$ as a submatrix of the Kronecker product $A \otimes B$ may have rank equal to the product of the ranks of A and B.

In connection with their study of unitarily invariant norms, Horn and Johnson [88] proved that

$$\sum_{t=1}^{m} s_t(A \circ B) \leqslant \sum_{t=1}^{m} s_t(A) s_t(B),$$

$m = 1, 2, \ldots, n$. (See also Horn's survey [87].) The same formula holds for the natural product AB, and is a consequence of the $\Pi_{t=1}^{m}$ inequality displayed above acting on AB (instead of $A \circ B$). This suggests that the Hadamard-Schur product has weaker properties than the natural product. Therefore the many other valid inequalities satisfied by the singular values of the natural product may not translate to valid inequalities for the singular values of the Hadamard-Schur product. Here is a conjecture for the Hadamard-Schur product based on the valid standard inequality [94] for the singular values of the natural product:

$$\sum_{t=1}^{m} s_{k_t}(A \circ B) \leqslant \sum_{t=1}^{m} s_{i_t}(A) s_{j_t}(B)$$

when $k_t = i_t + j_t - t$, $t = 1, \ldots, m$, and the indices i_t, j_t are strictly increasing. This conjecture, true for the natural product, may easily be tested using Matlab, an exercise the author leaves to the reader.

This attractive pair of inequalities resembling the Cauchy-Schwarz inequality is due to Horn and Mathias [91]:

$$\| A^*B \|^2 \leqslant \| A^*A \| \, \| B^*B \|, \qquad \| A \circ B \|^2 \leqslant \| A^*A \| \, \| B^*B \|,$$

where A and B are matrices and the norm is unitarily invariant.

Another, not quite so recent paper [92] studies iterates of the map $A \to A \circ (A^{-1})^t$.

Finally, a conjecture by Johnson and Bapat [90] is

$$\prod_{i=1}^{k} \lambda_i(AB) \leqslant \prod_{i=1}^{k} \lambda_i(A \circ B)$$

for the eigenvalues of a product and a Hadamard-Schur product of positive definite matrices, the eigenvalues λ_i being numbered in increasing order.

15. THE EXPONENTIAL FUNCTION

Let x and y be noncommutative indeterminates. It is reasonably easy to prove that an infinite series z of words in x and y with rational coefficients exists such that $e^x e^y = e^z$. The Campbell-Baker-Hausdorff theorem asserts that z is a Lie element, meaning that it is a formal infinite series of iterated commutators of x and y with rational coefficients. If x and y are replaced by matrices, the resulting series of matrices may or may not converge, and B. S. Mityagin [101] recently demonstrated a theorem solving the long-open question of the domain of analyticity of $z = \log(e^x e^y)$. His theorem is that z depends analytically on x and y if $\|x\| + \|y\| < \pi$, and this estimate is sharp. (The norm is the operator norm.) However, this does not end the investigation.

Indeed, motivated by the triangle inequality, a new version of the theorem was conjectured by me more than ten years ago,

$$e^x e^y = e^{sxs^{-1} + tyt^{-1}},$$

where s and t are themselves formal infinite series in x and y. This version of the theorem is not very hard to prove. More interesting is that an infinite series $\rho(x, y)$ exists such that

$$s = e^{\rho(x, y)} \quad \text{and} \quad t = e^{\rho(-y, -x)}.$$

Moreover, $\rho(x, y)$ is again a Lie element, that is, expressible as an infinite series of iterated commutators of x and y with rational coefficients. (All of this is proved, but only partly published.) That s and t should possess the representations just described was suggested by this theorem on exponentials of Hermitian matrices: If H and K are Hermitian matrices, then unitary matrices U and V (dependent on H and K) exist [106] such that

$$e^{iH} e^{iK} = e^{i(UHU^* + VKV^*)}.$$

A student working under my guidance [108] wrote a FORTRAN program to compute the terms in the $\rho(x, y)$ series, to degree 10. The program works well, but the exponential growth of combinatorial possibilities prevents it from being effective on a small computer beyond degree 10. The computer output suggests that the sum of the absolute values of the coefficients in the $\rho(x, y)$ series is $\frac{1}{3}$. This is an interesting conjecture, but I remain uncertain that it is correct. All of this, except the proof of the above formula for $e^{iH}e^{iK}$, was found using only elementary ideas. The proof of this formula heavily used B. V. Lidskii's theorem on the spectrum of a sum of Hermitian matrices.

The question of the convergence of the $\rho(x, y)$ series when x and y are replaced by matrices is an interesting one. The strategy for proving a local convergence theorem is to find a differential equation for the desired quantity, since differential equation theory guarantees the existence of a local solution as a coinvergent infinite series. But I could not discover the differential equation. However, one was found by F. Rovière [102, 103], using tools that I did not possess. He is an expert in Lie theory, and of course the problem is a Lie theoretical one. The coupled pair of differential equations for $s(x, y)$ and $t(x, y)$ that he found are not particularly pretty, but they are differential equations, and by standard differential equation theory there is a solution expressible in a series that converges near the initial point (the origin.) Moreover, by using certain facts on the continuability of the solution of a differential equation acting on a compact manifold, Rovière was able to give a proof [102, 103] of my exponential formula for $e^{iH}e^{iK}$. His proof, moreover, works in the full generality of Lie groups and algebras, so establishes much more that I was able to get.

Here we have an interesting example of the interplay of the low and high roads. The two together have achieved something that neither by itself found. The high road, of course, is the Lie theoretic one.

There is a bit more to this story. I felt it necessary to verify my student's computation of the coefficients in the $\rho(x, y)$ series. To do the verification, of course, it is only necessary to substitute the $\rho(x, y)$ series into the right hand side of the formula

$$e^x e^y = e^{sxs^{-1}+tyt^{-1}}, \qquad \text{where} \quad s = e^{\rho(x, y)} \text{ and } t = e^{\rho(-y, -x)},$$

and verify that the resulting expression is the series product $e^x e^y$, to degree 10. That is, just substitute the presumed solution into the equation it is supposed to satisfy and see if it works. Now Mathematica [100] has very good symbolic manipulation capabilities, including a noncommutative multiply operation. It seemed ideal for this computation. So another student assistant wrote the Mathematica program [108]. This proved to be a bit challenging, but was eventually done. It is less than one page long. Various tricks were employed to shorten the actual computation. And then the program was run on a Macintosh II computer. The outcome was this: (1) it worked, (2) it was exceptionally slow. We were able to verify $\rho(x, y)$ up to its degree 6 terms after a nearly 24 hour run on the Macintosh II. To verify the degree 7 terms we needed an overnight run on a Sun workstation. Each new degree increases the running time by a factor of about 10, so further computations on small computers are hopeless. The inherent difficulty is the exponential growth in the combinatorial complexity of the terms as the degree goes up, combined with a surprising slowness of Mathematica. We found that we could get some Cray time, and Mathematica was announced for the Cray. Going this route would have been an easy way to carry our computation to degree 10, and it would have been fun to play a little on a Cray. Unfortunately, the Cray version of Mathematica so far has been vaporware, but we are told that a compiler version of Mathematica is being prepared. This should be much speedier, and we are therefore awaiting its release.

Apart from the nice interplay of the high and low roads visible here, it is also interesting to be involved in the interplay between computers and combinatorial complexity. These wonderful machines can help us to gain so many good insights, but they do not solve all of our difficulties.

The exponential function is a rich source of questions [97–99, 104, 107], some hard, some easier. Here [105] is one. Let $s_1(A) \geqslant \cdots \geqslant s_n(A)$ denote the singular values of A. Then

$$s_1(e^A) \cdots s_k(e^A) \leqslant e^{s_1(A)} \cdots e^{s_k(A)}$$

for $k = 1, \ldots, n$. This is actually a weak inequality; a sharper statement is true [98], and there are many more inequalities like this one. Even proving this one, however, requires finding the right idea, not a completely elementary one.

16. THE EXPONENTIAL FUNCTION AND COMMUTATIVITY

If A and B are commutative matrices, then clearly $e^A e^B = e^B e^A = e^{A+B}$. The converse question has been asked several times in recent years, and it is not quite so easy to see what the correct theorem is. An attractive converse theorem was published by E. Wermuth [109], and is this: If $e^A e^B = e^B e^A$, and if A and B have algebraic numbers as entries, then $AB = BA$. Recall that an algebraic number is a complex number satisfying a polynomial equation with ordinary integer coefficients. The theorem is false without the hypothesis that A and B have algebraic number entries.

The proof is an excellent example of bringing in the right idea. Assume that e^A and e^B commute. Then $p(e^A)$ and $q(e^B)$ commute for any choice of polynomials $p(x)$ and $q(x)$. If $p(x)$ and $q(x)$ can be found such that $A = p(e^A)$ and $B = q(e^B)$, then A and B must commute. So it is enough to prove that a polynomial $p(x)$ exists such that $A = p(e^A)$ when A has algebraic entries.

For this we simply have to write down an interpolation polynomial $p(x)$ (more precisely, an osculating polynomial) that at the eigenvalues of e^A takes the eigenvalues of A, and at which $p'(x) = 1$ and $p^{(i)}(x) = 0$, $i = 2, \ldots, n$. To see this, take A in Jordan form, and observe that these conditions simply say that a Jordan block of e^A is carried to a Jordan block of A. By elementary interpolation theory, this polynomial will exist unless distinct eigenvalues of A become equal eigenvalues of e^A. And here is where the hypothesis is used: distinct eigenvalues of A would become coincident eigenvalues of e^A only if they differed by an integral multiple of $2\pi i$. But π is transcendental and the eigenvalues of A are algebraic (because the entries of A are algebraic), whence differences of eigenvalues of A could not be an integral nonzero multiple of $2\pi i$.

This fairly simple application of transcendental number theory is the high road, and it seems clear that its use cannot be avoided.

17. INTEGRAL QUADRATIC FORMS

Let A and B be positive definite symmetric matrices with integer entries. The question of when $B = UAU^t$ for some unimodular integral matrix U is a classical one, going back to Gauss, but still of very great current interest, with modern applications (for example) in topology and graph theory. (The superscript ^{t} denotes transpose, and unimodular means determinant ± 1.) We say that A and B are in the same class if $B = UAU^t$ for some integral unimodular matrix A. Of course, A and B must have the same determinant to be in the same class, but this is a far from sufficient condition. The most important case is when A and B are unimodular, and let us now assume this. The number of classes in the $n \times n$ case is known to be finite, for each fixed n, but it increases very rapidly for n more than 16. It is 1 for $n \leqslant 7$, 2 for $n = 8$, 8 for $n = 16$, 297 for $n = 24$ and more than 10^7 for $n = 32$, [110, 113]. An interesting but perhaps implausible conjecture [118] is that if n is a power of 2 then the class number also is a power of 2. Of course the evidence to support this conjecture is very limited, comprising only $n = 2, 4, 8, 16$. The vague numerology present in the class numbers also seeks an explanation. Example: $24 = 2^3 \times 3$, and the order 24 class number $297 = 3^3(2^3 + 3)$. This formula, surely not an accident, suggests a combinatorial map between two seemingly distinct objects. Are there others like it?

Of course, matrices A and B in the same class rarely have the same eigenvalues, since $A \to UAU^t$ is not a similarity transformation on A.

The class number is much better understood for integral unimodular matrices that are not definite, and there are in fact powerful techniques involving matrices over p-adic integers that yield a complete classification in the not definite unimodular case. See [110, 113].

S. Friedland [111] recently discovered a new and simple to describe theorem for the positive definite case that most experts in integral quadratic forms view as surprising. It states that within each class there is a matrix with all but three of its eigenvalues equal to 1. That is, by a good choice of U, B will have all but three of its eigenvalues equal to 1. Friedland's idea that led to this theorem really is the essence of simplicity, although turning it into a proof requires much skill (but the argument is not lengthy). Starting with positive definite A, form the direct sum $(-1) \oplus A$ of A with a single -1, and then apply the powerful theory for the not definite case.

The classes of positive definite matrices split into two types: the even and the odd. Even means that each principal diagonal element of a matrix in the class is an even integer, odd that at least one principal diagonal element is odd. It is known that even definite classes exist precisely when n is a multiple of eight. For the odd classes, Friedland has conjectured [112] that his eigenvalue theorem sharpens to the following: within an odd class there is a matrix with all but two of its eigenvalues equal to 1.

The very deep theory of definite integral forms surely belongs to the high road, involving as it does p-adic numbers, the Hasse-Minkowski theory, and the theory of genera and spinor genera of forms. By appealing to the high road in just the right way, Friedland was able to obtain a very dramatic result understandable at the low road. An open question now is whether a purely elementary proof of his result can be found. Can his proof be brought down to the low road?

An approach is to look for an especially easy matrix in a given class. For $n = 8$, there are two classes, one being the principal class (containing the identity matrix), and the other [114] being the (even) class containing the circulant with first row $(2, 1, 0, -1, -1, -1, 0, 1)$. But this matrix just fails to meet Friedland's conclusion: It has four eigenvalues 1, but not the five that his theorem claims can be obtained. Another matrix in the same class thus has the five eigenvalues 1. This low road approach, however, will not go far, because some classes do not contain a circulant, or more generally, do not contain a group matrix [116]. This happens, for example, at $n = 12$.

Nevertheless, the Friedland theorem is so striking that a low road proof should be sought. Dare we conjecture that one exists?

For those who like easy to state but not easy to solve problems that are finite in scope and elementary, but not too much so, here is one. Of the 297 classes [110] at $n = 24$, exactly 24 are even (why are these two 24's the same?), meaning that the associated quadratic form represents only even integers. (One of these belongs to the famous Leech lattice, having 4 as its arithmetical minimum.) There are also exactly 15 groups of order 24. That's the data. Here's the problem: For each of these 24 classes, which of the 15 types of group matrices are to be found in it? (A group matrix is a linear combination of the matrices from the regular representation of the group.) For example, which classes contain a circulant? This finite problem has exactly $24 \times 15 = 360$ cases to look at.

Let A and B be group matrices for the same group and in the same positive definite unimodular class, so that $B = UAU^t$ for some integral unimodular matrix U. It is not given that U is a group matrix, and it need not be, but if the class is the principal one, then U may always be chosen to be a group matrix. This fact was discovered by Newman and Taussky [114] for cyclic group matrices (circulants), and was extended to solvable groups by Thompson [117], and to any finite group by Kneser (unpublished). However, it still is true for the nonprincipal classes for all groups of order 13 or less, by a case by case analysis [116]. Is it true in general? Attempts to extend Kneser's proof to the nonprincipal classes so far have not succeeded.

18. THE MATRIX VALUED NUMERICAL RANGE

The numerical range, in its many different versions, has been a rich source of challenging questions, many of which have received answers. Recently, the kth matrix valued numerical range of a matrix A was introduced [126]. It is the subset of $2k^2$ real space comprising the leading $k \times k$ principal submatrices taken from the orbit of the $n \times n$ matrix A under the action of the unitary group: $A \to UAU^*$.

For $k = 1$, this becomes the classical numerical range. The higher matrix numerical ranges became interesting when it was observed [126] that the nth numerical range is so extremely nonconvex that it never has three collinear points. So here we have a theorem, analogous to Gerschgorin's and the ovals of Cassini, involving a parameter k that permits $k = 1$ or 2, but not 3. ($k = 2$ is permitted, since trivially two matrices are collinear.) And again, the right idea leads to an insight offering further opportunity.

The original proof that the nth numerical range never contains three collinear points was an easy determinantal calculation. A better proof was found by Li and Tsing [122], who observed that the nth matrix numerical range lies on the sphere around the origin in the matrix space with the Frobenius norm $(\mathrm{tr}\, AA^*)^{1/2}$ of A as the radius. Since this sphere is rounded, it can never have three collinear points. End of proof.

However, this insightful observation suggests that convexity should be viewed differently, in a way that involves the surface of this sphere. Let us say that two points on the surface of a sphere are on a line segment if they lie on the curve that is the intersection of the sphere surface with a two dimensional plane through the sphere center and the two points. Think, for example, of great circles on the surface of a three dimensional sphere, such as meridians of longitude on the earth's surface. Two points on the sphere surface of course lie on a unique great circle. And now, with this definition of the line segment connecting two points, is it true that the nth matrix numerical range is convex?

Unfortunately, no, for the nth matrix numerical range also lies[3] on the affine hyperplane for which each point X satisfies $\operatorname{tr} X = \operatorname{tr} A$. The intersection of a sphere and a hyperplane has little chance of being convex: in three dimensions it's a small circle (as opposed to a great circle). The only conjecture that seems possible (without much hope of being true) is that the nth matrix numerical range is the intersection of a convex set on the sphere with the trace hyperplane. Specht's theorem describing unitary similarity obviously applies to the nth matrix numerical range (but not to the kth for $k < n$). For the nth, the question amounts to finding a geometric description of unitary similarity.

For the kth matrix numerical range, Li and Tsing have shown [122] that convexity in the usual sense infrequently happens, nor does the star-shapedness property often hold.

Here we have a collection of theorems, counterexamples, and conjectures, mostly counterexamples, looking for the correct idea. The correct idea almost certainly is a high level one, and finding it may bring significant new understanding to the matrix numerical range.

Another type of matrix valued numerical range has been studied by Faranek [120], evolving from a concept introduced by Arveson [119].

If the matrix valued numerical range is not convex, perhaps a scalar function acting on its convex in the complex plane. An obvious one is the pth coefficient of the characteristic polynomial acting on the matrices in the kth matrix numerical range, $p \leqslant k$. (The leading coefficient has $p = 0$.) But it was already shown by Thompson [125] using Plücker coordinates and the quadratic p-relations that for $n = 4$ and $p = k = 2$ convexity fails for a certain normal matrix A. Later Marcus [123] used Plücker coordinates and the p-relations to show that convexity fails for $n \times n$ matrices and $p = k$, again for a fixed normal matrix, although convexity does hold when $p = 1$ or when $k = n - 1$ or n (and any p). The use of Plücker coordinates turns out to be unnecessary, though. By a simple though nonobvious calculation involving the eigenvalues of a principal submatrix of a Hermitian matrix, C. K. Li and N. K. Tsing [121] showed that convexity is not true for the normal matrix $\operatorname{diag}(1 + i, 1 + i, 1, 1, \ldots, 1)$, for all values of p and k other than those just listed. Independ-

[3] I thank C. K. Li for pointing this out to me.

ently, T. Y. Tam [124] later showed that convexity fails when $1 < p = k < n - 1$ for the normal matrix $A = \mathrm{diag}(i, i, 1, 1, \ldots, 1)$. Tam's argument is that the values of the pth coefficient are a fixed multiple of $(\lambda_1 + i)(\lambda_2 + i) \cdots (\lambda_p + i)$, where $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues of the leading $p \times p$ principal submatrix of a Hermitian matrix $U^* \mathrm{diag}(-1, -1, 1, 1, \ldots, 1)\, U$, U ranging over unitary matrices. It can assume values -1 and $+1$ (easily seen), but not 0, since the λ_i are real. Thus convexity fails. According to Li and Tsing [121], the star-shapedness property fails, too.

The high road in this last example is the application of Plücker coordinates, the low the reduction to eigenvalues of Hermitian matrices.

19. INEQUALITIES WITH SUBTRACTED TERMS

A recent trend in linear algebra is the increasing presence of inequalities with subtracted terms. A favorite example is the theorem linking the diagonal elements and singular values of a matrix, already mentioned, which we state again ([136]; see also [134]). Let $d_1, \ldots, d_n$ be the diagonal elements of a matrix having singular values $s_1, \ldots, s_n$, the numbering such that $|d_1| \geqslant \cdots \geqslant |d_n|$, $s_1 \geqslant \cdots \geqslant s_n$. Then the vector $(|d_1|, \ldots, |d_n|)$ is weakly majorized by $(s_1, \ldots, s_n)$, and furthermore

$$|d_1| + \cdots + |d_{n-1}| - |d_n| \leqslant s_1 + \cdots + s_{n-1} - s_n.$$

Weak majorization and this inequality are together necessary and sufficient conditions for the existence of a matrix with prescribed diagonal elements and prescribed singular values. Inequalities like the one displayed, with a subtracted term, seem to have first appeared in the linear algebra literature in the work of Horn [130]. Others have been found [137], and there is an interesting one in the spirit of the determinantal counterpart of Gerschgorin's theorem that was found by Johnson and Newman [131]. This last reads as follows. Let A be a real matrix; let R_i^+ be the sum of the nonnegative entries in row i, and R_i^- the sum of the nonpositive entries in the same row. Then

$$|\det A| \leqslant \prod_{i=1}^{n} \max\{R_i^+, |R_i^-|\} - \prod_{i=1}^{n} \min\{R_i^+, |R_i^-|\}.$$

In the context of eigenvalues and singular values, the subtracted term inequalities reflect the properties of the root systems associated with the simple Lie structures. See [132] for the connection. The Lie approach is the high road, and the low one is the elementary devices that have been used to prove the subtracted term inequalities, and sometimes to show they are sufficient conditions for the existence of matrices with certain properties. The high road has the merit of giving entire families of theorems with one proof, as well as insight. Yet it still is true that the low road has pointed the way to the high road, and without the low road the high road probably would not have been looked at.

Here is one low road theorem that was recently found because it was needed for a higher level insight involving reflection groups [127]. A well-known theorem of J. von Neumann gives the maximum of

$$|\mathrm{tr}(UAVB)|$$

for fixed matrices A and B as U and V range over the unitary group. The maximum is [128, 129]

$$\alpha_1\beta_1 + \cdots + \alpha_n\beta_n,$$

where the α_i and β_i are the singular values of A and B in weakly decreasing order. Now let A and B be real matrices, with det A positive and det B negative. Furthermore, restrict U and V to run over $SO(n)$, the proper orthogonal group. Then the maximum of $\mathrm{tr}(UAVB)$ becomes [133]

$$\alpha_1\beta_1 + \cdots + \alpha_{n-1}\beta_{n-1} - \alpha_n\beta_n.$$

Here we have an inequality exhibiting a subtracted term, with a relatively easy low level proof, but which was conjectured owing to its need in a higher level investigation involving reflection groups.

Although not many applications have been found for the subtracted term inequalities, here is one [135]. Consider the convex hull of the $n \times n$ proper orthogonal matrices, and ask for the greatest and least values of the determinant on this set. The maximum is easily found to be 1. The minimum can be negative, but never as low as

$$-e^{-2} = -0.13534\ldots.$$

This bound, although never achieved, is still sharp in that it is the limit of the achieved minima as n approaches infinity.

20. FURTHER USES OF THE COMPUTER

Use of the computer has played an important role in the study of the numerical range, in the hands of C. R. Johnson, M. Marcus, and some others [139–145]. Indeed, the graphical output the computer provides has repeatedly led to numerical range conjectures, many of which were eventually proved. However, the algorithms for generating graphical output seem to depend on the known convexity of the various numerical ranges. What happens for those situations in which convexity is not true, or not known? A good example is the numerical range for matrices with quaternion entries, known to be nonconvex in general [138]. Here we have a situation in which the desire to generate computer output may force the discovery of new ideas to make the output possible. Thus the availability of a powerful computational tool leads not to an answer but to a question.

The matrix numerical range offers another example: how can graphical output from a computer display an object lying on a sphere in a high dimensional space? Can a plethora of two dimensional sections through a multidimensional object give any useful understanding? We seem to need a concept of what a good graphical description of a multidimensional object is. Without it, a lack of graphics for the matrix numerical range may inhibit its study very severely. It would be foolish to predict whether the future holds an answer to this computational problem, since it involves two aspects: (1) finding an algorithm (for computing the matrix numerical range) that works provided enough computer power is available, and (2) overcoming the barrier that increasing dimension entails in the provision of adequate computing power.

The preparation of this paper was supported in part by the National Science Foundation, U.S.A.

I am indebted to D. Carlson, S. Friedland, C. R. Johnson, and C. K. Li for their comments on the first draft of this paper.

The chosen examples largely reflect the author's experience. Equally good (often better) examples abound in the literature, and the reader is invited to substitute his favorite examples for any of those described above.

REFERENCES

1. Introduction

1 R. C. Thompson, Core linear algebra, presented at First SIAM Linear Algebra Conference, Raleigh, N.C., 1983.

2 *Mathematical Reviews*, 1–58 and 80–89, Amer. Math. Soc., Providence, 1940–1989.

4. The Numerical Range

3 M. D. Choi, C. Laurie, H. Radjavi, and P. Rosenthal, On the congruence numerical range and related functions of matrices, *Linear and Multilinear Algebra* 22:1–5 (1987).

4 P. R. Halmos, *A Hilbert Space Problem Book*, 2nd ed., Springer-Verlag, New York, 1982.

5 G. Hughes, A note on the shape of the generalized numerical range, *Linear and Multilinear Algebra* 26:43–47 (1990).

6 Y. T. Poon, Another proof of a result of Westwick, *Linear and Multilinear Algebra* 9:35–37 (1980).

7 T. Y. Tam, Note on a paper of Thompson: The congruence numerical range, *Linear and Multilinear Algebra* 17:107–115 (1985).

8 R. C. Thompson, A note on normal matrices, *Canad. J. Math.* 15:220–225 (1963).

9 R. C. Thompson, The congruence numerical range, *Linear and Multilinear Algebra* 8:197–206 (1980).

10 N. K. Tsing, On the shape of the generalized numerical ranges, *Linear and Multilinear Algebra*, 10:179–185 (1981).

11 R. Westwick, A theorem on the numerical range, *Linear and Multilinear Algebra* 2:311–315 (1975).

5. Similarity Invariants of Principal Submatrices

12 P. Y. Cheng, Interlacing Divisibility of Similarity Invariant Factors and Its Application, Univ. of California, Davis, 1978.

13 G. N. de Oliveira, Conference report (Fourth Dublin Conference), *Linear Algebra Appl.* 117:146–150 (1989).

14 I. Gohberg, M. A. Kaashoek, F. van Schagen, Eigenvalues of completions of matrices, *Linear and Multilinear Algebra* 25:55–70 (1989).

15 E. Marques de Sa, Embedding conditions for λ matrices, *Linear Algebras Appl.* 25:33–50 (1979).

16 E. Marques de Sa, A convexity lemma on the interlacing inequalities for invariant factors, *Linear Algebra Appl.* 109:107–113 (1988).

17 R. C. Thompson, Interlacing inequalities for invariant factors, *Linear Algebra Appl.* 24:1–31 (1979).

18 H. Wimmer, Existenzsätze in der Theorie der Matrizen und lineare Kontroltheorie, *Monatsh. Math.* 78:256–263 (1974).

19 I. Zaballa, Matrices with prescribed rows and invariant factors, *Linear Algebra Appl.* 87:113–146 (1987).

20 I. Zaballa, Interlacing inequalities and control theory, *Linear Algebra Appl.* 101:9–31 (1988).

21 I. Zaballa, Matrices with prescribed invariant factors and off-diagonal submatrices, *Linear and Multilinear Algebra* 25:39–54 (1989).

6. Commutators

22 A. A. Albert and B. Muckenhaupt, On matrices of trace zero, *Michigan Math. J.* 4:1-3 (1957).

23 A. Brown and C. Pearcy, Structure of commutators of operators, *Ann. Math.* 82:112-127 (1965).

24 A. Brown and C. Pearcy, Multiplicative commutators of operators, *Canad. J. Math.* 18:737-749 (1966).

25 A. Brown and C. Pearcy, Operators of the form $PAQ - QAP$, *Canad. J. Math.* 20:853-859 (1968).

26 D. Estes, Scalar matrices as multiplicative commutators having prescribed determinants for the variables, *Linear and Multilinear Algebra* 8:213-217 (1980).

27 S. Friedland, A characterization of normal operators, *Israel J. Math.* 42:235-240 (1982).

28 H. Radjavi, Structure of $A^*A - AA^*$, *J. Math. Mech.* 16:19-26 (1966).

29 K. Shoda, Einige Sätze über Matrizen, *Japan. J. Math.* 13:361-365 (1936).

30 R. C. Thompson, On matrix commutators, *J. Wash. Acad. Sci.* 48:306-307 (1958).

31 R. C. Thompson, Commutators in the special and general linear groups, *Trans. Amer. Math. Soc.* 101:16-33 (1961).

32 R. C. Thompson, Commutators of matrices with coefficients from the field of two elements, *Duke Math. J.* 29:367-374 (1962).

33 R. C. Thompson, On matrix commutators, *Portugal. Math.* 21:143-153 (1962).

34 R. C. Thompson, Commutators of matrices with prescribed determinant, *Canad. J. Math.* 20:203-221 (1968).

35 R. C. Thompson, Some matrix factorization theorems, *Pacific J. Math.* 33:763-810 (1970).

36 L. Vaserstein and E. Wheland, Commutators and companion matrices over rings of stable rank 1, *Linear Algebra Appl.* 142:263-277 (1990).

37 Pei Yuan Wu, The operator factorization problems, *Linear Algebra Appl.* 117:35-63 (1989).

7. The Triangle Inequality

38 C. A. Akemann, J. Anderson, and G. K. Pederson, Triangle inequalities in operator algebras, *Linear and Multilinear Algebra* 11:167-178 (1982).

39 E. Seiler and B. Simon, An inequality among determinants, *Proc. Nat. Acad. Sci. U.S.A.* 72:3277-3278 (1975).

40 R. C. Thompson, Convex and concave functions of singular values of matrix sums, *Pacific J. Math.* 66:285-290 (1976).

41 R. C. Thompson, The case of equality in the matrix valued triangle inequality, *Pacific J. Math.* 82:279-280 (1979).

42 R. C. Thompson, The matrix valued triangle inequality: Quaternion version, *Linear and Multilinear Algebra* 25:85–91 (1989).

8. The Facial Structure of the Unit Ball

43 C. A. Akemann, private conversation.
44 W. So, Facial structure of Schatten p-norms, *Linear and Multilinear Algebra* 27:207–212 (1990).

9. The Gerschgorin Circle Theorem

45 A. Brauer, Limits for the characteristic roots of a matrix II, *Duke Math. J.* 14:21–26 (1947).
46 R. Brualdi, Matrices, eigenvalues, and directed graphs, *Linear and Multilinear Algebra* 11:143–165 (1982).
47 R. A. Horn and C. R. Johnson, *Matrix Analysis*, Vol. 1, Cambridge U.P., Cambridge, 1985.
48 C. R.Johnson, private communication.

10. Matrices, Graphs, Inertia, Number Theory

49 W. Barrett, C. R. Johnson, and M. Lundquist, Determinantal formulae for matrix completions associated with chordal graphs, *Linear Algebra Appl.* to appear.
50 N. Cohen, C. R. Johnson, L. Rodman, and H. Woerdman, Ranks of completions of partial matrices, *Integral Equations Operator Theory*, to appear.
51 D. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs—Theory and Application*, Academic, New York, 1982.
52 D. Cvetković, *Recent Results in the Theory of Graph Spectra*, North Holland, 1988.
53 S. Friedland, Quadratic forms and the graph isomorphism problem, preprint.
54 R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowitz, Positive definite completions of partial Hermitian matrices, *Linear Algebra Appl.* 58:109–124 (1984).
55 C. R. Johnson, R. B. Kellog, and A. B. Stephens, Complex eigenvalues of a nonnegative matrix with a specified graph II, *Linear and Multilinear Algebra*, 7:129–143 (1979); Erratum, 8:171 (1979).
56 C. R. Johnson and L. Rodman, Inertia possibilities for completions of partial Hermitian matrices, *Linear and Multilinear Algebra* 16:179–195 (1984).
57 C. R. Johnson and L. Rodman, Completions of partial matrices to contractions, *J. Funct. Anal.* 69:260–267 (1986).
58 C. R. Johnson and L. Rodman, Completion of Toeplitz partial contractions, *SIAM J. Matrix Anal.* 9:159–167 (1988).

59 C. R. Johnson and L. Rodman, Chordal inheritance principles and positive definite completions of partial matrices over function rings, in *Contributions to Operator Theory and Its Applications*, Birkhäuser, Basel, 1988, pp. 107–127.
60 C. R. Johnson, Combinatorial matrix analysis: An overview, *Linear Algebra Appl.* 107:3–15 (1988).
61 C. R. Johnson, Matrix completion problems: A survey, in *Matrix Theory and Applications*, Proc. Sympos. Appl. Math. 40, Amer. Math. Soc., Providence, 1990.
62 R. Merris, Special issue on algebraic graph theory, *Linear and Multilinear Alebra* 29:1–117 (1990).

11. Power Embeddings and Dilations

63 E. Egervary, On the contractive linear transformations of n-dimensional vector space, *Acta Sci. Math. (Szeged)* 15:178–182 (1954).
64 B. Sz-Nagy and C. Fois, *Harmonic Analysis of Operators on Hilbert Space*, North Holland, Amsterdam, 1970.
65 R. C. Thompson and Chi-Chin Tung Kuo, Doubly stochastic, unitary, unimodular, and complex orthogonal power embeddings, *Acta Sci. Math. (Szeged)* 44:345–357 (1982).

12. The Schubert Calculus

66 J. E. Dodd, *The Ideas of Particle Physics, an Introduction for Scientists*, Cambridge U.P., 1984, p. 57.
67 J. Hersch and B. P. Zwahlen, Évaluations per défaut pour une somme quelconque de valeurs propres γ_k d'un opérateur $C = A + B$ à l'aide de valeurs propres α_i de A et β_j de B, *C. R. Acad. Sci. Paris*, 1962, pp. 1559–1561.
68 W. V. D. Hodge and D. Pedoe, *Methods of Algebraic Geometry*, Cambridge, 1952, Chapter XIV.
69 A. Horn, Eigenvalues of a sum of Hermitian matrices, *Pacific J. Math.* 12:225–241 (1962).
70 S. Johnson, The Schubert Calculus and Eigenvalue Inequalities for Sums of Hermitian Matrices, Univ. of California, Santa Barbara, 1979.
71 S. L. Kleiman, Problem 15: Rigorous foundations of Schubert's enumerative calculus, in *Mathematical Developments Arising from Hilbert's Problems*. Proc. Sympos. Pure Math. 28, Amer. Math Soc., Providence, 1976, pp. 445–482.
72 I. Macdonald, *Symmetric Functions and Hall Polynomials*, Clarendon, Oxford, 1979.
73 R. C. Riddell, Minimax problems on Grassmann manifolds, *Adv. in Math.* 54:107–199 (1984).

74 R. C. Thompson, On a construction of B. P. Zwahlen, *Linear and Multilinear Algebra* 1:309–325 (1974).
75 H. Wielandt, An extremum property of sums of eigenvalues, *Proc. Amer. Math. Soc.* 6:106–110 (1955).
76 B. P. Zwahlen, Über die Eigenwerte der Summe zweier selbstadjungierter Operatoren, *Comment. Math. Helv.* 40:81–116 (1966).

13. The Spectrum of a Sum of Hermitian Matrices

77 F. Berezin and I. Gel'fand, Some remarks on spherical functions on symmetric Riemannian manifolds, *Trudy Moskow Mat. Obshch.* 5:311–351 (1956).
78 J. Day, W. So, and R. C. Thompson, manuscript.
79 A. Horn, Eigenvalues of a sum of Hermitian matrices, *Pacific J. Math.* 12:225–241 (1962).
80 B. V. Lidskii, Spectral polyhedron of a sum of two Hermitian matrices, *Functional Anal. Appl.* 10:76–77 (in Russian), 139–140 (in English) (1982).
81 V. B. Lidskii, Characteristic numbers of the sum and product of symmetric matrices, *Dokl. Akad. Nauk SSSR* 75:769–772 (1950).
82 R. C. Thompson and L. Freede, On the eigenvalues of a sum of Hermitian matrices, *Linear Algebra Appl.* 4:369–376 (1971).
83 R. C. Thompson, Smith invariants of a product of integral matrices, in *Comtemp. Math.* 47, Amer. Math. Soc., 1985, pp. 401–435.
84 H. Wielandt, An extremum property of sums of eigenvalues, *Proc. Amer. Math. Soc.* 6:106–110 (1955).

14. The Hadamard-Schur Product

85 T. Ando, R. Horn, and C. R. Johnson, The singular values of a Hadamard product: A basic inequality, *Linear and Multilinear Algebra* 21:345–365 (1987).
86 M. Eaton, *Lectures on Topics in Probability Inequalities*, 35, Centrum voor Wiskunde en Informatica, 1987.
87 R. Horn, The Hadamard product, in *Matrix Theory and Applications*, Proc. Sympos. Appl. Math., 40, Amer. Math. Soc., 1990, pp. 87–169.
88 R. Horn and C. R. Johnson, Hadamard and conventional submultiplicativity for unitarily invariant norms on matrices, *Linear and Multilinear Algebra* 20:91–106 (1987).
89 R. Horn and C. R. Johnson, *Matrix Analysis*, Vol. 2, Cambridge U.P., Cambridge, 1991.
90 C. R. Johnson and R. Bapat, A weak multiplicative majorization conjecture for Hadamard products, *Linear Algebra Appl.* 104:246–247 (1988).
91 R. Horn and R. Mathias, An analog of the Cauchy-Schwarz inequality of Hadamard products and unitarily invariant norms, preprint.

92 C. R. Johnson and H. Shapiro, Mathematical aspects of the relative gain array ($A \circ A^{-T}$), *SIAM J. Algebraic Discrete Methods* 7:627–644 (1986).
93 M. Perlman and M. Eaton, Reflection groups, generalized Schur functions, and the geometry and majorization, *Ann. Probab.* 5:829–860 (1977).
94 R. C. Thompson and S. Therianos, The singular values of matrix products, *Scripta Math.* 29:99–110 (1973).
95 R. C. Thompson, Singular values, diagonal elements, and convexity, *SIAM J. Appl. Math.* 32:39–63 (1977).
96 R. C. Thompson, Singular values and diagonal elements of complex symmetric matrices, *Linear Algebra Appl.* 26:65–106 (1979).

15. The Exponential Function

97 J. Cohen, S. Friedland, T. Kato, and F. Kelly, Eigenvalue inequalities for products of matrix exponentials, *Linear Algebra Appl.* 45:55–95 (1982).
98 J. Cohen, Spectral inequalities for matrix exponentials, *Linear Algebra Appl.* 111:25–28 (1988).
99 J. Day, W. So, and R. C. Thompson, Some properties of the Campbell-Baker-Hausdorff series, *Linear and Multilinear Algebra* 29, to appear.
100 Mathematica, Wolfram Research Inc.
101 B. Mityagin, Announcement, Calgary Operator Theory Conference, Aug. 1988.
102 F. Rovière, Invariant analysis and contractions of symmetric spaces, part I, manuscript, 1988.
103 F. Rovière, private communication.
104 W. So and R. C. Thompson, Products of exponentials of Hermitian and complex symmetric matrices, *Linear and Multilinear Algebra*, to appear.
105 W. So and R. C. Thompson, manuscript.
106 R. C. Thompson, Proof of a conjectured exponential formula, *Linear and Multilinear Algebra* 19:187–197 (1986).
107 R. C. Thompson, Special cases of a matrix exponential formula, *Linear Algebra Appl.* 107:283–292 (1988).
108 R. C. Thompson, J. Richards, and C. Papadapoulos, work in progress.

16. The Exponential Function and Commutativity

109 E. M. E. Wermuth, Two remarks on matrix exponentials, *Linear Algebra Appl.* 117:127–132 (1989).

17. Integral Quadratic Forms

110 J. Conway and N. J. A. Sloane, *Sphere Packings, Lattices, and Groups*, Grundlehren Math. Wiss. 290, Springer-Verlag, New York, 1988.

111 S. Friedland, Normal forms for definite integer unimodular quadratic forms, *Proc. Amer. Math. Soc.* 106:917–921 (1989).
112 S. Friedland, Quadratic forms and the graph isomorphism problem, preprint.
113 J. Milnor and D. Husemoler, *Symmetric Bilinear Forms*, Ergeb. Math. 73, 1973.
114 M. Newman and O. Taussky, Classes of positive definite unimodular circulants, *Canad. J. Math.* 9:71–73 (1956).
115 M. Newman and O. Taussky, A generalization of the normal basis in abelian algebraic number fields, *Pacific J. Math.* 9:85–91 (1956).
116 R. C. Thompson, Classes of definite group matrices, *Pacific J. Math.* 17:175–190 (1966).
117 R. C. Thompson, Unimodular group matrices with rational integers as elements, *Pacific J. Math.* 14:719–726 (1964).
118 R. C. Thompson, unpublished conjecture.

18. The Matrix Valued Numerical Range
119 W. Arveson, Subalgebras of C^* algebras II. *Acta Math. (Stockholm)* 128:279–308 (1972).
120 D. Faranek, presented at Auburn (Alabama) Matrix Conference, Mar. 1990.
121 C. K. Li and N. K. Tsing, The numerical range of derivations, *Linear Algebra Appl.* 119:97–121 (1989).
122 C. K. Li and N. K. Tsing, On the kth matrix numerical range, *Linear and Multilinear Algebra* 28:229–239 (1991).
123 M. Marcus, Derivations, Plücker relations, and the numerical range, *Indiana Univ. Math. J.* 22:1137–1149 (1973).
124 T. Y. Tam, Another proof of a result of Marcus without Plücker relations, *Linear and Multilinear Algebra* 29, to appear.
125 R. C. Thompson, A note on normal matrices, *Canad. J. Math.* 15:220–225 (1963).
126 R. C. Thompson, The matrix numerical range, *Linear and Multilinear Algebra* 21:313–314 (1987).

19. Inequalities with Subtracted Terms
127 L. C. Grove and C. T. Benson, *Finite Reflection Groups*, 1971; Springer-Verlag, New York, 1985.
128 Ky Fan, Maximum properties and inequalities for the eigenvalues of completely continuous operators, *Proc. Nat. Acad. Sci. U.S.A.* 37:760–766 (1951).
129 I. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Odessa, 1964; Amer. Math. Soc. Transl. 18, 1969.

130 A. Horn, Doubly stochastic matrices and the diagonal of a rotation matrix, *Amer. J. Math.* 76:620–630 (1954).
131 C. R. Johnson and M. Newman, A surprising determinantal inequality for matrices, *Math. Ann.* 247:179–185 (1980).
132 B. Kostant, On convexity, the Weyl group and the Iwasawa decomposition, *Ann. Sci. École Norm. Sup.* 6:413–455 (1973).
133 H. Miranda and R. C. Thompson, manuscript.
134 F. Y. Sing, Some results on matrices with prescribed diagonal elements and singular values, *Canad. Math. Bull.* 19:89–92 (1976).
135 R. C. Thompson, Extreme determinants on the convex hull of matrices with prescribed singular values, *Linear and Multilinear Algebra* 3:15–17 (1975).
136 R. C. Thompson, Singular values, diagonal elements, and convexity, *SIAM J. Appl. Math.* 32:39–63 (1977).
137 R. C. Thompson, Singular values and diagonal elements of complex symmetric matrices, *Linear Algebra Appl.* 26:65–106 (1979).

20. Further Uses of the Computer

138 Y. H. Au-Yeung, On the convexity of the numerical range in quaternionic Hilbert spaces, *Linear and Multilinear Algebra* 16:93–100 (1984).
139 C. Johnson, Numerical location of the field of values, *Linear and Multilinear Algebra* 3:9–14 (1975).
140 C. R. Johnson, Numerical determination of the field of values of a general complex matrix, *SIAM J. Numer. Anal.* 15:595–602 (1978).
141 C. R. Johnson, Numerical ranges of principal submatrices, *Linear Algebra Appl.* 37:23–34 (1981).
142 C. K. Li, C. H. Sung, and N. K. Tsing, c-Convex matrices: Characterizations, inclusion relations and normality, *Linear and Multilinear Algebra* 25:275–287 (1989).
143 M. Marcus and B. N. Shure, The numerical range of certain 0,1 matrices, *Linear and Multilinear Algebra* 7:111–120 (1979).
144 M. Marcus and C. Pesce, Computer generated numerical ranges and some resulting theorems, *Linear and Multilinear Algebra* 20:121–157 (1987).
145 J. C. Wang, M. K. H. Fan, and A. L. Tits, Structured singular value and geometric properties of the m-form numerical range, in *Proceedings of the 1987 MTNS* (R. Saeks, C. Byrne, and C. Martin, Eds.), North Holland, 1988.

Received 16 April 1990; final manuscript accepted 13 March 1991