of *Contemp. Math.*, Amer. Math. Soc., 1999, pp. 375–393.

Jacob E. Goodman
Richard Pollack

MSC 1991: 52A35

GERSHGORIN THEOREM, *Gerschgorin theorem*, *Geršgorin theorem* – Given a complex $(n \times n)$-matrix, $A = [a_{i,j}]$, with $n \geq 2$, then finding the eigenvalues of A is equivalent to finding the n zeros of its associated *characteristic polynomial*

$$p_n(z) := \det\{zI - A\},$$

where I is the identity $(n \times n)$-matrix (cf. also **Matrix**; **Eigen value**). But for n large, finding these zeros can be a daunting problem. Is there an 'easy' procedure which estimates these eigenvalues, without having to explicitly form the characteristic polynomial $p_n(z)$ above and then to find its zeros? This was first considered in 1931 by the Russian mathematician S. Gershgorin, who established the following result [2]. If $\Delta_\delta(\alpha) := \{z \in \mathbf{C} \colon |z - \alpha| \leq \delta\}$ denotes the closed complex disc having centre α and radius δ, then Gershgorin showed that for each eigenvalue λ of the given complex $(n \times n)$-matrix $A = [a_{i,j}]$ there is a positive integer i, with $1 \leq i \leq n$, such that $\lambda \in G_i(A)$, where

$$G_i(A) := \Delta_{r_i(A)}(a_{i,i}) \tag{1}$$

with

$$r_i(A) := \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{i,j}| .$$

($G_{r_i}(A)$ is called the ith *Gershgorin disc* for A.) As this is true for each eigenvalue λ of A, it is evident that if $\sigma(A)$ denotes the set of all eigenvalues of A, then

$$\sigma(A) \subseteq \bigcup_{i=1}^{n} G_i(A). \tag{2}$$

Indeed, let λ be any eigenvalue of $A = [a_{i,j}]$, so that there is a complex vector $\mathbf{x} = [x_1 \dots x_n]^T$, with $\mathbf{x} \neq \mathbf{O}$, such that $A\mathbf{x} = \lambda\mathbf{x}$. As $\mathbf{x} \neq \mathbf{O}$, then $\max_{1 \leq j \leq n} |x_j| > 0$, and there is an i, with $1 \leq i \leq n$, such that $|x_i| = \max_{1 \leq j \leq n} |x_j|$. Taking the ith component of $A\mathbf{x} = \lambda\mathbf{x}$ gives $\sum_{j=1}^{n} a_{i,j}x_j = \lambda x_i$, or equivalently

$$(\lambda - a_{i,i})\, x_i = \sum_{\substack{j=1\\ j\neq i}}^{n} a_{i,j}x_j .$$

On taking absolute values in the above expression and using the triangle inequality, this gives

$$|\lambda - a_{i,i}| \cdot |x_i| \leq \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{i,j}| \cdot |x_j| \leq r_i(A) \cdot |x_i| , \tag{3}$$

the last inequality following from the definition of $r_i(A)$ in (1) and the fact that $|x_j| \leq |x_i|$ for all $1 \leq j \leq n$. Dividing through by $|x_i| > 0$ in (3) gives that $\lambda \in G_i(A)$.

In the same paper, Gershgorin also established the following interesting result: If the n discs $G_i(A)$ of (2) consist of two non-empty disjoint sets S and T, where S consists of the union of, say, k discs and T consists of the union of the remaining $n-k$ discs, then S contains exactly k eigenvalues (counting multiplicities) of A, while T contains exactly $n-k$ eigenvalues of T. (The proof of this depends on the fact that the zeros of the characteristic polynomial $p_n(z)$ vary continuously with the entries $a_{i,j}$ of A.)

One of the most beautiful results in this area, having to do with the sharpness of the inclusion of (2), is a result of O. Taussky [4], which depends on the following use of directed graphs (cf. also **Graph, oriented**). Given a complex $(n \times n)$-matrix $A = [a_{i,j}]$, with $n \geq 2$, let $\{P_i\}_{i=1}^{n}$ be n distinct points, called vertices, in the plane. Then, for each $a_{i,j} \neq 0$, let $\overrightarrow{P_iP_j}$ denote an arc from vertex i to vertex j. The collection of all these arcs defines the *directed graph* of A. Then the matrix $A = [a_{i,j}]$, with $n \geq 2$, is said to be *irreducible* if, given any distinct vertices i and j, there is a sequence of abutting arcs from i to j, i.e.,

$$\overrightarrow{P_iP_{\ell_1}}, \overrightarrow{P_{\ell_1}P_{\ell_2}}, \dots, \overrightarrow{P_{\ell_m}P_{\ell_{m+1}}},$$

where $\ell_{m+1} = j$.

Taussky's theorem is this. Let $A = [a_{i,j}]$ be any irreducible complex $(n \times n)$-matrix, with $n \geq 2$. If λ is an eigenvalue of A which lies on the boundary of the union of the Gershgorin discs of (2), then λ lies on the boundary of each Gershgorin circle, i.e., from (1) it follows that

$$|\lambda - a_{i,i}| = r_i(A) \quad \text{for each } 1 \leq i \leq n.$$

Next, there is related work of A. Brauer [1] on estimating the eigenvalues of a complex $(n \times n)$-matrix $(n \geq 2)$, which uses Cassini ovals instead of discs. For any integers i and j $(1 \leq i, j \leq n)$ with $i \neq j$, the (i,j)th *Cassini oval* is defined by (cf. also **Cassini oval**)

$$K_{i,j}(A) := \tag{4}$$

$$:= \left\{z \in \mathbf{C} \colon |z - a_{i,i}| \cdot |z - a_{j,j}| \leq r_i(A) \cdot r_j(A)\right\}.$$

Then *Brauer's theorem* is that, for any eigenvalue λ of A, there exist i and j, with $i \neq j$, such that $\lambda \in K_{i,j}(A)$, and this now gives the associated eigenvalue inclusion

$$\sigma(A) \subseteq \bigcup_{\substack{i,j=1\\ i\neq j}}^{n} K_{i,j}(A). \tag{5}$$

Note that there are now $n(n-1)/2$ such Cassini ovals in (5), as opposed to the n Gershgorin discs in (2). But it is equally important to note that the eigenvalue inclusions that every $d+1$ sets of $\mathcal{A}$ are met by an oriented hyperplane consistently with the $(d-1)$-ordering induced on the corresponding points of S.

For this to hold, moreover, it is sufficient to find an acyclic rank-d oriented **matroid** structure on $\mathcal{A}$ whose $(d+1)$-tuples satisfy the consistency condition of the theorem. (For the notion of *oriented matroid*, which can be thought of as a 'locally realizable' generalization of the order type of a set of points, see also [2].)

This theorem was subsequently generalized in several directions, the most comprehensive statement (which subsumes intermediate results of M. Katchalski and of Pollack and R. Wenger) being:

Anderson–Wenger theorem. ([1]) Let $\mathcal{A}$ be a finite collection of connected sets in $\mathbf{R}^d$. $\mathcal{A}$ has a hyperplane transversal if and only if for some k, $0 \leq k < d$, there exists a rank-$(k+1)$ acyclic oriented matroid structure on $\mathcal{A}$ such that every $k+2$ members of $\mathcal{A}$ are met by an oriented k-flat consistently with that oriented matroid structure.

Other directions. An effort to understand intermediate-dimensional transversals leads to considering the set of all k-transversals to a collection of convex sets. While there are as yet (2000) very few non-trivial results about transversals of dimensions between 1 and $d-1$, there is a good deal known about the structure of these complete sets of transversals to collections of convex sets. It turns out that these subsets of the 'affine Grassmannian' themselves behave very much like convex point sets. Although they need not be connected, they nevertheless share many properties with convex sets, such as being defined by a convex hull operator satisfying the anti-exchange property that commutes with non-singular affine transformations, and satisfying the Kreĭn–Mil'man theorem [7] (cf. also **Locally convex space**).

Other streams in geometric transversal theory include:

- Gallai-type theorems, in which a Helly-type hypothesis leads to the conclusion that several transversals cover the entire collection (see [5], [18]);
- the more general theorems of (p, q) type, in which a hypothesis of the form 'for every choice of p sets from the collection, some q have a common transversal' leads to a Gallai-type conclusion (see [5], [18]);
- generalizations from transversal flats to transversal curves and surfaces (see [8]);
- ongoing work of L. Montejano on a topological generalization of Hadwiger's transversal theorem and on the related notion of a 'separoid';
- the problem of bounding the number of geometric permutations of a collection of n convex bodies in $\mathbf{R}^d$, and its generalization (via order types) to k-transversals (see [19]);
- algorithmic geometric transversal theory, a branch of **computational geometry** (see [19]).

Surveys include [4], [5], [8], [13], [18], [19], where many other references can be found.

References

[1] ANDERSON, L., AND WENGER, R.: 'Oriented matroids and hyperplane transversals', *Adv. Math.* **119** (1996), 117–125.

[2] BJÖRNER, A., LAS VERGNAS, M., WHITE, N., STURMFELS, B., AND ZIEGLER, G.M.: *Oriented matroids*, Cambridge Univ. Press, 1993.

[3] DANZER, L.: 'Über ein Problem aus der kombinatorischen Geometrie', *Archiv Math.* **8** (1957), 347–351.

[4] DANZER, L., GRÜNBAUM, B., AND KLEE, V.: 'Helly's theorem and its relatives', in V. KLEE (ed.): *Convexity*, Vol. 7 of *Proc. Symp. Pure Math.*, Amer. Math. Soc., 1963, pp. 101–180.

[5] ECKHOFF, J.: 'Helly, Radon, and Carathéodory type theorems', in P.M. GRUBER AND J.M. WILLS (eds.): *Handbook of Convex Geometry*, North-Holland, 1993, pp. 389–448.

[6] GOODMAN, J.E., AND POLLACK, R.: 'Hadwiger's transversal theorem in higher dimensions', *J. Amer. Math. Soc.* **1** (1988), 301–309.

[7] GOODMAN, J.E., AND POLLACK, R.: 'Foundations of a theory of convexity on affine Grassmann manifolds', *Mathematika* **42** (1995), 305–328.

[8] GOODMAN, J.E., POLLACK, R., AND WENGER, R.: 'Geometric transversal theory', in J. PACH (ed.): *New Trends in Discrete and Computational Geometry*, Vol. 10: Algorithms and Combinatorics, Springer, 1993, pp. 163–198.

[9] GRÜNBAUM, B.: 'On common transversals.', *Archiv Math.* **9** (1958), 465–469.

[10] GRÜNBAUM, B.: 'Common transversals for families of sets.', *J. London Math. Soc.* **35** (1960), 408–416.

[11] HADWIGER, H.: 'Über einen Satz Hellyscher Art.', *Archiv Math.* **7** (1956), 377–379.

[12] HADWIGER, H.: 'Über Eibereiche mit gemeinsamer Treffgeraden', *Portugal. Math.* **16** (1957), 23–29.

[13] HADWIGER, H., DEBRUNNER, H., AND KLEE, V.: *Combinatorial geometry in the plane*, Holt, Rinehart & Winston, 1964.

[14] HELLY, E.: 'Über Mengen konvexen Körper mit gemeinschaftlichen Punkten', *Jahresber. Deutsch. Math. Verein.* **32** (1923), 175–176.

[15] SANTALÓ, L.: 'Un teorema sobre conjuntos de paralelepipedos de aristas paralelas', *Publ. Inst. Mat. Univ. Nac. Litoral* **2** (1940), 49–60.

[16] TVERBERG, H.: 'Proof of Grünbaum's conjecture on common transversals for translates', *Discr. Comput. Geom.* **4** (1989), 191–203.

[17] VINCENSINI, P.: 'Figures convexes et variétés linéaires de l'espace euclidien à n dimensions', *Bull. Sci. Math.* **59** (1935), 163–174.

[18] WENGER, R.: 'Helly-type theorems and geometric transversals', in J.E. GOODMAN AND J.O'ROURKE (eds.): *Handbook of Discrete and Computational Geometry*, CRC, 1997, pp. 63–82 (Chap. 4).

[19] WENGER, R.: 'Progress in geometric transversal theory', in B. CHAZELLE, J.E. GOODMAN, AND R. POLLACK (eds.): *Advances in Discrete and Computational Geometry*, Vol. 223

of (2) and (5) use the exact same data from the matrix $A = [a_{i,j}]$, i.e., $\{a_{i,i}\}_{i=1}^n$ and $\{r_i(A)\}_{i=1}^n$. So, which of the eigenvalue inclusions of (2) and (5) is smaller and hence better? It turns out that

$$\bigcup_{\substack{i,j=1\\ i\neq j}}^{n} K_{i,j}(A) \subseteq \bigcup_{i=1}^{n} G_i(A), \tag{6}$$

for any complex $(n \times n)$-matrix A, so that the Cassini ovals are always at least as good as the Gershgorin discs. (The result (6) was known to Brauer, but was somehow neglected in the literature.)

Finally, as both eigenvalue inclusions (2) and (5) depend only on the row sums $r_i(A)$, it is evident that these inclusions apply not to just the single matrix A, but to a whole class of $(n \times n)$-matrices, namely,

$$\Omega(A) :=$$
$$:= \{B = [b_{i,j}] \colon b_{i,i} = a_{i,i}, \text{ and } r_i(B) = r_i(A),\ 1 \leq i \leq n\}.$$

Thus,

$$\sigma(B) \subseteq \bigcup_{\substack{i,j=1\\ i\neq j}}^{n} K_{i,j}(A) \subseteq \bigcup_{i=1}^{n} G_i(A)$$

for each B in $\Omega(A)$. Then, if $\sigma(\Omega(A))$ denotes the set of all eigenvalues of all B in $\Omega(A)$, it follows that

$$\sigma(\Omega(A)) \subseteq \bigcup_{\substack{i,j=1\\ i\neq j}}^{n} K_{i,j}(A) \subseteq \bigcup_{i=1}^{n} G_i(A). \tag{7}$$

How sharp is the first inclusion of (7)? It was shown in 1999 by R.S. Varga and A. Krautstengl [7] that

$$\sigma(\Omega(A)) = \begin{cases} \text{boundary of } K_{1,2}(A) & n = 2; \\ \bigcup_{\substack{i,j=1\\ i\neq j}}^{n} K_{i,j}(A) & n \geq 3. \end{cases} \tag{8}$$

Thus, for $n \geq 3$, it can be said that the Cassini ovals give 'perfect' results.

Gershgorin's discs and Brauer's Cassini ovals are mentioned in [5], [3]. A more detailed treatment of these topics can be found in [6].

References

[1] Brauer, A.: 'Limits for the characteristic roots of a matrix', *Duke Math. J.* **13** (1946), 387–395.

[2] Gerschgorin, S.: 'Ueber die Abgrenzung der Eigenwerte einer Matrix', *Izv. Akad. Nauk. SSSR Ser. Mat.* 1 (1931), 749–754.

[3] Horn, R.A., and Johnson, C.R.: *Matrix analysis*, Cambridge Univ. Press, 1985.

[4] Taussky, O.: 'Bounds for the characteristic roots of matrices', *Duke Math. J.* **15** (1948), 1043–1044.

[5] Varga, R.S.: *Matrix iterative analysis*, second ed., Springer, 2000.

[6] Varga, R.S.: *Geršgorin and his circles*, Springer, to appear.

[7] Varga, R.S., and Krautstengl, A.: 'On Geršgorin-type problems and ovals of Cassini', *Electron. Trans. Numer. Anal.* 8 (1999), 15–20.

Richard S. Varga

MSC1991: 15A18

Gleason–Kahane–Żelazko theorem – Let F be a non-zero linear and multiplicative **functional** on a complex **Banach algebra** $\mathcal{A}$ with a unit e, and let $\mathcal{A}^{-1}$ denote the set of all invertible elements of $\mathcal{A}$. Then $F(e) = 1$, and for any $a \in \mathcal{A}^{-1}$ one has $F(a) \neq 0$. A.M. Gleason [1] and, independently, J.P. Kahane and W. Żelazko [5], [6] proved that the property characterizes multiplicative functionals: If F is a linear functional on a complex unital Banach algebra $\mathcal{A}$ such that $F(e) = 1$ and $F(a) \neq 0$ for $a \in \mathcal{A}^{-1}$, then F is multiplicative. Equivalently: a **linear functional** F on a commutative complex unital Banach algebra $\mathcal{A}$ is multiplicative if and only if $F(a) \in \sigma(a)$ for all $a \in \mathcal{A}$, where $\sigma(a)$ stands for the spectrum of a (cf. also **Spectrum of an element**). As there is a one-to-one correspondence between linear multiplicative functionals and maximal ideals, the theorem can also be phrased in the following way: A codimension-one subspace M of a commutative complex unital Banach algebra $\mathcal{A}$ is an ideal if and only if each element of M is contained in a non-trivial ideal. The theorem is not valid for real Banach algebras.

The Gleason–Kahane–Żelazko theorem has been extended into several directions:

1) If φ is non-constant **entire function** and F is a linear functional on a complex unital Banach algebra $\mathcal{A}$, such that $F(e) = 1$ and $F(a) \neq 0$ for $a \in \varphi(\mathcal{A})$, then F is multiplicative [3].

2) Let M be a finite-codimensional subspace of the algebra $C(X)$ of all continuous complex-valued functions on a compact space X. If each element of M is equal to zero at some point of X, then the functions from M have a common zero in X [2]. It is not known if the analogous result is valid for all commutative unital Banach algebras.

3) The assumption of linearity of the functional F has been weakened, and the result has been extended to mappings between Banach and topological algebras.

See [4] for more information about the history, related problems, and further references.

References

[1] Gleason, A.M.: 'A characterization of maximal ideals', *J. d'Anal. Math.* **19** (1967), 171–172.

[2] Jarosz, K.: 'Finite codimensional ideals in function algebras', *Trans. Amer. Math. Soc.* **287**, no. 2 (1985), 779–785.

[3] Jarosz, K.: 'Multiplicative functionals and entire functions II', *Studia Math.* **124**, no. 2 (1997), 193–198.

[4] JAROSZ, K.: 'When is a linear functional multiplicative?': *Function Spaces: Proc. 3rd Conf. Function Spaces*, Vol. 232 of *Contemp. Math.*, Amer. Math. Soc., 1999.

[5] KAHANE, J.-P., AND ŻELAZKO, W.: 'A characterization of maximal ideals in commutative Banach algebras', *Studia Math.* **29** (1968), 339–343.

[6] ŻELAZKO, W.: 'A characterization of multiplicative linear functionals in complex Banach algebras', *Studia Math.* **30** (1968), 83–85.

K. Jarosz

MSC 1991: 46Hxx

GMANOVA, *generalized multivariate analysis of variance* – See **ANOVA**.

MSC 1991: 62Jxx