E0 219 Linear Algebra and Applications / August-December 2016

(ME, MSc. Ph. D. Programmes)

Download from : http://www.math.iisc.ernet.in/patil/courses/courses/Current Courses/...

Tel : +91-(0)80-2293 2239/(Maths Dept. 3212) **E-mails :** dppatil@csa.iisc.ernet.in / patil@math.iisc.ernet.in

Lectures : Monday and Wednesday ; 11:00–12:30 **Venue:** CSA, Lecture Hall (Room No. 117)

Corrections by : **Nikhil Gupta** (nikhil.gupta@csa.iisc.ernet.in ; **Lab No.: 303**) / **Vineet Nair** (vineetn90@gmail.com ; **Lab No.: 303**) / **Rahul Gupta** (rahul.gupta@csa.iisc.ernet.in ; **Lab No.: 224**) / **Sayantan Mukherjee** (meghanamande@gmail.com ; **Lab No.: 253**) / **Palash Dey** (palash@csa.iisc.ernet.in ; **Lab No.: 301 , 333 , 335**)

Midterms : 1-st Midterm : Saturday, September 17, 2016; 15:00 – 17:00 **2-nd Midterm :** Sunday, October 23, 2016; 15:00 – 17:00

Final Examination : Thursday, December 08, 2016, 09:00 – -12:00

Evaluation Weightage : Assignments : 20% **Midterms (Two) :** 30% **Final Examination :** 50%

Range of Marks for Grades (Total 100 Marks)						
	Grade S	**Grade A**	**Grade B**	**Grade C**	**Grade D**	**Grade F**
Marks-Range	> 90	76 — 90	61 — 75	46 — 60	35 — 45	< 35

	Grade A$^+$	**Grade A**	**Grade B$^+$**	**Grade B**	**Grade C**	**Grade D**	**Grade F**
Marks-Range	> 90	81 — 90	71 — 80	61 — 70	51 — 60	40 — 50	< 40

11. Eigenvalues[1], Characteristic Polynomials and Minimal Polynomials

Submit a solution of the $*$-**Exercise** ONLY. **Due Date :** Monday, 24-10-2016 (Before the Class)

• Highly recommended to solve the Exercise 11.5 to win 10 BONUS POINTS!!!
•• Complete Correct solution of the Exercise 11.8 carry 15 BONUS POINTS!!!
••• Complete Correct solution of the Exercise 11.10 carry 20 BONUS POINTS!!!

Let K be arbitrary field and let $\mathbb{K}$ denote either the field $\mathbb{R}$ or the field $\mathbb{C}$.

11.1 Let $V := \mathbb{K}^{\mathbb{R}}$ and let $T \in \mathbb{R}$ be a positive real number. Let $s_T : V \to V$ be the linear operator defined by $s_T(x)(t) := x(t+T)$ for $x \in V$.

---

[1] Eigenvalues and eigenvectors are introduced in of linear algebra or matrix theory. They are used in the investigation of linear transformations. The prefix eigen- is adopted from the German word eigen for “proper”, or “characteristic”. Historically, they arose in the study of quadratic forms and differential equations. Originally utilized to study principal axes of the rotational motion of rigid bodies, eigenvalues and eigenvectors have a wide range of applications.

In the 18-th century Euler, L. (1707 – 1783) studied the rotational motion of a rigid body and discovered the importance of the principal axes. Lagrange, J. L. (1736 – 1813) realized that the principal axes are the eigenvectors of the inertia matrix. In the early 19-th century, Cauchy, A. L. (1789 – 1857) saw how their work could be used to classify the quadric surfaces, and generalized it to arbitrary dimensions. Cauchy also coined the term “racine caractéristique” (characteristic root) for what is now called *eigenvalue* .

Fourier, J.-B. J. (1768 – 1830) used the work of Laplace and Lagrange to solve the heat equation by separation of variables in his famous 1822 book *Théorie analytique de la chaleur*. Sturm, J. K. F. (1803 – 1855) developed Fourier’s ideas further and brought them to the attention of Cauchy, who combined them with his own ideas and arrived at the fact that real symmetric matrices have real eigenvalues. This was extended by Hermite, C. (1822 – 1901) in 1855 to what are now called *Hermitian matrices*. Around the same time, Brioschi proved that the eigenvalues of orthogonal matrices lie on the unit circle and Clebsch, A. (1833 – 1872) found the corresponding result for skew-symmetric matrices. Finally, Weierstrass, K. (1815 – 1897) clarified an important aspect in the stability theory started by Laplace, P. S. (1749 – 1827) by realizing that defective matrices can cause instability.

In the meantime, Liouville, J. (1809 – 1882) studied eigenvalue problems similar to those of Sturm ; the discipline that grew out of their work is now called *Sturm-Liouville theory*. Schwarz, H. A. (1843 – 1921) studied the first eigenvalue of *Laplace’s equation* on general domains towards the end of the 19-th century.

In the beginning of the 20-th century, Hilbert, D. (1862 – 1943) studied the eigenvalues of integral operators by viewing the operators as infinite matrices. He was the first to use the German word eigen, which means "own", to denote eigenvalues and eigenvectors in 1904, though he may have been following a related usage by Helmholtz, H. (1821 – 1894). For some time, the standard term in English was "proper value", but the more distinctive term "eigenvalue" is standard today.

The first numerical algorithm for computing eigenvalues and eigenvectors appeared in 1929, when Von Mises (1893 – 1973) published the power method. One of the most popular methods today, the QR algorithm, was proposed independently by John G. F. Francis (1934 –) and Vera Kublanovskaya (1920 2012) in 1961.

(a) Show that 0 is neither a spectral value nor an eigenvalue for s_T and the eigenspace of s_T at 1 is $V_{s_T}(1) = V_{\text{per},T} := \{x \in V \mid x \text{ is periodic with period } T\}$.

(b) For $\mathbb{K} = \mathbb{C}$, show that every $\lambda \in \mathbb{C}^\times$ is an eigenvalue of s_T with eigenfunction $\exp(\ln(\lambda)/Tt)$, where, if λ is a negative real number then we put $\ln(\lambda) := \ln(|\lambda|) + i\pi$ and the eigenspace of s_T at λ is $\exp(\ln(\lambda)/Tt)V_{\text{per},T}$.

(c) For $\mathbb{K} = \mathbb{R}$, show that every positive real number λ is an eigenvalue of s_T and the eigenspace of s_T at λ is $\lambda^{t/T}V_{\text{per},T}$.

(d) For $\mathbb{K} = \mathbb{R}$, the eigenspace of s_T at the eigen-value -1 is called the *half periodic functions* and is usually denoted by $V_{\text{hper},T}$. Show that

(i) Every half periodic function is period with period $2T$.

(ii) $V_{\text{hper},T} = \cos(\pi t/T)V_{\text{per},T} + \sin(\pi tT/)V_{\text{per},T}$.

(iii) For a positive real number λ, the eigenspace of s_T at $-\lambda$ is $V_{f_T}(-\lambda) = \lambda^{t/T}V_{\text{hper},T}$.

(e) Eigenfunction corresponding to an eigenvalue $\lambda \neq 1$ are called *periodic functions of second kind with multiplicator* λ. Show that if λ is a n-th root of unity then every eigen-function of second kind with multiplicator λ is periodic with period nT. (**Remark :** The same assertions (a) to (e) hold for the restriction of s_T to the subspaces $C^k_{\mathbb{K}}(\mathbb{R})$, $k \in \mathbb{N} \cup \{\infty, \omega\}$.)

*__11.2__ Let $\mathfrak{A} \in M_n(K)$, $n \geq 2$ be a nilpotent matrix.

(a) If $\mathfrak{A}^{n-1} \neq 0$, then there does not exists any matrix $\mathfrak{B} \in M_n(K)$ with $\mathfrak{B}^2 = \mathfrak{A}$.

(b) The following statements are equivalent: (i) $\mu_{\mathfrak{A}} = \chi_{\mathfrak{A}} \,(= X^n)$. (ii) $\mathfrak{A}^{n-1} \neq 0$. (iii) Rank $\mathfrak{A} = n-1$. (iv) There exists a $x \in K^n$ such that $\mathfrak{A}^i x$, $i = 0, \ldots, n-1$ is a basis of K^n. (**Hint :** Since $\mathfrak{A}$ is nilpotent, the characteristic polynomial $\chi_{\mathfrak{A}} = X^n$ and the minimal polynomial of $\mu_{\mathfrak{A}} = X^m$ with $m \leq n$. Prove the implications (i) $\Longleftrightarrow$ (ii) $\Longleftrightarrow$ (iv) and (ii) $\Longleftrightarrow$ (iii). The matrix $\mathfrak{A}$ defines K-linear map $f := f_{\mathfrak{A}} : K^n \to K^n$, $f(x) := \mathfrak{A}x$. Then $\text{Rank} f = \text{Rank}\,\mathfrak{A}$. Since $\mathfrak{A}$ (and hence f) id nilpotent, $\text{Rank} f \leq n-1$ and $\text{Dim}_K \text{Ker} f = n - \text{Rank} f \geq n-(n-1) =$ by Rank-Theorem. For (iii)$\Rightarrow$(ii) by induction on n. — **Remark :** This Exercise give the characterization of the cyclic nilpotent operators, where an operator (resp. a matrix) is called *cyclic* if it satisfies the condition (iv). In general, this is further equivalent to the condition that the characteristic and minimal polynomials are equal, see Exercise 11.8 (e) below.)

11.3 Let $f : V \to V$ be an operator on the K-vector space V. The following statements are equivalent: (i) f is a homothecy. (ii) Every subspace of V is f-invariant. (iii) Every non-zero vector in V is an eigen-vector of f.

11.4 Let $\mathfrak{A}$ and $\mathfrak{B}$ be two $n \times n$-matrices over the field K, assume that one of them is invertible. Then there exists at most n distinct elements $a \in K$ such that the matrix $a\mathfrak{A} + \mathfrak{B}$ is not invertible. (**Hint :** Suppose that $\mathfrak{A}$ is invertible, then $\text{Det}\,\mathfrak{A} \neq 0$. Now, since $\text{Det}(a\mathfrak{A}+\mathfrak{B}) = \text{Det}(a\mathfrak{E}_n + \mathfrak{B}\mathfrak{A}^{-1}) \cdot \text{Det}(\mathfrak{A}) = \chi_{-\mathfrak{B}\mathfrak{A}^{-1}}(a) \cdot \text{Det}(\mathfrak{A})$, only for at most n eigenvalues a of $-\mathfrak{B}\mathfrak{A}^{-1}$, $\text{Det}(a\mathfrak{A}+\mathfrak{B}) = 0$.
Now suppose that $\mathfrak{B}$ is invertible, then $a\mathfrak{A}+\mathfrak{B}$ is invertible for $a = 0$ and for $a \neq 0$, $a\mathfrak{A}+\mathfrak{B}$ is not invertible only for the n eigenvalues of $-\mathfrak{A}\mathfrak{B}^{-1}$, since $\text{Det}(a\mathfrak{A}+\mathfrak{B}) = \text{Det}(\mathfrak{A}\mathfrak{B}^{-1} + a^{-1}\mathfrak{E}_n) \cdot \text{Det}(\mathfrak{B}) = a \cdot \chi_{-\mathfrak{A}\mathfrak{B}^{-1}}(a^{-1}) \cdot \text{Det}(\mathfrak{B})$.)

**__11.5__ Let $n \in \mathbb{N}$ and let K be a field with $k \cdot 1_K \neq 0$ for all $k = 1, \ldots, n$.

(a) An operator f on the n-dimensional K-vector space V is nilpotent if and only if $\text{Tr} f = \text{Tr} f^2 = \cdots = \text{Tr} f^n = 0$. (**Hint** If f is nilpotent, then so are $f^2, f^3, \ldots, f^n$ and hence the characteristic polynomials $\chi_{f^i} = X^n$, in particular, $\text{Tr} f^i = 0$ for all $i = 1, \ldots, n$. Prove the converse by induction on n. Since $\text{Tr}(f^i) = 0$ for all $i = 1, \ldots, n$, by Cayley-Hamilton Theorem $0 = \chi_f(f) = f^n - (\text{Tr}(f))f^{n-1} + \cdots + (-1)^n \text{Det}\,\text{id}_V$ and hence applying the trace map, we get $0 = \text{Tr}(\chi_f(f)) = \text{Tr}(f^n) - (\text{Tr}(f))\text{Tr}(f^{n-1}) + \cdots + (-1)^n \text{Det}\,\text{Tr}(\text{id}_V) = (1)^n n \text{Det}(f)$. It follows that $\text{Det} f = 0$ and hence f is not injective and $\text{Dim}_K \overline{V} < n = \text{Dim}_K V$, where $\overline{V} := V/\text{Ker} f$. Now use Test-Exercise T10.24 and apply induction.)

(b) Suppose that $a_1, \ldots, a_n$ are elements in K with

$$a_1^1 + \cdots + a_n^1 = 0$$
$$\cdots\cdots\cdots\cdots$$
$$a_1^n + \cdots + a_n^n = 0.$$

Then $a_1 = \cdots = a_n = 0$. (**Hint** Let $f : K^n \to K^n$ be the linear map defined by the diagonal matrix $\operatorname{Diag}(a_1,\ldots,a_n)$ (with respect to the standard basis $e_1,\ldots,e_n$ of K^n). Then for every $k = 1,\ldots,n$, the matrix of f^k (with respect to the standard basis) is the diagonal matrix $\operatorname{Diag}(a_1^k,\ldots,a_n^k)$ and by hypothesis $\operatorname{Tr}(f) = \operatorname{Tr}(f^2) = \cdots = \operatorname{Tr}(f^n) = 0$. Now apply the part (a) above, to conclude that $\mathfrak{A}$ is nilpotent. — **Remark :** The parts (a) and (b) are equivalent: There exists (by *Kronecker's Theorem*[2]) a field extension $K \subseteq L$ such that the characteristic polynomial χ_f of f splits into liner factors $\chi_f = (X-a_1)\cdots(X-a_n)$ in $L[X]$. Then the trace $\operatorname{Tr}(f^k) = a_1^k + \cdots + a_n^k$, see Example 11.B.13.)

11.6 Find the characteristic polynomial of the following matrices :

(a) $$\mathfrak{A} := \begin{pmatrix} a_1 & 0 & \cdots & 0 & 0 & \cdots & 0 & b_1 \\ 0 & a_2 & \cdots & 0 & 0 & \cdots & b_2 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \iddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_n & b_n & \cdots & 0 & 0 \\ 0 & 0 & \cdots & b_n & a_n & \cdots & 0 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & b_2 & \cdots & 0 & 0 & \cdots & a_2 & 0 \\ b_1 & 0 & \cdots & 0 & 0 & \cdots & 0 & a_1 \end{pmatrix} \in \mathrm{M}_{2n}(K).$$

(**Ans :** $\chi_{\mathfrak{A}} = \prod_{k=1}^n (X-a_k-b_k)(X-a_k+b_k)$.) (**Hint :** See Supplement S10.64 (c).)

(b) $$\mathfrak{A} := \begin{pmatrix} a & b_2 & \cdots & b_n \\ c_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_n & 0 & \cdots & 0 \end{pmatrix} \in \mathrm{M}_n(K).$$ (**Ans :** $\chi_{\mathfrak{A}} = X^n - aX^{n-1} - \left(\sum_{k=2}^n b_k c_k\right) X^{n-2}\ \ n \geq 2.$)

(c) $$\mathfrak{F}_n := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \in \mathrm{M}_n(\mathbb{R}).$$

(**Ans :** $2^n U_n(X/2)$, where U_n is the n-th Tchebychev polynomial of second kind (see Supplement S10.61 (c).) In particular, $\lambda_k := 2\cos(k\pi/(n+1))$, $k = 1,\ldots,n$ are eigenvalue s of $\mathfrak{F}_n$. The vector with components $\sin(k\pi i/(n+1))$, $i = 1,\ldots,n$ is an eigenvector corresponding to λ_k.)

11.7 Let f and g be operators on the K-vector space V.

(a) If either fg or gf is algebraic, then both fg and gf are algebraic and the minimal polynomials of fg and gf are either equal or differ by the factor X. Moreover, if either f or g is invertible, then $\mu_{fg} = \mu_{gf}$. Give examples of operators f and g on K^2 such that $\mu_{fg} \neq \mu_{gf}$.

(b) Suppose that V is finite dimensional. Then $\chi_{fg} = \chi_{gf}$. (**Hint :** Use Exercise 8.4 (b) to assume that either f is invertible or f is a projection.)

****11.8** Let f be an operator on the K-vector space V and let $x \in V$. Show that :

(a) $V_x := \sum_{m \in \mathbb{N}} K f^m(x)$ is the smallest f-invariant subspace of V which contain x. (**Remark :** The subspace V_x is called the *f-cyclic subspace generated by x*.)

[2] **Kronecker's Theorem** *Let K be a field and let $P \in K[X]$ be a non-zero polynomial. Then there exists a field extension $K \subseteq L$ such that P factores into linear factors in $L[X]$. Moreover, one can also choose L such that L has finite dimension over K (as an K-algebra).*

(b) V_x is finite dimensional if and only if there exists a monic polynomial $P \in K[X]$ such that $P(f)(x) = 0$. Moreover, in this case, if P_x is the monic polynomial of the smallest degree with $P_x(f)(x) = 0$, then P_x is the minimal polynomial and the characteristic polynomial of $f|V_x$.

(**Remark :** This polynomial P_x is called the f-annihilator of x and denoted by $\text{Ann}_f(x)$. With this $\text{Deg Ann}_f(x) = \text{Dim}_K V_x$.)

(c) If V is finite dimensional and $x_1, \ldots, x_r$ is a generating system for V, then μ_f is equal to $\text{lcm}(P_{x_1}, \ldots, P_{x_r})$. (**Hint :** Follows from (b) and the following more general assertion : If $f : V \to V$ is a K-lienar operator and $V = V_1 + \cdots + V_r$ is a sum of f-invariant subspaces $V_1, \ldots, V_r$, then $\mu_f = \text{lcm}(\mu_{f|V_1}, \ldots, \mu_{f|V_r})$. See also Supplement S11.14.)

(d) Suppose that V is finite dimensional. Then the following statements are equivalent :
(i) $V_{x_0} = V$ for some $x_0 \in V$.
(ii) There exists a K-basis $\mathfrak{v} = \{v_1, \ldots, v_n\}$ of V such that the matrix of f with respect to the basis $\mathfrak{v}$ is of the form

$$\mathfrak{A}_P := \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in \text{M}_n(K)$$

(iii) $\chi_f = \mu_f$.

(**Remark :** If any one of the above equivalent statements hold, then the operator f is called a cyclic operator and the element x_0 is called a cyclic element for f. The matrix $\mathfrak{A}_P$ is called the companion matrix of the polynomial P. — A matrix $\mathfrak{A} \in \mathbb{M}_I(K)$ is called cyclic if the operator $f_{\mathfrak{A}} : K^I \to K^I$ defined by $\mathfrak{A}$ is cyclic. *A matrix $\mathfrak{A} \in \mathbb{M}_I(K)$ is cyclic if and only if $\mathfrak{A}$ is similar to the companion matrix of its characteristic polynomial $\chi_{\mathfrak{A}}$.*)

(e) If χ_f has only *simple prime factors*, then f is cyclic. (**Hint :** In this case $\chi_f = \mu_f$ by 11.A.14.)

11.9 Let V be a finite dimensional K-vector space of dimension n.

(a) Let f and g be invertible operators on V. Then all operators $\lambda f - \mu g$, $(\lambda, \mu) \in K^2 - \{(0,0)\}$ are invertible if and only if the characteristic polynomial $\chi_{f^{-1}g}$ of $f^{-1}g$ has no zeroes, i.e., $f^{-1}g$ has no eigenvalue.

(b) Let $\Phi \colon V \times V \to V$ be bilinear. If K is algebraically closed and $n \geq 2$, then Φ has a zero divisor, i.e., there exist $x, y \in V$ with $x \neq 0 \neq y$ and $\Phi(x,y) = 0$. If $K = \mathbb{R}$ and n is odd and ≥ 3, then Φ has a zero divisor. (**Hint :** For $x \in V$ consider the operators $f_x \colon y \mapsto \Phi(x,y)$ on V. — **Remark :** A well-known deep Theorem of Adams which states that : *if $K = \mathbb{R}$ and $n \neq 0, 1, 2, 4, 8$, then Φ has a zero divisor*, i.e., *there exist $x, y \in V \smallsetminus \{0\}$ with $\Phi(x,y) = 0$.*)

*****11.10** Let $\lambda \in \mathbb{K}$ be an eigenvalue of the matrix $\mathfrak{A} = (a_{ij}) \in \text{M}_n(\mathbb{K})$. Then show that $|\lambda - a_{ii}| \leq z_i := \sum_{j \neq i} |a_{ij}|$ for at least one $i \in \{1, \ldots, n\}$ and also $|\lambda - a_{jj}| \leq s_j := \sum_{i \neq j} |a_{ij}|$ for at least one $j \in \{1, \ldots, n\}$. In particular, the (eigen) spectrum $\text{Spec}\,\mathfrak{A}$ of $\mathfrak{A}$ is contained in $\left(\bigcup_{i=1}^n \overline{\text{B}}(a_{ii}; z_i)\right) \cap \left(\bigcup_{j=1}^n \overline{\text{B}}(a_{jj}; s_j)\right)$, where $\text{D}_i(\mathfrak{A}) := \overline{\text{B}}(a_{ii}, z_i)$ (resp. $\overline{\text{B}}(a_{jj}, s_j)$) are the closed discs centered at a_{ii} (resp. a_{jj}) and radius z_i, $i = 1, \ldots, n$ (resp. s_j, $j = 1, \ldots, n$), — called the Gershgorin discs. — For a diagonal matrix $\mathfrak{D}$, the union of the Gershgorin discs $\cup_{i=1}^n \overline{\text{B}}(a_{ii}, z_i)$ coincides with the spectrum $\text{Spec}\,\mathfrak{D}$, and conversely. (**Hint :** On the contrary, suppose that $|\lambda - a_{jj}| > s_j$ for all $j = 1, \ldots, n$. Then the matrix $\lambda \mathfrak{E}_n - \mathfrak{A}$ is invertible by Exercise 4.3, see also Exercise 10.7 (a) which contradicts the fact that λ is an eigenvalue of $\mathfrak{A}$. The first assertion proves the second one by applying the first to the transpose matrix ${}^t\mathfrak{A}$ (which has the same eigenvalues as $\mathfrak{A}$). — **Remark :** This assertion is also known as the Gershgorin circle theorem[3] which is useful in solving matrix equations of the form

[3] It was first published by the Belarusian mathematician Gershgorin, S. (1901 – 1933) in 1931, see [**Gerschgorin, S.** Über die Abgrenzung der Eigenwerte einer Matrix, *Izv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk*, **7** (1931), 749-754]. He studied at Petrograd Technological Institute from 1923, becoming Professor in 1930, and from 1930 he worked in the Leningrad Mechanical Engineering Institute on algebra, theory of functions of complex variables, numerical methods and differential equations.

$\mathfrak{A}\mathfrak{x} = \mathfrak{b}$ for $\mathfrak{x}$, where $\mathfrak{b}$ is a vector and $\mathfrak{A}$ is a matrix with a large *condition number*[4]. The Gershgorin circle theorem can be strengthened as follows : *If the union* $\mathrm{D}(\mathfrak{A}) := \mathrm{D}_{i_1} \cup \cdots \cup \mathrm{D}_{i_k}$ *of k Gershgorin-discs is disjoint from the union* $\mathrm{D}'(\mathfrak{A}) := \cup_{i \in \{1,\ldots,n\} \smallsetminus \{i_1,\ldots,i_k\}} \mathrm{D}_i$ *of the other* $n-k$ *Gershgorin-discs then* $\mathrm{D}(\mathfrak{A})$ *contains exactly* k *and* $\mathrm{D}'(\mathfrak{A})$ $n-k$ *eigenvalues of* $\mathfrak{A}$. — **Proof :** The assertion is obviously true for diagonal matrices. For a proof consider $\mathfrak{B}(t) := (1-t)\mathfrak{D} + t\mathfrak{A}$, $t \in [0,1]$, where $\mathfrak{D} := \mathrm{Diag}(a_{11},\ldots,a_{nn})$. Note that the hypothesis $\mathrm{D}(\mathfrak{A}) \cap \mathrm{D}'(\mathfrak{A}) = \emptyset$, yields $\mathrm{D}(\mathfrak{B}(t)) \cap \mathrm{D}'(\mathfrak{B}(t)) = \emptyset$ for all $t \geq 0$, since the centers of the Gershgorin discs of $\mathfrak{B}(t)$ are same as those of $\mathfrak{A}$ and the radii are t times those of $\mathfrak{A}$. Let $d(t) := d(\mathrm{D}(\mathfrak{B}(t)), \mathrm{D}'(\mathfrak{B}(t)))$ denote the distance between $\mathrm{D}(\mathfrak{B}(t))$ and $\mathrm{D}'(\mathfrak{B}(t))$. Then $d(0) = d(\mathfrak{D}) \geq d(t) \geq d(\mathfrak{A}) = d(1) > 0$ (since the discs are closed and the function $t \mapsto d(t)$ is decreasing). Since the eigenvalues of $\mathfrak{B}(t)$ are continuous functions of t (this is proved below), for any eigenvalue $\lambda(t)$ of $\mathfrak{B}(t)$ in $\mathrm{D}(\mathfrak{B}(t))$, its distance $\delta(t) := d(\lambda(t), \mathrm{D}'(t))$ is also continuous. Obviously $\delta(t) \geq d(t) \geq d(1) > 0$ for all $t \in [0,1]$ and in particular, $\delta(0) \geq d(1) > 0$. Note that since the assertion is obviously true for the diagonal matrices, there are exactly k eigenvalues $\lambda_1(0),\ldots,\lambda_k(0)$ of $\mathfrak{D}$ in $\mathrm{D}(\mathfrak{D})$. We shall use this and the continuity of the function $,\delta$ to show that the eigenvalues $\lambda_1(1),\ldots,\lambda_k(1)$ of $\mathfrak{A}$ are in $\mathrm{D}(\mathfrak{D})$. For this we fix $i \in \{1,\ldots,k\}$ and put $\lambda(t) := \lambda_i(t)$. Suppose on the contrary that $\lambda(1) \in \mathrm{D}'(\mathfrak{A}) = \mathrm{D}'(\mathfrak{B}(1))$. Then $\delta(1) = 0$, and hence $\delta(0) \geq d(0) > d(1) > 0 = \delta(1)$. Therefore by *Intermediate value Theorem* (see Footnote 4 in Exercise 10.7) there exists a $t_0 \in (0,1)$ such that $\delta(t_0) = d(1)$. But, then $\delta(t_0) = d(1) < d(t_0) \leq \delta(t_0)$, which is impossible. This proves the assertion.

Now we shall indicate the proof of the assertion : *The zeros of a monic complex polynomial are continuous functions of its coefficients*, which is used in the above proof. More precisely :

Lemma *Let* λ *be a zero of the polynomial* $X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in \mathbb{C}[X]$ *of multiplicity* m. *Further, let* $\varepsilon > 0$ *be given. Then there exists a* $\delta > 0$ *such that all polynomials* $X^n + b_{n-1}X^{n-1} + \cdots + b_0 \in \mathbb{C}[X]$ *with* $|b_i - a_i| \leq \delta$ *for* $i = 0,\ldots,n-1$ *have at least* m, *zeroes in the (open) disc* $\mathrm{B}(\lambda\,;\varepsilon)$, *every zero is counted with its multiplicity.*

Proof. We consider the continuous map $\Phi\colon \mathbb{C}^n \to \mathbb{C}^n$, which maps every n-tuple of complex numbers $(\lambda_1,\ldots,\lambda_n)$ to the n-tuple $(a_0,\ldots,a_{n-1})$ of the coefficients (other than the leading coefficient) of the polynomial $(X-\lambda_1)\cdots(X-\lambda_n)$. Then Φ is surjective by the *Fundamental Theorem of Algebra*[5], and the fibre of Φ passing through the n-tuple $(\lambda_1,\ldots,\lambda_n)$ is the set of all n-tuples $\sigma(\lambda_1,\ldots,\lambda_n) = (\lambda_{\sigma^{-1}1},\ldots,\lambda_{\sigma^{-1}n})$, $\sigma \in \mathfrak{S}_n$. Further, *if* $A \subseteq \mathbb{C}^n$ *is a closed subset, then its image* $\Phi(A)$ *is also closed subset.* For, if $\Phi(x_\nu)$, $\nu \in \mathbb{N}$, $x_\nu \in A$, is a convergent sequence in $\Phi(A)$, then $x_\nu \in A$, is a bounded sequence by the Exercise[6] and hence by the *Bolzano-Weierstrass Theorem*[7] x_ν, $\nu \in \mathbb{N}$, has a convergent subsequence. We may therefore assume that x_ν, $\nu \in \mathbb{N}$, is already convergent. Then, if $x := \lim x_\nu \in A$, then $\Phi(x) = \lim \Phi(x_\nu) \in \Phi(A)$. Therefore it follows that: *If* $U \subseteq \mathbb{C}^n$ *open, then its image* $\Phi(U)$ *is also open.* The complement of $\Phi(U)$ in $\mathbb{C}^n$ is $\Phi\big(\mathbb{C}^n - \bigcup_{\sigma \in \mathfrak{S}_n} \sigma(U)\big)$ and hence it is closed by the above proof.
Let $X^n + a_{n-1}X^{n-1} + \cdots + a_0 = (X-\lambda_1)\cdots(X-\lambda_n)$ and $\varepsilon > 0$ be given. Then $\Phi\big(\mathrm{B}(\lambda_1\,;\varepsilon) \times \cdots \times \mathrm{B}(\lambda_n\,;\varepsilon)\big)$ is an open neighbourhood of $(a_0,\ldots,a_{n-1})$, which contains a product $\overline{\mathrm{B}}(a_0\,;\delta) \times \cdots \times \overline{\mathrm{B}}(a_{n-1}\,;\delta)$ of discs with $\delta > 0$. This proves the assertion. •)

[4]The condition number of a square non-singular matrix $\mathfrak{A}$ is defined by $\mathrm{cond}\,\mathfrak{A} = ||\mathfrak{A}|| \cdot ||\mathfrak{A}^{-1}||$. By convention, $\mathrm{cond}\,\mathfrak{A} = \infty$ if $\mathfrak{A}$ is singular. It is therefore a measure of how close a matrix is to being singular. A matrix with large condition number is nearly singular, whereas a matrix with condition number close to 1 is far from being singular. It is obvious from the definition that a nonsingular matrix and its inverse have the same condition number.

[5]**Fundamental Theorem of Algebra** (d'Alembert–Gauss) *Every non-constant polynomial* $f \in \mathbb{C}[X]$ *has a zero in* $\mathbb{C}$. — d'Alembert, J. (1717–1783) was a a French mathematician who was a pioneer in the study of differential equations and their use of in physics. He studied the equilibrium and motion of fluids. — Gauss, C. F. (1777–1855) was a German mathematician who worked in a wide variety of fields in both mathematics and physics including number theory, analysis, differential geometry, geodesy, magnetism, astronomy and optics. His work has had an immense influence in many areas.

[6]**Exercise** Let $f = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n$ be a monic polynomial in $\mathbb{C}[X]$. Then for every zero α of f in $\mathbb{C}$ prove the estimates :

(a) $|\alpha| \leq \mathrm{Max}\,(1, |a_0| + \cdots + |a_{n-1}|)$.

(b) $|\alpha| \leq \mathrm{Max}\,(|a_0|, 1 + |a_1|, \ldots, 1 + |a_{n-1}|)$.

(c) (Cauchy's Estimates) $|\alpha| \leq 2R$ mit $R := \mathrm{Max}\,(|a_\nu|^{1/(n-\nu)}, \nu = 0,\ldots,n-1)$. (**Hint :** From $|\alpha| > 2R$ and $f(\alpha) = 0$, we get $|\alpha|^n = |a_0 + \cdots + a_{n-1}\alpha^{n-1}| \leq \sum_{\nu=0}^{n-1} R^{n-\nu}|\alpha|^\nu = R\,\big(|\alpha|^n - R^n\big)/\big(|\alpha| - R\big) < |\alpha|^n$, a contradiction.)

[7] **Theorem** (Bolzano-Weierstrass) *Every bounded sequence of real numbers has a limit point.*