The proof requires some preliminary results. First we need inequality due to Hadamard which goes as follows:

6.2 Hadamard's inequality. *For every matrix A with a non-zero determinant one has the inequality*

$$\bigl|\det(A)\bigr| \cdot \mathrm{Norm}(A^{-1}) \leq \frac{||A||^{n-1}}{(n-1)^{n-1)/2}}$$

Exercise. Prove this result. The hint is to use expansions of certain determinants while one considers $\det(A) \cdot \langle A^{-1}(x), y\rangle$ for all pairs of unit vectors x and y.

We are given the simple algebra A. For ech non-ero elemebt $x \in A$ we put

$$\rho(x) = \dim_{\mathbf{C}} Ax$$

where Ax is the left principal ideal generated by x. Choose some x_* such that

$$\rho(x_*) = \min_{x \in A} \rho(x)$$

if $a \in A$ gives $ax_* \neq 0$ the obvious inclusion $Aax_* \subset Ax_*$ gives

$$\rho(ax_*) \leq \ rho(x_*)$$

and the minimal, choice in 8i) entials that eqauality holds which implies that the complex vector apaces Ax and Aax are equal. This can be expressed by saying that the left ideal Ax is minimal, i.e. it does not contain a proper left ideal than the zero-ideal. Next, since the algebra A is since, the *two-sided ideal* generated by x_* is equal to A. This means that the identty 1 can be wirtten as

$$1 = a_1 x b_1 + \ldots + a_m x b_m$$

where m is some positive integer and $\{a_\nu\}$ and $\{b_\nu\}$ two m-tuples of A-elements. Here m can be chosen minimal Let us then consider one of the left prinicipal ideals

$$L_\nu = A a_\nu x_* b_\nu$$

Firsot, since Ax_* is minimal we see that

$$L_\nu = Ax_* b_\nu$$

and (xx) above gives $\rho(x_* b_\nu) \leq \rho(x_*$ so the minmal choice of $\rho(x_)*$ entials by 8xx) above that L_ν is a minimal left ideal. Put

$$e_\nu = a_\nu x_* b_\nu \quad : 1 \leq \nu \leq m$$

Suppose one has an equaution

$$y_1 e_1 + \cdots + y_m e_m = 0$$

where $\{y_\nu\}$ is some m-tuple in A. If $y_1 e_1 \neq 0$ the fact that L_1 is minimal entails that there exists an element u_1 such that

$$u_1 y_1 e_1 = e_1$$

This would give

$$1 = (u_1 y_2 + 1)e_2 + \ldots + (u_1 y_m + 1)e_m$$

which contradicts the minimality of m in (xx). Hence $y_1 e_= 0$ and in a similar way one proves that $y_\nu e_\nu = 0$ for every ν. So one has a direct sum decomposition

$$A = \oplus A e_\nu \quad \& \quad 1 = e_1 + \ldots + e_m$$

Exercise. Show that (x) implies that $\{e_\nu\}$ are idempotents and that

$$e_k \cdot e_\nu = 0 \quad : k \neq \nu$$

Moreover, from the above

$$\rho(e_\nu) = \rho(x_*) = m$$

hold for every ν and hence the diirect sum decompsotion 8xx) gives

$$\dim A = m \cdot \rho(x_*)$$

Let $n \geq 2$ and A is an $n \times n$-matrix whose elements are complex numbers. Let λ be a complex parameter and E_n the identity matrix of order n. The charactwraitic polynomial of AQ is defined by

$$P(\lambda) = \det(\lambda \cdot E_n - A) \tag{0.1}$$

The zeros of P is a finite set of comppex numbers denoted by $\sigma(A)$ and called the spectrum of A. Since matrices with non-zero determintns are inveritble, it follows that there exists inverse matrices

$$R_A(\lambda) = (\lambda \cdot E_n - A)^{-1} \quad : \quad \lambda \in \mathbf{C} \setminus \sigma(A) \tag{0.2}$$

Moreover, the construction of inverse matrices via Cramer's rule entails gives an n-tuple of matrices $\{Q_\nu\}$ such that

$$R_A(\lambda) = \frac{1}{P_A(\lambda)} \cdot \sum_{\nu=0}^{n-1} \lambda^\nu \cdot Q_\nu \tag{0.3}$$

Next, when $|\lambda|$ is large we can express $R(\lambda$ by the Neuann series:

$$R(\lambda) = (\lambda \cdot E_n - A)^{-1} \quad : \quad \lambda \in \mathbf{C} \setminus \sigma(A) \tag{0.4}$$

1. Exercise. We can construct line integrals over circles $|\lambda| = w$ where w is strictly larger than the absolute value of every root of $P_A(\lambda)$. Residue calcuus teaches that

$$\frac{1}{2\pi i} \int_{|\lambda|=w} \lambda^\nu = 0 \quad : \quad \nu \neq -1$$

Show that (0.4) gives

$$A^k = \frac{1}{2\pi i} \int_{|\lambda|=w} \lambda^k \cdot R_A(\lambda) \cdot d\lambda \quad : \quad k = 1, 2, \ldots \quad \& \quad E_n = \frac{1}{2\pi i} \int_{|\lambda|=w} R_A(\lambda) \cdot d\lambda$$

More generally, if $q(\lambda)$ is an arbitrary polynomial then

$$q(A) = \frac{1}{2\pi i} \int_{|\lambda|=w} q(\lambda) \cdot R_A(\lambda) \cdot d\lambda \tag{1.1}$$

2. Some idempotent matrices. Consider a zero α of $P_A(\lambda)$ whose multiplicity is some positive integer e_α. Put

$$P_\alpha(\lambda) = \prod_{\beta \neq \alpha} (\lambda - \beta)^{e_\beta}$$

where the profict is taken over the reamining zeros of $P(\lambda)$. Euclidian divsions give a unique polynomial $\rho_\alpha(\lambda)$ of degree $\leq e_\alpha$ such that

$$\rho_\alpha(\lambda) \cdot P_\alpha(\lambda) = 1 + (\lambda - \alpha)^{e_\alpha} \cdot S_\alpha(\lambda)$$

where S_α is another polynomial.

Exercise. Show that

$$\rho_\alpha(A) \cdot P_\alpha(A = \rho = \frac{1}{2\pi i} \int_{|\lambda-\alpha|=\delta} R(\lambda)\, d\lambda \tag{1}$$

hold for all $\delta > 0$ for which the distance from the root α to other points in $\sigma(A)$ are $> \delta$. Moroeover, put

$$E_\alpha = \rho = \frac{1}{2\pi i} \int_{|\lambda-\alpha|=\delta} R(\lambda)\, d\lambda \tag{1}$$

and show that this matrix is idempotent, i.e.

$$E_\alpha = E_\alpha^2$$

Finally, show that

$$(A - \alpha \cdot E_n)^k \cdot E_\alpha = \frac{1}{2\pi i} \int_{|\lambda - \alpha| = \delta} \lambda^k \cdot R(\lambda)\, d\lambda \tag{1}$$

hold for every $k \geq 1$.

The minimal polynomial p_* of A. To begin with (1.1) and (0.3) give

$$p(A) = \sum \frac{1}{2\pi i} \int_{|\lambda| = w} \lambda^\nu \, d\lambda \cdot Q_\nu$$

and the right hand side is zero because the complex ine integrals over λ-monomials are all zero. Hence we have

$$p(A) = 0$$

Next, in the rpniciapl ideal domain of pilynomials of one variable over the complex field, we find a unique monic plynomia p_* of smallest degree such that $p_*(A) = 0$. Now one has a favtorisation

$$p(\lambda) = \prod (\lambda - \alpha)^{e_\alpha}$$

with the prodict taken over all α in $\sigma(A)$. Since $p(A) =$ it follows that

$$p_*(\lambda) = \prod (\lambda - \alpha)^{j_\alpha}$$

where $\{j_\alpha\}$ are integers and $j_\alpha \leq e - \alpha$ for every α. It turnd out that these j-integers acan be found via the meromorphic matrix-valued function $R(\lambda)$.

Theorem. *For each $\alpha \in \sigma(A)$ the integer j_α is equal to the order ot the ple of $R(\lambda)$ at α.*

Exercise Prove Theorem xx.

Jordan's theorem.

We are given a finite dimensional complex vector space V and $A \colon V \to V$ is a linear operator which is nilptent, i.e. there exists an integer N such that A^N is the zero operator. Every non-zero vector $x \in V$ generates an A-invariant subspace denoted by $\mathcal{C}(x)$. It menas that $\mathcal{C}(x)$ is the vector space generated by $x, Ax, A2x, \ldots$. Let $j(x)$ be the smallest positive integer such that

$$A^{j(x)}(x) = 0$$

Exercise. Show that $\mathcal{C}(x)$ is has dimensionl $j(x)$ where the vectors $x, Ax, \ldots, A^{j(x)-1}$ is a basis.

Now we anounce Jordan's theorem.

Theorem. *The ector sopace V is a direct sum of cyclic subspaces, i.e. there exists a finite set of vectors $\{x_\nu\}$ siuch that*

$$V = \oplus\, \mathcal{C}(x_\nu)$$

Proof.) we shall use an induction over $\dim V$. To begin with we choose a vector $x_* \in V$ such that

$$j(x_*) = \max_x j(x)$$

Consider he vector space

$$W = \frac{W}{\mathcal{C}(x_*)}$$

Since $\mathcal{C}(x)$ is an A-invariant subspace of V, it follows that A induces a linear operator $\bar{A}$ on W whic hobvousoy also is nilpotent. By the iunduction over $\dim V$ we can assume that Jordan's theorem holds for W. Hence

$$W = \oplus\, \mathcal{C}(\bar{x}_\nu)$$

where $\{\bar{x}_\nu\}$ are images of V-vectors $\{x_\nu\}$. By the condtriution of the quitoent space W, the j.numers of the vectors $\bar{x}_\nu$ are given by integers $\bar{j}(\nu)$ whosen to be minimal so that

$$A^{\bar{j}(\nu)}(x_\nu) \in \mathcal{C}(x_*)$$

hold in V. So with ν kept fixed we can wirte

$$A^{\bar{j}(\nu)}(x_\nu) = c_1 x_* + \ldots + c_{j(x_*)-1} \cdot x_*^{j(x_*)-1}$$

Here

$$j(\nu) \leq j(x_\nu) \leq j(x_*)$$

So with $k = j(x_*) - j(\nu)$ we get

$$0 = A^k(A^{\bar{j}(\nu)}(x_\nu) = c_1 A^k(x_*) + \ldots + c_{j(x_*)-1} \cdot A^{k+j(x_*)-1}(x_*)$$

Since $\{A^\nu(x_*) : 0 \leq \nu \leq j(x_*)\}$ are inearlu indep+endent it follows that

$$c_1 = \ldots c_{j(\nu)-1} = 0 \implies$$

$$A^{\bar{j}(\nu)}(x_\nu) = A^{j(\nu)}(\xi_\nu) \quad : , \xi_\nu \in \mathcal{C}(x_*)$$

Replacin x_ν by $x_\nu - \xi_\nu$ does not change its image in W and here

$$j(\nu) = j(x_\nu - \xi_\nu)$$

So now one has a direct sum decompostion

$$W = \oplus \mathcal{C}(\bar{x}_\nu) \quad : \ j(\nu) = j(x_\nu)$$

At this tage the reade can check that V is the direft sium of the cyuc lic suspaces $\{\mathcal{C}(x_\nu)\}$ in (x) together wit $\mathcal{C}(x_*)$. This finihses the proof of Jordan's theorem.

Resturing to (0.3) we notice that since $P_A(\lambda)$ is a polynomial of degree n with highest coefficient equal to one, it follows that

$$Q_{n-1} = E_n$$

Next, with $k = 1$ in (0.6) one has

$$A = Q_{n-2} + \lim_{R\to\infty} \frac{1}{2\pi i} \cdot \int_{|\lambda|=R} \frac{\lambda^n}{P_A(\lambda)}\, d\lambda \tag{0.10}$$

Let us write

$$P_A(\lambda) = \lambda^n + c_{n-1} \cdot \lambda^{n-1} + \ldots + c_0$$

The reader can check that the last term in (0.10) is c_{n-1} and hence

$$Q_{n-2} = A - c_{n-1} \tag{0.11}$$

If one continues in this way it follows that each $j \geq 2$ gives

$$Q_{n-j} = q_j(A)$$

where $q_j(A)$ is a polynomial in A of degree $\leq j - 1$. When $j = n$ the reader can check that Cauchy'sresidue formula gives

$$Q_0 = \frac{1}{2\pi i} \cdot \int_{|\lambda|=w} \frac{P_A(\lambda) \cdot R_A(\lambda)}{\lambda}\, d\lambda = A^{n-1} + c_{n-1}A^{n-2} + \ldots + c_2 A + c_1 \cdot E_n$$

1. The case when $P_A(\lambda)$ has simple roots. Let $\alpha_1, \ldots, \alpha_n$ be the simle roots. To each $1 \leq k \leq n$ we put

$$\mathcal{C}_k(A) = \frac{1}{\prod_{\nu\neq k} (\alpha_k - \alpha_\nu)} \cdot \prod_{\nu\neq k} (A - \alpha_\nu E_n) \tag{1.1}$$

When λ is outside $\sigma(A)$ we get the matrix

$$\mathcal{C}(\lambda) = \sum_{k=1}^{k=n} \frac{1}{\lambda - \alpha_k} \cdot \mathcal{C}_k(A) \tag{1.2}$$

With these notations one has the equation below which is due to Cayley, Hamilton and Sylvester:

$$\mathcal{C}(\lambda) = R_A(\lambda) \tag{1.3}$$

Exercise. Prove (1.3) using residuye calculas and the previous equations.

2. The Cayley-Hamilton polynomial. It is by definition the unique monic polynomial $p_*(\lambda]$ in the polynomial ring $\mathbf{C}[\lambda]$ of smallest possible degree such that the associated matrix $p_*(A) = 0$. It is found as follows: Let $\alpha_1, \ldots, \alpha_k$ be the distinct roots of $P_A(\lambda)$ so that

$$P_A(\lambda) = \prod_{\nu=1}^{\nu=k} (\lambda - \alpha_\nu)^{e_\nu}$$

where $e_1 + \ldots + e_k = n$. From (0.3) and (0.7) it is clear that

$$P_A(A) = 0$$

Hence $p_*(\lambda)$ is a factor of the chacteristic polynomial $P_A(\lambda)$. If P_A has multiple zeros it can occur that the degree of $p_*(\lambda)$ is strictly smaller thann. To get the exact formula for $p_*\lambda$ one needs Jordan's theorem in § 3 where we also explain how to compiute the minimal polynomial p_* attached to our given matrix A.

2.1 Residue matrices. Let $\alpha_1, \ldots, \alpha_k$ be the distinct zeros of $P_A(\lambda)$. For a given root, say α_1 of multiplicity $p \geq 1$ we have a local Laurent series expansion

$$R_A(\alpha_1 + \zeta) = \frac{G_p}{\zeta^p} + \ldots + \frac{G_1}{\zeta} + B_0 + \zeta \cdot B_1 + \ldots \tag{i}$$

which converges in a disc $\{|\zeta| < \delta\}$. One refers to $G_1, \ldots, G_p$ as the residue matrices at α_1. Choose a polynomial $q(\lambda)$ in $\mathbf{C}[\lambda]$ which vanishes up to the multiplicity at all the the remaining roots $\alpha_2, \ldots, \alpha_k$ while it has a zero of order $p-1$ at α_1, i.e. locally

(i) $$q(\alpha_1 + \zeta) = \zeta^{p-1}(1 + q_1\zeta + \ldots)$$

2.2 Exercise. Use residue calculus and show that:

(*) $$q(A) = \frac{1}{2\pi} \int_{|\lambda - \alpha_1| = \epsilon} q(\lambda) \cdot R_A(\lambda) \cdot d\lambda = G_p$$

Hence the matrix G_p is a polynomial of A. In a similar way one proves that every G-matrix in the Laurent series (i) is a polynomial in A.

2.3 Some idempotent matrices. Consider a zero α_j and choose a polynomial Q_j such that $Q_j(\lambda) - 1$ has a zero of order $e(\alpha_j)$ at α_j while Q_j has a zero of order $e(\alpha_\nu)$ at the remaining roots. Set

(1) $$E_A(\alpha_j) = \frac{1}{2\pi i} \int_{|\lambda| = w} Q_j(\lambda) \cdot R_A(\lambda) \cdot d\lambda$$

where w is large as in 2.5. Since the polynomial $S = Q_j - Q_j^2$ vanishes up to the multiplicities at all the roots of $P_A(\lambda$ we have $S(A) = 0$ from (0.9) in which entails that

(2.3.1) $$E_A(\alpha_j) = E_A(\alpha_j) \cdot E_A(\alpha_j)$$

In other words, we have constructed an idempotent matrix.

2.4 The Cayley-Hamilton decomposition. Recall the equality

$$E_n = \frac{1}{2\pi i} \cdot \int_{|\lambda| = w} R_A(\lambda) \cdot d\lambda$$

where the radius w is so large that the disc D_w contains the zeros of $P_A(\lambda)$. The previous construction of the E-matrices at the roots of $P_A(\lambda)$ entail that

$$E_n = E_A(\alpha_1) + \ldots + E_A(\alpha_k)$$

Identifying A with a $\mathbf{C}$-linear operator on $\mathbf{C}^n$ we obtain a direct sum decomposition

(*) $$\mathbf{C}^n = V_1 \oplus \ldots \oplus V_k$$

where each V_ν is an A-invariant subspace given by the image of $E_A(\alpha_\nu)$. Here $A - \alpha_\nu$ restricts to a *nilpotent* linear operator on V_ν and the dimension of this vector space is equal to the multiplicity of the root α_ν of the characteristic polynomial. One refers to (*) as the *Cayley-Hamilton decomposition* of $\mathbf{C}^n$.

2.5 About invertible matrices. Consider the characteristic polynomial $P_A(\lambda)$ and let us write

$$P_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0$$

Notice that $c_0 = (-1)^n \cdot \det(A)$. So if the determinant of A is $\neq 0$ the vanhsi ng of $P_A(A)$ gives the equation

$$A \cdot [A^{n-1} + c_{n-1}A^{n-2} + \ldots + c_1] = (-1)^{n-1}\det(A) \cdot E_n$$

Hence the inverse A^{-1} is a polynomial in A.

2.6 Similarity of matrices. Recall that the determinant of a matrix A does not change when it is replaced by SAS^{-1} where S is an arbitrary invertible matrix. This implies that the coefficients of the characterstic polynomial $P_A(\lambda)$ are intrinsically defined via the associated linear operator, i.e. if another basis is chosen in $\mathbf{C}^n$ the given A-linear operator is expressed by a matrix SAS^{-1} where S effects the change of the basis. Let us now draw an interesting consequence of the previous operational calculus. Let us give the following:

2.6.1 Definition. *A pair of $n \times n$-matrices A, B are* similar *if there exists some invertible matrix S such that*

$$B = SAS^{-1}$$

Since the product of two invertible matrices is invertible this yield an equivalence relation on $M_n(\mathbf{C})$ and $P_A(\lambda)$ depends only on its equivalence class. The question arises if to matrices A and B whose characteristic polynomials are equal are similar in the sense of Definition 2.6. This is not true in general. More precisely, *Jordan normal form* determines the eventual similarity between a pair of matrices with the same characteristic polynomial.

I:C Complex vector spaces

Contents

Introduction.

The modern era about matrices and determinants started around 1850 with major contributions by Hamilton, Sylvester and Cayley. In § 1 and § 2 we expose results which foremost are due to these mathematicians. Matrices and their determinants of arbitary high order were defined around 1810. One result from this early period is Cauchy's spectral theorem for a symmetric $n\times$n-matrix A whose characteristic polynomials have simple zeros. More precisely, when the elements of a are real, Cacuhy proved that there exists an orthogonal matrix U such that U^*AU is a a diagonal matrix. His result was later extended by Weierstrass to the general case when multiple zeros appear. Considerable credit to the whoke subjhect dealing with linear systems of equations must also be given to Cramer. Already in 17xx he gave the general inversion fornula for 4×4-matrices. More refined results were later established by Laplace around 1820 , such as his general expansion theorem to calculate determinants. Another major achievement is due to Camille Jordan. His theorem from 1850 is exposed in § 3.

Some facts will be taken for granted. The reader is expected to be familiar with the construction of determinants of $n\times n$-matrices $A=\{a_{pq}\}$. But let us recall Cramér's formula to solve systems of linear equations

$$\sum_{p=n}^{p=n} a_{pq}\cdot x_q=y_p \quad : \quad 1\leq p\leq n$$

Under the hypothesis that $\det(A)\neq 0$ this system has a unique solution $x_\bullet$ for every complex n-vector $y_\bullet$. It is obtained as follows: For each pair $1\leq p,q\leq n$ one deletes the p:th row and the q:th colum from A which gives an $(n-1)\times(n-1)$-matrix denoted by $A[p,q]$. Put

$$C_{p,q}=(-1)^{j+q}\cdot\det(A[p,q])$$

Computing the determinant of A via an expansion along the p:th row gives

$$\text{(i)}\qquad \sum_{p=1}^{p=n} a_{pq}\cdot C_{p,q}=\det(A)$$

At the same time one has

$$\text{(ii)}\qquad \sum_{p=1}^{p=n} a_{pq}\cdot C_{j,q}=0 \quad : j\neq p$$

which follows because a matrix with two equal rows has a zero determinant. Given an n-vector y we set

$$x_q=\frac{1}{\det(A)}\cdot\sum_{j=1}^{p=n} C_{j,q}\cdot y_j$$

Now (i-ii) entail that n-vector $x_\bullet$ solves (*).

Example. Take $n=2$ and let

$$A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\implies \det(A)=a_{11}a_{22}-a_{12}a_{21}$$

Consider the 2×2-matrix

$$B=\begin{pmatrix}C_{11} & C_{21}\\ C_{12} & C_{22}\end{pmatrix}=\begin{pmatrix}a_{22} & -a_{12}\\ -a_{21} & a_{11}\end{pmatrix}$$

The rule for products of matrices show that $B\cdot A=E_2$, i.e. B is the inverse of A. If $n\geq 3$, Cramer's invesrion formula for an $n\times n$-matrix A with a non-zero determinant is given by

$$(*)\qquad A^{-1}=\frac{1}{\det A}\cdot\begin{pmatrix}C_{11} & C_{21} & \dots & C_{n1}\\ C_{12} & C_{22} & \dots & C_{n2}\\ \dots & \dots & \dots & \dots\\ C_{1n} & C_{2n} & \dots & C_{nn}\end{pmatrix}$$

Another fundamental fact is the product formula for determinants:

$$\det(A) \cdot \det(B) = \det(AB)$$

which hold for every pair of $n \times n$-matrices. The verification is exposed in many text-books and therefore left to the reader. A consequence is that a determinant can be associated in an intrinsic way to a linear operator L on an n-dimensional complex vector space V where one from the start has not fixed a basis. Given a basis $e_1, \ldots, e_n$ one gets a matrix $A = \{a_p q\}$ where

$$L(e_p) = \sum_{q=1}^{q=n} a_{qp} e_q$$

i.e. the columns of the A-matrix give the vectors $\{L(e_p)\}$. If $f_1, \ldots, f_n$ is another basis then L is represented by another matrix B and one has the equality

$$B = SAS^{-1}$$

where S is an invertible matrix which interchanges the e-basis with the f-basis. The product rule gives $\det(A) = \det(B)$ and this common determinant is therefore associated to the linear operator L. A perspective on the construction of determinants arises via exterior products of a given vector space V, i.e. for every $1 \leq p \leq n$ one considers elements of the form

$$v_1 \wedge \ldots \wedge v_p$$

with the rule that

$$v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(p)} = \text{sign}(\sigma) \cdot v_1 \wedge \ldots \wedge v_p$$

where $\text{sign}(\sigma)$ is the signature of the permutation σ acting on $1, \ldots, p$. Now one get a vector space denoted by $\wedge^p(V)$ whose dimension os $\binom{n}{p}$. It is called the p-fold exterior product of V. A linear operator L yields a linear operator $\wedge^p(L)$ on this vector space. When $p = n$ one has a 1-dimensional vector space and $\det(L)$ corresponds to a scalar multiplication on $\wedge^n(V)$ by this complex number.

Since many text-books expose linear algebra we shall not dwell upon general constructions, but pay attention to results ehichminvolve *inequalities* since this is the central issue in mathematics, at least if one is concerned with analysis. To attain this one often needs quite involved results about determinants. An example is Hadamard's theorem in the chapter devoted to series which gives conditions on order that a power series with a finite positive radius of convergence extends to a meromorphic function in a larger disc. Lack of space prevents us from a more detailed account about the positions of zeros of analytic functions vis their Taylor series. Let us only remark that if $P(z) = 1 + c - 1z + \cdots + c_n z^n$ is a polynomial whose zeros are simple, then the sequence of their absolute values $\{|\alpha_\nu\}$ can be recaptured via determinants. This goes back to work by Runge in his article [Acta Mathemaca vol 6] and is also treated in Hadamard's cited thesis. One should add that Bernoulli found the smallest absolute value of the roots of $P(z)$ by investigating its logarithmic derivative. See § xxx. It is therefore no surprise that analytic function theory has become a very useful tool to study matrices and their determinants, and conversely calculations with determinants lead to some quite remarkable results in analytic function theory. Recall also that the postion of roots to characteristic polynomials are used to study asymptotic properties of solutions to ordinary differential equations. An example where analytic function theory was used to achieve conditions in order tht the roots of a charaterstic polynomia,stay in a half-space $\mathfrak{Re}\,\lambda < 0$ was given by Eduard Routh in his famous treatise from 1876. Since his proof relies on analytic function theory it is given in my notes on this topic.

An open problem. While we review classic results in this chapter it may be of interest to give an example of an open problem in the spirit of the subsequent material. Let $m \geq 2$ and consider a differential operator

$$T = q_m(x)\partial^m + \ldots + x\partial$$

where $\partial = \frac{d}{dx}$ and $\{q_\nu\}$ are polynomials such that

$$\deg q_\nu \leq \nu - 1 \quad : 2 \leq \nu \leq m$$

It is easily seen that for evety positive integer n there exists a unique monic plynomial $p_n(x)$ of degree n such that

$$T(p_n) = n \cdot p_n$$

Denote by r_n^* the maximum of the absolute values of the zeros of p_n. A result due to T. Bergquist in the phd-thesis [Berg] shows that there exists a constant $c_* > 0$ such that

$$r_n^* \geq c_*$$

hold for every $n \geq 1$. But the following is an open:

Conjercture. Does there exists a constant c^* which depends on ∞t nnoly such that

$$r_n^* \leq c^* \cdot n \tag{*}$$

hold for every $n \geq 1$. Let us remark that numerical experiments indicate that c^* exists. But so far no proof has been found, except for special cases. For example, in situations where the zeros of the eigenpolynomials above are real then it is easily verified via Sturm chains that c^* exists. It is tempting to try to apply the studiues by Hadamard and Ringe to settle the conjecture. But I have bern unable to do this and perhaps some rader is able to find the answer. For a proof of the lower boud c_* I refer to a joint preprint by T. Bergquist and myself.

Integral equations. The calculus with determinants is used to study integral equations. The *Gram-Fredholm formula* given in § xx is the starting point to analyze spectra of linear operators expresse via kernel functions. Results about matrices of finite order and their determinants have paved the way to operator theory on normed linear spaces which in general are infinite dimensional.

An example. As alreafy said the aim of this chapter is to expose precise inequalities.

An important result is the *spectral theorem* for symmetric and real $n \times n$-matrices, and its counterpart for complex Hermitian matrices. In both these cases eigenvalues are found by regarding maxima and minima of quadratic forms. Far-reaching studies of quadratic forms appear in the colected work by Weierstrass which contains a wealth of results related to the spectral theorem for hermitian matrices and their interplay with quadratic forms. One should also mention later investigations by Frobenius about quadratic forms. Here is an example from Weierstrass' studies which goes as follows: Let $N \geq 2$ and $\{c_p q : 1 \leq p, q \leq N\}$ a doubly indexed sequence of positive numbers which is symmetric, i.e. $c_{qp} = c_{pq}$ hold for all pairs $1 \leq p, q \leq N$. Suppose that

$$\sum_{q=1}^{q=N} c_{p,q} \leq 1 \quad : 1 \leq p \leq N$$

Then

$$\sum_{p=1}^{p=N} \Big[\sum_{q=1}^{q=N} c_{p,q} \cdot x_q \Big]^2 \leq \sum_{p=1}^{p=N} x_p^2 \tag{0.1}$$

hold for every N-tuple $\{x_p\}$ of non-negative real numbers. The proof uses the spectral theorem for symmetric matrices and is given in § xx. A result with a wide range of applications due to frobebnius goes as follows: Let $\{a_{pq} : 1 \leq p, q \leq N\}$ be a double indexed family of positive real numbers where symmetry is not assumed. This double indexed family are elements of an $N \times N$-matrix A which yields a linear operator on $\mathbf{R}^N$. We shall learn how to contruct determinants and the zeros of the polynomial $P_A(\lambda) = \det(\lambda \cdot E_N - A)$ are in general complex numbers. When the elements of A are positive real numbers Frobenius proved that there exists a unique N-vector $x^* = (x_1^*, \ldots, x_N^*)$ where every $x_\nu^* > 0$ and $\sum x_\nu^* = 1$ which is an eigenvector for A, i.e.

$$A(x^*) = \rho \cdot x^*$$

holds for a positive real number ρ. Moreover, ρ is a simple zero of $P_A(\lambda)$ and the absolute value of every other root is $< \rho$. A result where the calcukus based upon determinants and solutions to systems of linear equations using the rule of Cramer becomes useful appears in Hadamard's theorem exposed in §xx which give necessary and sufficiet conditions for in order that a complex powere series $\sum c_n \cdot z^n$ which from the start have some finite radius of convergence. $\rho > 0$ extends to a meromrphic function in a large disc $|z| < \rho^*$. Among other important results one should mention the theorem due to Camille Jordan which shows that a linear operator after a suitable linear transformation is represented by a matirx of special form.

Using Lagrange's interpolation formula Sylvester exhibited extensive classes of matrix-valued functions by residue calculus and more delicate results were achieved by Frobenius who treated the general case when a characteristic polynomial of a matrix has multiple roots. Thjs is exposed in § 0.4. Passing to inifinite dimensions, the usefulness of matrices and their determinants was put forward by Fredholm in his studies of integral equations. Here estimates are needed to control determinants of matrices of large size to study resolvents of linear operators acting on infinite dimensional vector spaces. To handle cases where singular kernels appear in an integral operator, modified Fredholm determinants were introduced by Hilbert whose text-book *Zur Theorie der Integralgleichungen* from 1904 laid the foundations for spectral theory of linear operators on infinite dimensional spaces. A systematic study of matrices with infinitely many elements was done by Hellinger and Toeplitz in their joint article *Grundlagen für eine theorie der undendlichen matrizen* from 1910 and applied to solve integral equations of the Fredholm-Hilbert type. During these investigations Carleman's inequality for norms of resolvents in § 6 is a veritable cornerstone. Let me remark that Carelan's proof ffers a very insructive lesson in the subject dealing with matrices and their determinants.

Outline of the content.

Here follows a brief presentation of basic material. To each integer $n \geq 2$ we denote by $M_n(\mathbf{C})$ the set of $n \times n$-matrices with complex elements. As a complex vector space $M_n(\mathbf{C})$ has dimension n^2 and it is an associative $\mathbf{C}$-algebra defined by the usual matrix product where the identity E_n is the matrix whose elements outside the diagonal are zero while $e_{\nu\nu} = 1$ for every $1 \leq \nu \leq n$. When $n \geq 2$ a pair of $n \times n$-matrices A and B do not commute in general which means that $M_n(\mathbf{C})$ is a non-commutative algebra over the complex field. In § 1 we prove Wedderburn's theorem which asserts that the matrix algebras $\{M_n(\mathbf{C})\}$ are the sole finite dimensional complex algebra with no other two-sided ideals than the zero ideal and the whole algebra. Resovents are studied in § 2. They consist of inverse matrices $R_\lambda(A) = \lambda \cdot E_n - A)^{-1}$ when λ is outside the spectrum $\sigma(A)$ of a matrix A defined as the set of zeros of the characteristic polynomial

$$P_A(\lambda) = \det(\lambda \cdot E_n - A) \tag{0.1}$$

A fundamental fact is that $P_A(\lambda)$ only depends upon the associated linear operator defined by the A-matrix. More precisely, if S is an invertible matrix the product formula for determinants give the equality

$$P_A(\lambda) = P_{SAS^{-1}}(\lambda) \tag{0.2}$$

Using some analytic function theory one gets a certai calculus with resolvents which was carried out by Cayley, Hamilton and Sylvester and was later extended by Carl Neumann to study inverses of linear operators on normed vector spaces which in general need not be bounded, but only densely defined. So inspired by the material in the present chapter which deals with finite-dimensional situations, the Neumann calculus which atarted in 1880 has become a corner stone in operator theory and exposed in my notes about functional analysis.

Final remark. The subsewuent material os foremost devoted to establish *inequalities*, i.e. reader who prefer formal and more abstract theories may refrain from studying our account about linear operators in finite dimensional vector spaces, while they are of great importance in applications to various problems in analysis. An example of the psirit in the sections below is the followinfg inequality for iterated functions.

A. Matrices and determinants.

Let A be an $n \times n$-matrix whose elements $\{a_{pq}\}$ are complex numbers. The Hilbert-Schmidt norm is defined by

$$(*) \qquad \mathbf{HS}(A) = \sqrt{\sum\sum |a_{pq}|^2}$$

where the doube sum extends over all pairs $1 \leq p, q \leq n$. The operator norm is defined by:

$$(**) \qquad ||A|| = \max_{z_1,\ldots z_n} \sqrt{\sum_{p=1}^{p=n} |\sum_{q=1}^{q=n} a_{pq} z_q|^2}$$

with the maximum taken over n-tuples of complex numbers such that $\sum |z_p|^2 = 1$. Introduce the Hermitian inner product on $\mathbf{C}^n$ and identify A with the linear operator which sends a basis vector e_q into

$$A(e_q) = \sum_{p=1}^{p=n} a_{pq} \cdot e_p$$

If z and w is a pair of complex n-vectors one gets:

$$\langle Az, w\rangle = \sum\sum a_{pq} z_q \bar{w}_p$$

The Cauchy-Schwarz inequality gives

$$(1) \qquad \big|\langle Az, w\rangle\big|^2 \leq \big(\sum_{p=1}^{p=n} |\sum_{q=1}^{q=n} a_{pq} z_q|^2\big) \cdot \sum_{p=1}^{p=n} |w_p|^2$$

So if both z and w have length ≤ 1 The definition of the operator norm entails that

$$(2) \qquad ||A|| = \max_{z,w} |\langle Az, w\rangle|$$

where the maximum is taken over vectors z and w of unit length. Next, another application of the Cauchy-Schwarz inequality shows that if z has unit length, then

$$(3) \qquad \sum_{p=1}^{p=n} |\sum_{q=1}^{q=n} a_{pq} z_q|^2 \leq \sum_{p=1}^{p=n}\sum_{q=1}^{q=n} |a_{pq}|^2$$

Then (1) and (3) give the inequality

$$(4) \qquad ||A|| \leq \mathbf{HS}(A)$$

5. Example. Consider a matrix A whose elements are non-negative real numbers. Then it is clear that the operator norm is found when we use complex n-vectors with real and non-negative components. Thus,

$$(5.1) \qquad ||A|| = \max_{x_1,\ldots x_n} \sqrt{\sum_{p=1}^{p=n} \big(\sum_{q=1}^{q=n} a_{pq} x_q\big)^2}$$

taken over real n-vectors for which $\sum x_p^2 = 1$ with every $x_p \geq 0$. The A-norm is found via Lagrange's multiplier which shows that (5.1) is maximized by a real non-negative n-vector x satisfying the linear system of equations

$$(\mathrm{i}) \qquad \lambda \cdot x_j^* = \sum_{p=1}^{p=n} a_{pj} \cdot \sum_{q=1}^{q=n} a_{pq} x_q^*$$

Introducing the double indexed numbers

$$(\mathrm{ii}) \qquad \beta_{jq} = \sum_{p=1}^{p=n} a_{pj} a_{pq}$$

Lagrange's equations corresponds to the system

(iii) $$\lambda \cdot x_j^* = \sum_{q=1}^{q=n} \beta_{jq} \cdot x_q^*$$

Notice that the β-mstrix is symmetric, i.e. $\beta_{jq} = \beta_{qj}$ hold for ech pair. So (iii) amounts to find an eigenvector to the symmetric β-matrix with an eigenvector x^* for which $x_j^* \geq 0$ hold for each j. In "generic" cases the $\{\beta_{jq}\}$ are strictly positive numbers, and for such special matrices methods to attain the largest eigenvalue have been studied extensively by Frobenius and Perron. As a specific example we consider an $n \times n$-matrix of the form

$$T_s[n] = \begin{pmatrix} 1 & s & s & \dots & s \\ 0 & 1 & \dots & s & s \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \\ 0 & 0 & \dots & 1 & s \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

In spite of the explicit expression above the computation of its operator norm is not so obvious. For example, with $s = 2$ a result due to Hankel and Frobenius gives

(*) $$||T_2[n]|| = \cot \frac{\pi}{4n}$$

Of course, classic formulas of this kind are nowadays implemented in computer programs where the reader can "check" (*). But of course it is more instructive to try to prove it. An excellent text which offers detailed proofs of many delicate results about matrices and deterjinants is Gerhard Kovalevski's text-book *Determinantenheorie* from 1909. Pereonally I have not fonund any more recent text-book whch contains such challenging results, and it goes without saying that theorems like thos in § A.6 below cannot be derived by trivial abstract non nsense åresented in elementary books in linear algebra.

A.6 Sylvester's determinant formula.

Let A be some $n \times n$-matrix with elements $\{a_{ik}\}$. For eacg integer $1 \leq h \leq n-1$ one constructs the $(n-h \times (n-h)$-matrix whose elements are

$$b_{rs} = \det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1h} & a_{1s} \\ a_{21} & a_{22} & \dots & a_{2h} & a_{2s} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{h1} & a_{h2} & \dots & a_{hh} & a_{hs} \\ a_{r1} & a_{r2} & \dots & a_{rh} & a_{rs} \end{pmatrix} \quad : \quad h+1 \leq r, s \leq n$$

With these notation one has the Sylvester equation:

(*) $$\det \begin{pmatrix} b_{h+1,h+1} & b_{h+1,h+2} & \dots & b_{h+1,n} \\ b_{h+2,h+1} & b_{h+2,h+2} & \dots & b_{h+2,n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ b_{n,h+1} & b_{n,h+2} & \dots & b_{n,n} \end{pmatrix} = \Big[\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1h} \\ a_{21} & a_{22} & \dots & a_{2h} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{h1} & a_{h2} & \dots & a_{hh} \end{pmatrix} \Big]^{n-h-1} \cdot \det(A)$$

For a proof of (*) we refer to original work by Sylvester or [Kovalevski: page xx-xx].

A result about symmetric matrices. The next result is also due to Sylvester. Let $n \geq 2$ and consider a symmetric matrix

$$S = \begin{pmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ s_{21} & s_{22} & \dots & s_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ s_{n1} & a_{N2} & \dots & s_{nn} \end{pmatrix}$$

Now we construct three matrices as follows. First we get three $(n-1) \times (n-1)$-matrices

$$S_1 = \begin{pmatrix} s_{22} & s_{23} & \dots & s_{2n} \\ s_{32} & s_{33} & \dots & s_{3n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ s_{n2} & s_{N3} & \dots & s_{nn} \end{pmatrix} \quad : \quad S_2 = \begin{pmatrix} s_{12} & s_{13} & \dots & s_{1n} \\ s_{22} & s_{23} & \dots & s_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ s_{n-1,2} & s_{N-1,3} & \dots & s_{n-1,n} \end{pmatrix}$$

$$S_3 = \begin{pmatrix} s_{11} & s_{12} & \dots & s_{1,n-1} \\ s_{21} & s_{22} & \dots & s_{2,n-1} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ s_{n-1,1} & s_{N-1,2} & \dots & s_{n-1,n-1} \end{pmatrix}$$

We have also the $(n-2) \times (n-2)$-matrix when extremal rows and columns are removed:

$$S_* = \begin{pmatrix} s_{22} & s_{23} & \dots & s_{2,n-1} \\ s_{32} & s_{33} & \dots & s_{3,n-1} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ s_{2,n-1} & s_{3,n-1} & \dots & s_{n-1,n-1} \end{pmatrix}$$

A.6.1 Sylvester's identity. *One has the equation:*

$$\det S \cdot \det S_* = \det S_1 \cdot \det S_3 - \big(\det S_2\big)^2$$

The proof is left as an exercise. If necessay, consult the literature where the most elegant proofs of course are found in Sylvester's original work.

A.7 The Franke-Sylvester theorem.

Let $n \geq 2$ and $A = \{a_{ik}\}$ an $n \times n$-matrix. Each integer $1 \leq m < n$ gives the family of minors of size m. Thus, one one picks m columns and m rows which give an $m \times m$-matrix whose determinant is a minor of size m. The total number of such minors is equal to

$$\binom{n}{m}^2$$

Namely, with $N = \binom{n}{m}$ one has N many strictly increasing sequences $1 \leq \gamma_1 < \dots \gamma_m \leq n$ where a γ-sequence corresponds to preserved columns, and similarly there exist N strictly increasing sequences which correspond to preserved rows. With this in mind we get for each pair $1 \leq r, s \leq N$ a minor $\mathfrak{M}_{rs}$ where the enumerated r:th γ-sequence preserve columns and s corresponds to the enumerated sequence of rows. Now we obtain the $N \times N$-matrix

$$\mathcal{A}_m = \begin{pmatrix} \mathfrak{M}_{11} & \mathfrak{M}12 & \ldots & \mathfrak{M}_{1N} \\ \mathfrak{M}_{21} & \mathfrak{M}22 & \ldots & \mathfrak{M}_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \mathfrak{M}_{N1} & \mathfrak{M}22 & \ldots & \mathfrak{M}_{NN} \end{pmatrix}$$

It is called the Franke-Sylvester matrix of order m.

A.7.1 Theorem. *For every $1 \leq m < n$ one has the equality*

$$\mathcal{A}_m = \det(A)^{\binom{n-1}{m-1}}$$

Example. Consider the diagonal 3×3-matrix:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

With $m = 2$ we have 9 minors of size 2 and the reader can recognize that when they are arranged so that we begin to remove the first column, respectively the first row, then the resulting $\mathfrak{M}$-matrix becomes

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Its determinant is $4 = 2^2$ which is in accordance with the general formula since $n = 3$ and $m = 2$ give $\binom{n-1}{m-1} = 2$. For a proof of Theorem A.7.1 the reader can consult [Kovalevski: page102-105].

A.8 Hankel determinants.

Let $\{c_0, c_1, \ldots\}$ be a sequence of complex numbers. For each integer $p \geq 0$ and every $n \geq 0$ we get the symmetric $(p+1) \times (p+1)$-matrix:

$$\mathcal{C}_n^{(p)} = \begin{pmatrix} c_n & c_{n+1} & \ldots & c_{n+p} \\ c_{n+1} & c_{n+2} & \ldots & c_{n+p} \\ \ldots & \ldots & \ldots & \ldots \\ c_{n+p} & c_{n+p+1} & \ldots & c_{n+2p} \end{pmatrix}$$

Set

(A.8.1) $$\mathcal{D}_n^{(p)} = \mathcal{C}_n^{(p)}$$

One refers to $\{\mathcal{D}_n^{(p)}\}$ as the recursive Hankel determinants. They are used to describe properties of the given c-sequence. To begin with we define the rank r^* of $\{c_n\}$ as follows: To every non-negative integer n one has the infinite vector

$$\xi_n = (c_n, c_{n+1}, \ldots)$$

We say that $\{c_n\}$ has finite rank if there exists an integer N such that N many ξ-vectors are linearly independent and the rest are linear combinations of these.

A.8.2 Taylor series of rational functions. A given sequence $\{c_n\}$ gives the formal power series

(*) $$f(x) = \sum_{\nu=0}^{\infty} c_\nu x^\nu$$

For every $n \geq 1$ we get a new formal power series:

(*) $$\phi_n(x) = x^{-n} \cdot (f(x) - \sum_{\nu=0}^{n-1} c_\nu x^\nu) = \sum_{\nu=0}^{\infty} c_{n+\nu} x^\nu$$

Regariding the right hand side it is clear that the sequence $\{c_\nu\}$ has finite rank if and only if $\{\phi_n(x)\}$ generates a finite dimensional complex subspace of the vector space $\mathbf{C}[[x]]$ whose elements are formal power series. If this dimension is finite we find a positive integer p and a non-zero $(p+1)$-tuple $(a_0,\ldots,a_p)$ of complex numbers such that the power series

$$a_0\cdot\phi_0(x)+\ldots+a_p\cdot\phi_p(x)=0$$

Multiplying this equation with x^p it follows from the first equaity in (**) that

$$(a_p+a_{p-1}x+\ldots+a_o x^p)\cdot f(x)=q(x)$$

where $q(x)$ is a polynomial of degree $p-1$ at most. So when $\{c_n\}$ has finite rank the power series (*) represents a rational function.

Exercises.

Assume that

$$\sum c_\nu x^\nu=\frac{q(x)}{g(x)}$$

for some pair of polynomials. Show that $\{c_n\}$ has finite rank. Show also that a sequence $\{c_n\}$ has finite rank if and only if there exists an integer p such that

$$(4)\qquad \mathcal{D}_0^{(p)}\neq 0\quad\text{and}\quad D_0^{(q)}=0\quad:\quad q>p$$

and check also that p is equal to the rank N of the given sequence.

A.8.3 specific example. Suppose that the degree of q is strictly less than that of g in Exercise B.1 and that the rational function $\frac{q}{g}$ is a sum of simple fractions, i.e.

$$\sum c_\nu x^\nu=\sum_{k=1}^{k=p}\frac{d_k}{1-\alpha_k x}$$

where $\alpha_1,\ldots,\alpha_p$ are distinct and every $d_k\neq 0$. The reader can check that $c_0=\sum d_k$ and $n\geq 1$ gives:

$$c_n=\sum_{k=1}^{k=p}d_k\cdot\alpha_k^n$$

A.8.4 The reduced rank. Assume that $\{c_n\}$ has a finite rank N. To each $k\geq 0$ we denote by r_k the dimension of the vector space generated by $\xi_k,\xi_{k+1},\ldots$. It is clear that $\{r_k\}$ decrease as k increases. Hence there exists a non-negative integer N_* such that $r_k=N_*$ for large k. One refers to N_* as the reduced rank. It is obvious that $N_*\leq N$. The relation between N_* and N is related to the representation

$$f(x)=\frac{q(x)}{g(x)}$$

where q and g are polynomials without common factor. We shall not pursue this study in detail but refer to the literature. See for example the exercises in [Polya-Szegö : Chapter VII:problems 17-34].

A.8.5 Hankel's formula for Laurent series. Let $p\geq 2$ and consider a rational function of the form

$$R(z)=\frac{z^{p-1}}{z^p-[a_1z^{p-1}+\ldots+a_{p-1}z+a_p]}$$

The rational function R has a simple pole at inifinity and when $|z|$ is large it is expressed via a Laurent series

$$R(z)=\frac{c_0}{z}+\frac{c_1}{z^2}+\ldots$$

Consider the $p \times p$-matrix

$$A = \begin{pmatrix} 0 & 0 & & \dots & 0 & a_p \\ 1 & 0 & 0 & \dots & 0 & a_{p-1} \\ 0 & 1 & 0 & \dots & \dots & c_{a-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & a_1 \end{pmatrix}$$

Exercfise. Prove Hankel's equations which assert that for every $n \geq 1$ one has:

$$\mathcal{D}_n^{(p)} = \mathcal{D}_0^{(p)} \cdot \big[\det(A)\big)^n$$

A.8.6 The Hadamard-Kronecker identity. For all pairs of positive integers p and n one has the equality:

(*) $$\mathcal{D}_n^{(p+1)} \cdot \mathcal{D}_{n+2}^{(p-2)} = \mathcal{D}_n^{(p+1)} \mathcal{D}_{n+2}^{(p-1)} - \big[\mathcal{D}_{n+1}^{(p)}\big]^2$$

A.9 The Gram-Fredholm formula.

A result whose discrete version is due to Gram was extended to integrals by Fredholm and goes as follows: Let $\phi_1, \ldots, \phi_p$ and $\psi_1, \ldots, \psi_p$ be two p-tuples of continuous functions on the unit interval. We get the $p \times p$-matrix with elements

$$a_{\nu k} = \int_0^1 \phi_\nu(x)\psi_k(x) \cdot dx$$

At the same time we define the following functions on $[0,1]^p$:

$$\Phi(x_1,\ldots,x_p) = \det \begin{pmatrix} \phi_1(x_1) & \cdots & \phi_1(x_p) \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ \phi_p(x_1) & \cdots & \phi_1(x_p) \end{pmatrix} \quad : \quad \Psi(x_1,\ldots,x_p) = \det \begin{pmatrix} \psi_1(x_1) & \cdots & \psi_1(x_p) \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ \psi_p(x_1) & \cdots & \psi_1(x_p) \end{pmatrix}$$

Product rules for determinants give the Gram-Fredholm equation

$$\text{(A.9.1)} \qquad \det(a_{\nu k}) = \frac{1}{p!} \int_{[0,1]^p} \Phi(x_1,\ldots,x_p) \cdot \Psi(x_1,\ldots,x_p) \cdot dx_1 \ldots dx_p$$

Exercise. Prove (A.9.1) or consult the literature. See for example the text-book [Bocher] which contains a detailed account about Fredholm determinants and their role for solutions to integral equations.

A.9.2 Resolvents of integral operators.

Fredholm studied integral equations of the form

$$(*) \qquad \phi(x) - \lambda \cdot \int_\Omega K(x,y) \cdot \phi(y) \cdot dy = f(x)$$

where Ω is a bounded domain in some euclidian space and the kernel function K is complex-valued. In general $K(x,y) \neq K(y,x)$. Various regularity conditions can be imposed upon the kernel. The simplest is when $K(x,y)$ is a continuous function in $\Omega \times \Omega$. The situation becomes more involved when singularities occur, for example when K is $+\infty$ on the diagonal, i.e. $|K(x,x)| = +\infty$. This occurs for example when K is derived from Green's functions which yield fundamental solutions to elliptic PDE-equations where corresponding boundary value problems are solved via integral equations. To obtain square integrable solutions in (*) for such singular kernels, functions, the determinants used by Fredholm were modified by Hilbert which avoid the singularities and are used to obtain resolvents of the integral operator $\mathcal{K}$ defined by

$$\mathcal{K}(\phi)(x) = \int_\Omega K(x,y) \cdot \phi(y) \cdot dy$$

where ϕ belong to the Hilbert space $L^2(\Omega)$. Hilbert studied the case when K is square integrable, i.e. when

$$(**) \qquad \iint_{\Omega\times\Omega} |K(x,y)|^2 \, dxdy < \infty$$

An eigenvalue to $\mathcal{K}$ is a complex number $\lambda \neq 0$ for which there exists a non-zero L^2-function ϕ such that

$$\mathcal{K}(\phi) = \lambda \cdot \phi$$

It is not difficult to show that (**) entails that the set of eigenvalues form a discrete set $\{\lambda_n\}$ with a sole cluster point at the origin in the complex λ-plane. In a famous article from 1909, Schur proved that (**) gives the inequality

$$(***) \qquad \sum \frac{1}{|\lambda_n|^2} \leq \iint_{\Omega\times\Omega} |K(x,y)|^2 \, dxdy$$

It means that the non-increasing sequence $\{|\lambda_1 \geq |\lambda_2| \geq \ldots\}$ miust tend to zero rather rapidly, where these eigenvalues are repeated when the eigenspaces have dimension ≥ 2.

A.9.3 Hilbert's determinants. Let K be a kernel function whose integral in (**) is finite. A typical case is that K is singular on the diagonal subset of $\Omega \times \Omega$ while it is contonous outside the diagonal. To each positive integer m one associates a pair of matrices of size $(m+1) \times m(+1)$ whose elements depend upon a pair $(\xi, \eta) \in \Omega \times \Omega$ and an m-tuple of distinct points $x_1, \ldots, x_m$ in Ω:

$$C_m^* = \begin{pmatrix} 0 & K(\xi,x_1) & K(\xi,x_2) & \ldots & \ldots & K(\xi,x_m) \\ K(x_1,\eta) & 0 & K(x_1,x_2) & \ldots & \ldots & K(x_1,x_m) \\ K(x_2,\eta) & K(x_2,x_1) & 0 & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ K(x_m,\eta) & K(x_m,x_1) & K(x_m,x_2) & \ldots & \ldots & 0 \end{pmatrix}$$

$$C_m = \begin{pmatrix} 0 & K(x_1,x_2) & & \ldots & 0 & K(x_1,x_m) \\ K(x_2,x_3) & 0 & K(x_2,x_3) & \ldots & & K(x_2,x_m) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ K(x_m,x_1) & K(x_m,x_2) & K(x_m,x_3) & \ldots & K(x_m,x_{m-1}) & 0 \end{pmatrix}$$

Put:

$$\text{(i)} \qquad D^*_m(\xi,\eta) = \int_{\Omega^m} C^*_m(\xi,\eta : x_1,\ldots,x_m)\cdot dx_1\cdots dx_m$$

$$\text{(ii)} \qquad D_m = \int_{\Omega^m} C_m(x_1,\ldots,x_m)\cdot dx_1\cdots dx_m$$

Thus, we take the integral over the m-fold product of Ω. Next, let λ be a new complex parameter and set

$$\mathcal{D}^*(\xi,\eta,\lambda) = \sum_{m=1}^{\infty} \frac{(-\lambda)^m}{m!}\cdot D^{**}_m(\xi,\eta)$$

$$\mathcal{D}(\lambda) = 1 + \sum_{m=1}^{\infty} \frac{(-\lambda)^m}{m!}\cdot D_m$$

Some results by Carleman

Using Fredholm-Hilbert determinants some conclusive facts about integral operators were established by Carleman in the article *Zur Theorie der Integralgleichungen* from 1921 when the kernel K is of the Hilbert-Schmidt type, i.e. (**) holds. Outside a discrete set in the complex λ-plane we have the inverse operator

$$R(\lambda) = (\lambda\cdot E - \mathcal{K})^{-1}$$

A.9.4 Theorem. *The kernel function of $R(\lambda)$ is for each complex λ outside the spectrum given by*

$$\Gamma(\xi,\eta;\lambda) = K(\xi,\eta) + \frac{\mathcal{D}^*(\xi,\eta,\lambda)}{\mathcal{D}(\lambda)}$$

Remark. Let $\{\lambda_\nu\}$ be the discrete spectrum of $\mathcal{K}$ where multiple eigenvalues are repeated when the corresponding eigenspaces have dimension ≥ 2. This spectrum constitutes the zeros of $\mathcal{D}(\lambda)$ and out turns out that thid is an entire function, i.e. analytic in the whole complex λ-plane. When λ is outside the spectrum the inverse operator $(\lambda\cdot E - \mathcal{K})^{-1}$ is an integral operator defined by

$$\phi \mapsto \int_\Omega \Gamma(\xi,\eta;\lambda)\cdot\phi(\eta)\,d\eta$$

where ξ and η are variable points in Ω. Using inequalities of Fredholm-Hadamard type for determinants, a first result from Carelan's cited article asserts that

$$\text{(A.9.5)} \qquad \int_\Omega \Gamma(\xi,\xi;\lambda)\cdot d\xi = -\lambda\cdot\sum_{\nu=1}^{\infty}\frac{1}{\lambda_\nu(\lambda-\lambda_\nu)}$$

where Schur's inequality from (***) in A.9.2 entails that the right hand side indeed represents a meromorphic function. A major result in Carelan'as article is about the entire function $\mathcal{D}(\lambda)$.

A.9.5 Theorem. *$\mathcal{D}(\lambda)$ is an entire function of the complex parameter λ given by a Hadamard product*

$$\text{(1)} \qquad \mathcal{D}(\lambda) = \prod(1-\frac{\lambda}{\lambda_n})\cdot e^{\frac{\lambda}{\lambda_n}}$$

where $\{\lambda_n\}$ satisfy

$$\text{(2)} \qquad \sum|\lambda_n|^{-2} < \infty$$

Remark. Prior to this Schur proved a representation as in (1) adding a factor $e^{b\lambda^2}$ in the right hand side. So the novelty in Carleman's work is that $b=0$ always hold. A crucial step in Carleman's proof of (1) was to use an inequality for determinants which goes as follows: Let

$q > p \geq 1$ be a pair of integers and $\{a_{k,\nu}\}$ a doubly-indexed sequence of complex numbers which appear as elements in a $p+q$-matrix of the form:

$$(*) \qquad \begin{pmatrix} 0 & \dots & 0 & a_{1,p+1} & \dots & a_{1,p+q} \\ 0 & \dots & 0 & a_{2,p+1} & \dots & a_{1,p+q} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{p,p+1} & \dots & a_{p,p+q} \\ a_{p+1,1} & \dots & a_{p+1,p} & a_{p+1,p+1} & \dots & a_{p+1,p+q} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{p+q,1} & \dots & a_{p+q,p} & a_{p+q,p+1} & \dots & a_{p+q,p+q} \end{pmatrix}$$

For each pair $1 \leq m \leq p$ we put

$$L_m = \sum_{\nu=1}^{\nu=q} |a_{m,p+\nu}|^2 \quad : \quad S_m = \sum_{\nu=1}^{\nu=q} |a_{p+\nu,m}|^2 \quad : \quad N = \sum_{j=1}^{j=q} \sum_{\nu=1}^{\nu=q} |a_{p+j,p+\nu}|^2$$

A.9.6 Theorem. *Let Δ be the determinant of the matrix (*). Then*

$$|\Delta| \leq (L_1 \cdots L_p)^{\frac{1}{p}} \cdot \sqrt{M_1 \cdots M_p} \cdot \frac{N^{\frac{q-p}{2}}}{(q-p)^{\frac{q-p}{2}}}$$

Proof. After unitary transformations of the last q rows and the last q columns respectively, the proof is reduced to the case when $a_{jk} = 0$ for pairs (j,k) with $j \leq p$ and $k > p+j$ or with $k \leq p$ and $j > p+k$. Here L_m, S_m and N are unchanged and we get

$$\Delta = (-1)^p \cdot \prod_{j=1}^{j=p} a_{j,p+j} \cdot \prod_{k=1}^{k=p} a_{p+k,k} \cdot \det \begin{pmatrix} a_{p+1,2p+1} & \dots & a_{p+1,p+q} \\ \dots & \dots & \dots \\ a_{p+q,p+1} & \dots & a_{p+q,p+q} \end{pmatrix}$$

The absolute value of the last determinant is majorized by Hadamard's inequality in § F.XX and the requested inequality in Theorem A.9.6 follows.

§ 1. Wedderburn's theorem.

A finite dimensional $\mathbf{C}$-algebra $\mathcal{A}$ is an associative ring which contain $\mathbf{C}$ as a central subfield, i.e. $\lambda\cdot a = a\cdot\lambda$ for pairs $a \in \mathcal{A}$ and complex numbers λ. The ring product gives the family of left ideals. They consist of complex subspaces L which are stable under left multiplication, i.e. $aL \subset L$ hold for every element a in $\mathcal{A}$. One may also regard two-sided ideals J where one requires that both aJ and Ja are contained in J for every $a \in \mathcal{A}$. One says that $\mathcal{A}$ is called a simple algebra if the sole 2-sided ideals are $\mathcal{A}$ and the trivial zero ideal. Examples of finite dimensional $\mathbf{C}$-algebras are matrix-algebras $\{M_n(\mathbf{C} : \ n \geq 1$. It turns out that they give the sole simple algebras.

1.1 Theorem. *Let A be a finite dimensional and simple $\mathbf{C}$-algebra. Then there exists an integer n such that*

$$A \simeq M_n(\mathbf{C})$$

The proof requires several steps. Let us first show that the matrix algebras are simple. With $n \geq 2$ we put $\mathcal{A} = M_n(\mathbf{C})$ and identify every matrix with a linear operator on $\mathbf{C}^n$. Conioder the special matrices $e_1, \ldots, e_n$ where the elements of e_p are zero except for the diagonal element $a_{pp} = 1$. Product rukes for matrices show that

$$e_p e_q = 0 \quad : p \neq q$$

At thes same time we notice that the identity element is $e_1 + \ldots + e_n$ and that $e_p = e_p^2$. This means that $\{e_p\}$ are pairwise orthogonal idempotent elements. Wirh p fixed we have the left ideal

$$L_p = Ae_p$$

The reader can check that it consists of all matrices whose columns of degree $q \neq p$ are zero. The left ideal L_p is minimal. For let ξ be a non-zero matrix in L_p which means that there exists at least some ν where the matrix element $a_{\nu p} \neq 0$. mutlplying with a sclar we can assume that $a_{\nu p} = 1$ and then we see that

$$e_\nu \cdot \xi = e_p$$

Hence the principal left ideal generated by ξ is equal to L_p, i.e. every non-zero element in L_p generates L_p which means that this left ideal is minimal. Next, let $\xi = \{a_{qp}\}$ be a non-zero matrix. and choose p so that $a_{qp} \neq 0$ for at least one q. Then $\xi \cdot e_p$ is a non-zero element in L_p so the 2-sided ideal generated by ξ contains the minimal left ideal L_p. If we consider another integer q we take the matrix ξ with a single non-zero element placed at (p, q) which is equal to one. Then we see that $e_p \cdot \xi = e_q$ and hence the 2-sided ideal contains $L_q = Le_q$, Since this hold for every $1 \leq q \leq n$ the reader mat conclude that the 2-sided ideal generated by ξ is the whole ring A. This proves that A is simple.

Next, let ξ be a non-zero matrix which commutes with all other matrices. To prove that ξ is a complex multiple of the identity matrix one argues as follows: The matrix elements of ξ are $\{a_{\nu k}\}$. For a given p the product $\xi \cdot e_p$ gives a matrix with a single non-zero column put in place p with elements $\{a_{\nu p}\}$. At the same time $e_p \cdot \xi$ is a matrix with a single non-zero row placed in degree p. So the equality $e_p\xi = \xi e_p$ entails that

$$a_{qp} = 0 \quad : q \neq p$$

So when ξ commnutes with all the e-matrices we concude that ξ is a diagonal matrix, i.e. the elements outside the diagonal are all zero. There remain to see that the diagonal elements are all equal. Suppose for example that $a_{11} \neq a_{22}$. Now there exists the matrix β where $b_{12} = b_{21} = 1$ and all other elements are zero. Then we see that

$$\beta \cdot |xi = a_{11}e_2 + a_{22}e_1 \quad : \quad \xi \cdot \beta = a_{11} \cdot e_1 + a_{22} \cdot e_2$$

Hence the equality $\xi\beta = \beta\xi$ entails that $a_{11} = a_{22}$. In the same way one proves that all diagoal elements are equal. This proves that the center of the matrix algebra is reduced to complex multiples of the identity.

1.1. Exercise. Set $\mathcal{A} = M_n(\mathbf{C})$ and identify every matrix with a $\mathbf{C}$-linear operator on $\mathbf{C}^n$. To each left ideal L we assign the null space

$$L^\perp = \{v \in \mathbf{C}^n : L(v) = 0\}$$

Thus one takes the intersection of the null spaces of operators from L. Show that $L^\perp$ determines L in the sense that a matrix Q belongs to L if and only if its null-space contains $L^\perp$. So by

$$L \mapsto L^\perp$$

one has a bijective map between the family of left ideals in the matrix algebra and subspaces of $\mathbf{C}^n$.

Next, let p is the dimension of the vector space $L^\perp$. Show that dimension of L regared as a complex vector space is equal to $n(n-p)$.

For each $1 \leq p \leq n-1$ we get the family $\mathcal{L}_p$ of left ideals L for which $L^\perp$ has dimension $n-p$. The family of $(n-p)$-dimensional subspaces of $\mathbf{C}^n$ is denoted by $\mathcal{G}(n-p;n)$ and called the Grassmannian of degree $n-p$. From the above one has the set-theoretic equality

$$\mathcal{G}(n-p;n) = \mathcal{L}_p \tag{1.1.2}$$

Next, show that for every left ideal L there exists an idempotent matrix Π such that

$$L = \mathcal{A} \cdot \Pi \tag{i}$$

Thus, every left ideal is generated by a single matrix and (i) means that

$$L^\perp = \mathrm{Ker}(\Pi)$$

When $L \in \mathcal{L}_p$ the Π-kernel has dimension $n-p$ and this implies that the range of π is p-dimensional. Moreover, since $\Pi^2 = \Pi$ one has a direct sum decomposition

$$\mathbf{C}^n = L^\perp \oplus \Pi(\mathbf{C}^n) \tag{ii}$$

Proof of 1.1 Theorem.

Denote by $\mathcal{L}_*$ the family of non-zero left ideals L in A which are minimal in the sense that every non-zero left ideal $L_1 \subset L$ is equal to L. Since A is a finite dimensional vector space it is clear that there exists at least one minimal left ideal L. Identifying L with a complex vector space it has some dimension k, and we get the $\mathbf{C}$-algebra

$$\mathcal{M} = \mathrm{Hom}_{\mathbf{C}}(L, L)$$

Choosing a basis in the complex vector space L one has

$$\mathcal{M} \simeq M_k(\mathbf{C})$$

Wedderburn's theorem follows if we prove that

(*) $$\mathcal{M} \simeq A$$

To get (*) we first consider some $a \in A$ which by left multiplication gives a map

$$a^* \colon x \mapsto ax \quad : \quad x \in L$$

Since $\mathbf{C}$ by assumption is a central subfield of A these maps are complex linear and hence a^* is an element in $\mathcal{M}$. If b is another element in A we get the composed linear operator $b^* \circ a^*$ defined by

$$x \mapsto bax = (ba)^*(x)$$

Hence

(i) $$a \mapsto a^*$$

is an algebra homomorphism from A into $\mathcal{M}$. We claim that this map is injective. For if a^* is the zero map we use that L is a left ideal which gives

$$ax\xi = 0$$

for all $x \in A$ and $\xi \in L$. This gives $a^* \circ x^* = 0$ and since it is obvious that $x^* \circ a^* = 0$ also holds, we conclude that the kernel of (i) is a 2-sided ideal in A. Since A is simple this kernel is zero which proves that (i) is injective.

Proof of surjectivity. If $x \in A$ is such that $Lx \neq 0$ then this is a non-zero left ideal and since L is minimal the reader can check that we also have $Lx \in \mathcal{L}_*$ and the simple left A-modues L and Lx are isomorphic, i.e.

(ii) $$L \simeq Lx$$

hold whenever $Lx \neq 0$. Next, by assumption the 2-sided ideal generated by L is the whole ring A. Hence there exists a finite set of A-elements $\{x_\nu\}$ such that

(iii) $$A = Lx_1 + \ldots + Lx_m$$

Above we can choose m minimal which gives a direct sum

(iv) $$A = Lx_1 \oplus \ldots \oplus Lx_m$$

For suppose that

$$\xi_1 x_1 + \ldots + \xi_m x_m = 0 \quad : \xi_\nu \in L$$

where $\xi_k x_k \neq 0$ for some k. Since Lx_k is minimal the resder can chek that Lx_k now can be deleted in (iii) which contradicts the minmal chose of m. Hence one has the direct sum in (iv). By (ii) the vector spaces L and Lx_k are isomorphic for every $1 \leq k \leq m$. Counting dimensions we conclude that

(v) $$\dim_{\mathbf{C}} A = m \cdot k$$

Since the map from A into $\mathcal{M}$ was injective and $\mathcal{M}$ is a matrix algebra of dimension k^2, thr injectivity entails that $m \leq k$, and there remains to prove the opposite inequality

(*) $$k \leq m$$

To get (*) we take the identity element 1 in A and (iv) gives an m-tuple $\{\xi_\nu\}$ in L so that

(1) $$1 = \xi_1 x_1 + \cdots + \xi_k x_m$$

Put $e_\nu = \xi_\nu \cdot x_\nu$. Multiplying to the left by some e_k in (1) we get

$$e_k = e_k e_1 + \ldots + e_k e_m$$

The direct sum in (iv) entials that

$$e_k e_\nu = 0 : \nu \neq k \quad \& \quad e_k e_k = e_k$$

Hence $\{e_\nu\}$ are pairwise orthogonal idempotent elements in A. For a fixed k the equality $e_k = e_k^2$ entails that $e_k \cdot A \cdot e_k$ is a subalgebra of A. If $x = e_k \cdot x \cdot e_k$ is an element in this subalgebra then right multiplication by x on the left ideal Ae_k is left A-linear, i.e. one has a map

(2) $$e_k \cdot A \cdot e_k \to \mathrm{Hom}_A(Ae_k, Ae_k)$$

Now we use that Ae_k is a minimal left ideal, i.e. as a left A-module it is simple. This implies that the right hand side in (2) is a division ring, i.e. every non-zero element is invertible, Since the complex field is algebraically closed this division ring is equal to $\mathbf{C}$. Moreover, if $\xi = e_k x e_k$ is such that its image in (2) is zero, then

$$e_k \xi = e_k^2 x e_k = e_k x e_k = \xi = 0$$

So (2) is injective and hence

(3) $$e_k A e_k = \mathbf{C}$$

Let us now take some $j \neq k$ and consider the space

(4) $$\mathrm{Hom}_A(Ae_j, Ae_k)$$

Every left A-linear map from Ae_j into Ae_k is induced by right multiication with an element ξ and since e_j and e_k are idempotens one has

$$\xi = e_j \xi e_k$$

Conversly, every $x \in A$ gives gets a ξ-element $a_k x e_j$. Hence the vector space (4) can be identfied with the subset of A given by

(5) $$e_j A e_k$$

We have already seen that the left A-modukes generated by e_k and e_j are isomorphic and then (3) entails that

(6) $$\dim_{\mathbf{C}}(e_j A e_k) = 1$$

Now (*) follows because with $L = Ae_1$ one has

$$L = \sum_{j=1}^{j=m} e_j A e_1$$

which proves thst the k-dimensional vector space L has dimension m at most which gives the inequality (*) and finishes the proof of Wedderburn's theorem.

Remark. For readers interested in ring theory we recall that Wedderburn's theorem is more general, i.e. it describes every simple associative ring A which in addition is *artinian*, i.e. the descendinig chain condition holds for the family of left ideals. Every simple and artinian ring is isomorphic to a matrix ring $M_k(D)$ where D in general is a division ring. The proof is verbatim the same as above where the sole difference is that we now get the division ring D via (3) in the previous proof. A more extensive family of rings is found as follows: An associative ring B with a unit elementis called *semi-primsry* if it contains a tweo-sided ifeal $J(B)$ which is nilpotent while the quotent ring

$$\frac{B}{J(B)}$$

is a semi-simle artinian ring which via Wdderburin's theorem means thst this ring contains a finite family of central and orthogonal idempotents $e_1, \ldots, e_k$ such that the identity $1 = e_1 + \ldots + e_k$ and the subrings $\{e_\nu \cdot \frac{B}{J(B)} \cdot e_\nu\}$ are simple artinian rings. One reason why this extended family of rings is interesting is that they srise via simle artinian rings while one regards invariant subrings. A general result of this kind which is due to the present author goes as follows:

Theorem. *Let A be a simple artinian ring and $\mathcal{E}$ a family of left A-linear maps on A into itself. Then the invariant subring below is semi-priumary.*¨

$$A_{\mathcal{E}} = \{x \in A : E(x) = x : \; E \in \mathcal{E}\}$$

2. Resolvents

Let A be some matrix in $M_n(\mathbf{C})$. Its characteristic polynomial is defined by

$$(*)\qquad P_A(\lambda) = \det(\lambda \cdot E_n - A)$$

By the fundamental theorem of algebra P_A has n roots $\alpha_1, \ldots, \alpha_n$ where eventual multiple roots are repeated. The union of distinct roots is denoted by $\sigma(A)$ and called the spectrum of A. Since matrices with non-zero determinants are invertible we obtain a matrix valued function defined in $\mathbf{C} \setminus \sigma(A)$ by:

$$(**)\qquad R_A(\lambda) = (\lambda \cdot E_n - A)^{-1} \quad : \quad \lambda \in \mathbf{C} \setminus \sigma(A)$$

One refers to $R_A(\lambda)$ as the resolvent of A. The map

$$\lambda \mapsto R_A(\lambda)$$

yields a matrix-valued analytic function defined in $\mathbf{C} \setminus \sigma(A)$. To see this we take some $\lambda_* \in \mathbf{C} \setminus \sigma(A)$ and set

$$R_* = (\lambda_* \cdot E_n - A)^{-1}$$

Since R_* is a 2-sided inverse we have the equality

$$E_n = R_*(\lambda_* \cdot E_n - A) = (\lambda_* \cdot E_n - A) \cdot R_* \implies R_* A = A R_*$$

Hence the resolvent R_* commutes with A. Next, construct the matrix-valued power series

$$(1)\qquad \sum_{\nu=1}^{\infty} (-1)^\nu \cdot \zeta^\nu \cdot (R_* A)^\nu$$

which is convergent when $|\zeta|$ are small enough.

2.1 Exercise. Prove the equality

$$R_A(\lambda_* + \zeta) = R_* + \sum_{\nu=1}^{\infty} (-1)^\nu \cdot \zeta^\nu \cdot R_* \cdot (R_* A)^\nu$$

The local series expansion () above therefore shows that the resolvents yield a matrix-valued analytic function in $\mathbf{C} \setminus \sigma(A)$.

We are going to use analytic function theory to establish results which after can be extended to an operational calculus for linear operators on infinite dimensional vector spaces. The analytic constructions are also useful to investigate dependence upon parameters. Here is an example. Let A be an $n \times n$-matrix whose characteristic polynomial $P_A(\lambda)$ has n simple roots $\alpha_1, \ldots, \alpha_n$. When λ is outside the spectrum $\sigma(A)$. residue calculus gives the following expression for the resolvents:

$$(*)\qquad (\lambda \cdot E_n - A)^{-1} = \sum_{k=1}^{k=n} \frac{1}{\lambda - \alpha_k} \cdot \mathcal{C}_k(A)$$

where each matrix $\mathcal{C}_k(A)$ is a polynomial in A given by:

$$\mathcal{C}_k(A) = \frac{1}{\prod_{\nu \neq k} (\alpha_k - \alpha_\nu)} \cdot \prod_{\nu \neq k} (A - \alpha_\nu E_n)$$

The formula (*) goes back to work by Sylvester, Hamilton and Cayley. The resolvent $R_A(\lambda)$ is also used to construct the Cayley-Hamilton polynomial of A which by definition this is the unique monic polynomial $P_*(\lambda]$ in the polynomial ring $\mathbf{C}[\lambda]$ of smallest possible degree such that the associated matrix $p_*(A) = 0$. It is found as follows: Let $\alpha_1, \ldots, \alpha_k$ be the distinct roots of $P_A(\lambda)$ so that

$$P_A(\lambda) = \prod_{\nu=1}^{\nu=k} (\lambda - \alpha_\nu)^{e_\nu}$$

where $e_1 + \ldots + e_k = n$. Now the meromorphic and matrix-valued resolvent $R_A(\lambda)$ has poles at $\alpha_1, \ldots, \alpha_k$. If the order of a pole at root α_j is denoted by ρ_j one has the inequality $\rho_j \leq e(\alpha_j)$ which in general can be strict. The Cayley-Hamilton polynomial becomes:

$$P_*(\lambda) = \prod_{\nu=1}^{\nu=k} (\lambda - \alpha_\nu)^{\rho_\nu} \tag{**}$$

Now we begin to prove results in more detail. To begin with one has the Neumann series expansion:

Exercise. Show that if $|\lambda|$ is strictly larger than the absolute values of the roots of $P_A(\lambda)$, then the resolvent is given by the series

$$R_A(\lambda) = \frac{E_n}{\lambda} + \sum_{\nu=1}^{\infty} \lambda^{-\nu-1} \cdot A^\nu \tag{*}$$

A differential equation. Taking the complex derivative of $\lambda \cdot R_A(\lambda)$ in (*) we get

$$\frac{d}{d\lambda}(\lambda R_A(\lambda)) = -\sum_{\nu=1}^{\infty} \nu \cdot \lambda^{-\nu-1} \cdot A^\nu \tag{1}$$

Exercise. Use (1) to prove that if $|\lambda|$ is large then $R_A(\lambda)$ satisfies the differential equation:

$$\frac{d}{d\lambda}(\lambda R_A(\lambda)) + A[\lambda^2 R_A(\lambda) - E_n - \lambda A] = 0 \tag{2}$$

Now (2) and the analyticity of the resolvent outside the spectrum of A give:

2.3 Theorem *Outside the spectrum $\sigma(A)$ $R(\lambda)$ satisfies the differential equation*

$$\lambda \cdot R_A'(\lambda) + R_A(\lambda) + \lambda^2 \cdot A \cdot R_A(\lambda) = A + \lambda \cdot A^2$$

2.4 Residue formulas. Since the resolvent is analytic we can construct complex line integrals and apply results in complex residue calculus. Start from the Neumann series (*) above and perform integrals over circles $|\lambda| = w$ where w is large.

2.5 Exercise. Show that when w is strictly larger than the absolute value of every root of $P_A(\lambda)$ then

$$A^k = \frac{1}{2\pi i} \int_{|\lambda|=w} \lambda^k \cdot R_A(\lambda) \cdot d\lambda \quad : \quad k = 1, 2, \ldots$$

It follows that when $Q(\lambda)$ is an arbitrary polynomial then

$$Q(A) = \frac{1}{2\pi i} \int_{|\lambda|=w} Q(\lambda) \cdot R_A(\lambda) \cdot d\lambda \tag{*}$$

In particular we take the identity $Q(\lambda) = 1$ and obtain

$$E_n = \frac{1}{2\pi i} \cdot \int_{|\lambda|=w} R_A(\lambda) \cdot d\lambda \tag{**}$$

Finally, show that if $Q(\lambda)$ is a polynomial which has a zero of order $\geq e(\alpha_\nu)$ at every root then

$$Q(A) = 0 \tag{***}$$

2.6 Residue matrices. Let $\alpha_1, \ldots, \alpha_k$ be the distinct zeros of $P_A(\lambda)$. For a given root, say α_1 of multiplicity $p \geq 1$ we have a local Laurent series expansion

$$R_A(\alpha_1 + \zeta) = \frac{G_p}{\zeta^p} + \ldots + \frac{G_1}{\zeta} + B_0 + \zeta \cdot B_1 + \ldots \tag{i}$$

We refer to $G_1,\ldots,G_p$ as the residue matrices at α_1. Choose a polynomial $Q(\lambda)$ in $\mathbf{C}[\lambda]$ which vanishes up to the multiplicity at all the the remaining roots $\alpha_2,\ldots,\alpha_k$ while it has a zero of order $p-1$ at α_1, i.e. locally

(i) $$Q(\alpha_1+\zeta)=\zeta^{p-1}(1+q_1\zeta+\ldots)$$

2.7 Exercise. Use residue calculus and (*) from Exercise 2.5 to show that:

(*) $$Q(A)=\frac{1}{2\pi}\int_{|\lambda-\alpha_1|=\epsilon}Q(\lambda)\cdot R_A(\lambda)\cdot d\lambda=G_p$$

Hence the matrix G_p is a polynomial of A. In a similar way one proves that every G-matrix in the Laurent series (i) is a polynomial in A.

2.7 Some idempotent matrices. Consider a zero α_j and choose a polynomial Q_j such that $Q_j(\lambda)-1$ has a zero of order $e(\alpha_j)$ at α_j while Q_j has a zero of order $e(\alpha_\nu)$ at the remaining roots. Set

(1) $$E_A(\alpha_j)=\frac{1}{2\pi i}\int_{|\lambda|=w}Q_j(\lambda)\cdot R_A(\lambda)\cdot d\lambda$$

where w is large as in 2.5. Since the polynomial $S=Q_j-Q_j^2$ vanishes up to the multiplicities at all the roots of $P_A(\lambda$ we have $S(A)=0$ from (***) in 2.5 which entails that

(*) $$E_A(\alpha_j)=E_A(\alpha_j)\cdot E_A(\alpha_j)$$

In other words, we have constructed an idempotent matrix.

2.8 The Cayley-Hamilton decomposition. Recall the equality

$$E_n=\frac{1}{2\pi i}\cdot\int_{|\lambda|=w}R_A(\lambda)\cdot d\lambda$$

where the radius w is so large that the disc D_w contains the zeros of $P_A(\lambda)$. The previous construction of the E-matrices at the roots of $P_A(\lambda)$ entail that

$$E_n=E_A(\alpha_1)+\ldots+E_A(\alpha_k)$$

Identifying A with a $\mathbf{C}$-linear operator on $\mathbf{C}^n$ we obtain a direct sum decomposition

(*) $$\mathbf{C}^n=V_1\oplus\ldots\oplus V_k$$

where each V_ν is an A-invariant subspace given by the image of $E_A(\alpha_\nu)$. Here $A-\alpha_\nu$ restricts to a *nilpotent* linear operator on V_ν and the dimension of this vector space is equal to the multiplicity of the root α_ν of the characteristic polynomial. One refers to (*) as the *Cayley-Hamilton decomposition* of $\mathbf{C}^n$.

2.9 The vanishing of $P_A(A)$. Consider the characteristic polynomial $P_A(\lambda)$. By definition it vanishes up to the order of multiplicity at every point in $\sigma(A)$ and hence (***) in 2.5 gives $P_A(A)=0$. Let us write:

$$P_A(\lambda)=\lambda^n+c_{n-1}\lambda^{n-1}+\ldots+c_1\lambda+c_0$$

Notice that $c_0=(-1)^n\cdot\det(A)$. So if the determinant of A is $\neq 0$ we get

$$A\cdot[A^{n-1}+c_{n-1}A^{n-2}+\ldots+c_1]=(-1)^{n-1}\det(A)\cdot E_n$$

Hence the inverse A^{-1} is expressed as a polynomial in A. Concerning the equation

$$P_A(A)=0$$

it is in general not the minimal equation for A, i.e. it can occur that A satisfies an equation of degree $<n$. More precisely , if α_ν is a root of some multiplicity $k\geq 2$ there exists a Jordan decomposition which gives an integer $k_*(\alpha_\nu)$ for the largest Jordan block attached to the nilpotent operator $A-\alpha_\nu$ on V_{α_ν}. The *reduced* polynomial $P_*(\lambda)$ is the product where the factor $(\lambda-\alpha_\nu)^{k_\nu}$ is replaced by $(\lambda-\alpha_\nu)^{k_*(\alpha_\nu)}$ for every α_ν where $k_\nu<k_*(\alpha_\nu)$ occurs. Then P_* is the polynomial

of smallest possible degree such that $P_*(A) = 0$. One refers to P_* as the *Hamilton polynomial* attached to A. This result relies upon Jordan's result in § 3.

2.10 Similarity of matrices. Recall that the determinant of a matrix A does not change when it is replaced by SAS^{-1} where S is an arbitrary invertible matrix. This implies that the coefficients of the characterstic polynomial $P_A(\lambda)$ are intrinsically defined via the associated linear operator, i.e. if another basis is chosen in $\mathbf{C}^n$ the given A-linear operator is expressed by a matrix SAS^{-1} where S effects the change of the basis. Let us now draw an interesting consequence of the previous operational calculus. Let us give the following:

2.11 Definition. *A pair of $n \times n$-matrices A, B are* similar *if there exists some invertible matrix S such that*

$$B = SAS^{-1}$$

Since the product of two invertible matrices is invertible this yield an equivalence relation on $M_n(\mathbf{C})$ and from 2.2 above we conclude that $P_A(\lambda)$ only depends on its equivalence class. The question arises if to matrices A and B whose characteristic polynomials are equal also are similar in the sense of Definition 2.6. This is not true in general. More precisely, *Jordan normal form* determines the eventual similarity between a pair of matrices with the same characteristic polynomial.

3. Jordan's normal form

Introduction. Theorem 3.1 below is due to Camille Jordan. It plays an important role when we discuss multi-valued analytic functions in punctured discs and is also used in ODE-theory. Jordan's theorem says that every equivalence class in $M_n(\mathbf{C})$ contains a matrix which is built up by Jordan blocks which are defined below. The proof employs the Cayley-Hamilton decomposition from 2.7. which shows that an arbitrary $n \times n$-matrix A has a similar matrix $B = S^{-1}AS$ represented in a block form. More precisely, to every root α_ν of $P_A(\lambda)$ of some multiplicity e_ν there occurs a square matrix B_ν of size e_ν and α_ν is the only root of $P_{B_\nu}(\lambda)$. It follows that for every fixed ν one has

$$B_\nu = \alpha \cdot E_{k_\nu} + S_\nu$$

where E_{k_ν} is an identity matrix of size k_ν and S_ν is nilpotent, i.e. there exists an integer m such that $S_\nu^m = 0$. Jordan's theorem gives a further decomposition of these nilpotent S-matrices.

3.0 Jordan blocks. An *elementary* Jordan matrix of size 4 is matrix of the form

$$\begin{pmatrix} \lambda & 0 & 0 & 0 \\ 1 & \lambda & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & \lambda \end{pmatrix}$$

where λ is the eigenvalue. For $k \geq 5$ one has similar expressions. In general several elementary Jordan block matrices build up a matrix which is said to be in Jordan's normal form.

3.1 Theorem. *For every matrix A there exists an invertible matrix u such that UAU^{-1} is in Jordan's normal form.*

Proof. By the remark after Proposition 2.12 it suffices to prove Jordan's result when A has a single eigenvalue α. Replacing A by $A - \alpha$ there remains only to consider the nilpotent case, i.e when $P_A(\lambda) = \lambda^n$ so that $A^n = 0$ and then we must find a basis where A is represented in Jordan's normal form.

3.2 Nilpotent operators. Let S be a nilpotent $\mathbf{C}$-linear operator on some n-dimensional complex vector space V. So for each non-zero vector in $v \in V$ there exists a unique integer m such that

$$S^m(v) = 0 \quad \& \quad S^{m-1}(v) \neq 0$$

The unique integer m is denoted by $\text{ord}(S, v)$. The case $m = 1$ occurs if $S(v) = 0$. If $m \geq 2$ the reader can check that the vectors $v, S(v), \ldots, S^{m-1}(v)$ are linearly independent The vector space generated by this m-tuple is denoted by $\mathcal{C}_S(v)$ and called the *cyclic* subspace of V generated by v. With these notations the reader cabn check that Jordan's theorem amounts to prove the following:

3.3 Proposition *Let S be a nilpotent linear operator. Then V is a direct sum of cyclic subspaces.*

Proof. Set

$$m^* = \max_{v \in V} \operatorname{ord}(S, v)$$

and choose a vector $v^* \in V$ such that $\operatorname{ord}(S, v^*) = m^*$. On the quotient space $W = \frac{V}{\mathcal{C}_S(v^*)}$ we notice that S induces a linear operator which we denote by $\bar{S}$. By induction over $\dim(V)$ we may assume that W is a direct sum of cyclic subspaces. Hence we can pick a finite set of vectors $\{v_\alpha\}$ in V such that if $\{\bar{v}_\alpha\}$ are the images in W, then

$$W = \oplus \mathcal{C}_{\bar{S}}(\bar{v}_\alpha) \tag{1}$$

For each v_α we put

$$k_\alpha = \operatorname{ord}(\bar{S}, \bar{v}_\alpha)$$

The construction of a quotient space means that

$$S^{k_\alpha}(v_\alpha) \in \mathcal{C}_S(v^*) \tag{2}$$

Hence there exists a unique m^*-tuple $c_0, \ldots, c_{m-1}$ in $\mathbf{C}$ such that

$$S^{k_\alpha}(v_\alpha) = c_0 \cdot v^* + c_1 \cdot S(v^*) + \ldots + c_{m^*-1} \cdot S^{m^*-1}(v^*) \tag{3}$$

Put

$$k^*_\alpha = \operatorname{ord}(S, v_\alpha) \tag{4}$$

It is obvious that $k^*_\alpha \geq k_\alpha$ and (3) gives

$$0 = S^{k^*_\alpha}(v_\alpha) = \sum c_\nu \cdot S^{k^*_\alpha - k_\alpha + \nu}(v^*) \tag{5}$$

The maximal choice of m^* entails that $k^*_\alpha \leq m^*$. Since the vectors $v^*, S(v^*), \ldots S^{m^*-1}(v^*)$ are linearly independent we see that (5) gives

$$c_0 = \ldots = c_{k_\alpha - 1} = 0 \tag{5}$$

Hence (3) gives a vector $w_\alpha \in \mathcal{C}_S(v^*)$ such that

$$S^{k_\alpha}(v_\alpha) = S^{k_\alpha}(w_\alpha) \tag{6}$$

The images of v_α and $v_\alpha - w_\alpha$ are equal in $\mathcal{C}(v^*)$. Replacing $\{v_\alpha\}$ by the vectors $\{\xi_\alpha = v_\alpha - w_\alpha\}$ one still has

$$W = \oplus \mathcal{C}_{\bar{S}}(\bar{\xi}_\alpha) \tag{7}$$

and the construction of the ξ-vectors give

$$\operatorname{ord}(\bar{S}, \bar{\xi}_\alpha) = \operatorname{ord}(S, v_\alpha) \tag{8}$$

for each α. At this stage an obvious counting of dimensions give the requested direct sum decomposition

$$V = \mathcal{C}_S(v^*) \oplus \mathcal{C}_S(\xi_\alpha)$$

Remark. The proof was bit cumbersome. The reason is that the direct sum decomposition in Jordan's Theorem is not unique. Only the individual *dimensions* of the cyclic subspaces which appear in a direct sum decomposition are unique. It is instructive to perform Jordan decompositions of specific matrices using an implemented program which for example can be found in *Mathematica*.

4. Hermitian operators.

The n-dimensional vector space $\mathbf{C}^n$ is equipped with the hermitian inner product:

$$\langle x, y\rangle = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n$$

A basis $e_1, \ldots, e_n$ is orthonormal if $\langle e_i, e_k\rangle =$ Kronecker's delta function. To each linear operator A the adjoint A^* is the linear opeartor for which

$$\langle A(x), y\rangle = \langle x, A^*(y)\rangle$$

4.0. Hermitian operators. A linear operator A is called Hermitian if

$$\langle A(x), y\rangle = \langle x, A(y)\rangle$$

holds for all x and y. An equivalent condition is that A is equal to its adjoint A^*. Therefore one also refers to a self-adjoint operator, i.e the notion of a hermitian respectively self-adjoint matrix is the same.

4.1 Unitary operators. A linear operator U is unitary if it preserves the inner product:

$$\langle U(x), U(y)\rangle = \langle x, y\rangle$$

for all x and y. It is clear that a unitary operator U sends an orthonormal basis to another orthonormal basis and the reader may verify that a linear operator U is unitary if and only if

$$U^{-1} = U^*$$

4.3 Self-adjoint projections. Let V be a subspace of $\mathbf{C}^n$ of some dimension $1 \leq k \leq n-1$. Its orthogonal complement is denoted by $V^\perp$ and we have the direct sum decomposition

$$\mathbf{C}^n = V \oplus V^\perp$$

To V we associate the linear operator E whose kernel is $V^\perp$ while it restricts to the identity on V. Here

$$E = E^2 \quad \text{and} \quad E = E^*$$

One refers to E as a self-adjoint projection.

4.4 Orthonormal bases. Let $V_1 \subset V_2 \subset \ldots V_n = \mathbf{C}^n$ be a strictly increasing sequence of subspaces. So here each V_k has dimension k. The *Gram-Schmidt orthogonalisation* yields an orthonormal basis $\xi_1, \ldots, \xi_n$ such that

$$V_k = \mathbf{C} \cdot \xi_1 + \ldots + \mathbf{C} \cdot \xi_k$$

hold for every k. The verification of this wellknown construction is left to the reader. Next, if A is an arbitrary $n \times n$-matrix the fundamental theorem of algebra implies that there exists a sequence $\{V_k\}$ as above such that every V_k is A-invariant, i.e.

$$A(V_k) \subset V_k$$

hold for each k. We find the orthonormal basis $\{\xi_k\}$ and construct the unitary operator U which sends the standard basis in $\mathbf{C}^n$ onto this ξ-basis. In this ξ-basis we see that the linear operator A is represented by an upper triangular matrix. Hence we have

4.5 Theorem. *For every $n \times n$-matrix A there exists a unitary matrix U such that U^*AU is upper triangular.*

The spectral theorem.

It asserts the following:

4.6 Theorem. *If A is Hermitian there exists an orthonormal basis $e_1, \ldots, e_n$ in $\mathbf{C}^n$ where each e_k is an eigenvector to A whose eigenvalue is a real number. Thus, A can be diagonalised in an orthonormal basis and expressed by matrices this means that there exists a unitary matrix U such that*

$$(*) \qquad U^* A U = S$$

where S is a diagonal matrix and every s_{ii} is a real number. In particular the roots of the characteristic polynomial $det(P_A(\lambda))$ are all real.

Proof. Since A is self-adjoint we have a real-valued function on $\mathbf{C}^n$ defined by

$$(1) \qquad x \mapsto \langle Ax, x \rangle$$

Let m^* be the maximum of (1) as x varies over the compact unit sphere of unit vectors in $\mathbf{C}^n$. The maximum is attained by some complex vector x_* of unit length. Suppose y is a unit vector where that $y \perp x_*$ and let λ be a complex number. Since A is self-adjoint we have:

$$(2) \qquad \langle A(x_* + \lambda y), x_* + \lambda y \rangle = m^* + 2 \cdot \mathfrak{Re}\big(\lambda \cdot \langle Ax_*, y \rangle\big) + |\lambda|^2 \cdot \langle Ay, y \rangle$$

Now $x + \lambda y$ has norm $\sqrt{1 + \lambda|^2}$ and the maximality gives:

$$(3) \qquad m^* + 2 \cdot \mathfrak{Re}\big(\lambda \cdot \langle Ax_*, y \rangle\big) + |\lambda|^2 \cdot \langle Ay, y \rangle \leq \sqrt{1 + |\lambda|^2} \cdot m^*$$

Suppose now that $\langle Ax_*, y \rangle \neq 0$ and set

$$\langle Ax_*, y \rangle = s \cdot e^{i\theta} \quad : \quad s > 0$$

With $\delta > 0$ we take $\lambda = \delta \cdot e^{-i\theta}$ and (3) entails that

$$(4) \qquad 2s \cdot \delta \leq (\sqrt{1 + \delta^2} - 1) \cdot m^* - \langle Ay, y \rangle \cdot \delta^2$$

Next, by calculus one has $2 \cdot \sqrt{1 + \delta^2} - 1) \leq \delta^2$ so after division with δ we get

$$(5) \qquad 2s \leq \delta \cdot \big(\frac{m_*}{2} - \langle Ay, y \rangle\big)$$

But this is impossible for arbitrary small δ and hence we have proved that

$$(6) \qquad y \perp x_* \implies \langle Ax_*, y \rangle = 0$$

This means that $x_*^\perp$ is an invariant subspace for A and the restricted operator remains self-adjoint. At this stage the reader can finish the proof to get a unitary matrix U such that (*) holds.

5. Normal operators.

An $n \times n$-matrix A is normal if it commutes with its adjoint, i.e. $A^* A = AA^*$ holds in $M_n(\mathbf{C})$.

5.1 Exercise. Let A be a normal matrix. Show that every equivalent matrix is normal, i.e. if S is invertible then SAS^{-1} is also normal. The hint is to use that

$$(S^{-1})^* = (S^*)^{-1}$$

holds for every invertible matrix. Conclude that one can refer to normal linear operators the hermitian vector space $\mathbf{C}^n$.

5.2 Exercise. Let A and B be two Hermitian matrices which commute, i.e. $AB = BA$. Show that the matrix $A + iB$ is normal.

Next, let R be normal and assume that the its characteristic polynomial has simple roots. This means that there exists a basis $\xi_1, \ldots, \xi_n$ formed by eigenvectors to R with eigenvalues $\lambda_1, \ldots, \lambda_n$. Thus:

$$(*) \qquad R(\xi_\nu) = \lambda_\nu \cdot \xi_\nu \quad : \quad 1 \leq \nu \leq n$$

Notice that R is invertible if and only if al the eigenvalues are $\neq 0$. It turns out that the normality gives a stronger conclusion.

5.3 Proposition. *Assume that the eigenvalues are* $\neq 0$. *Then the* ξ*-vectors in (*) are orthogonal.*

Proof. Consider some eigenvector, say ξ_1. Now we get

$$\text{(i)} \qquad R(R^*(\xi_1)) = R^*(R(\xi_1)) = \lambda_1 \cdot R^*(\xi_1)$$

Hence $R^*(\xi_1)$ is an eigenvector to R with eigenvalue λ_1. By hypothesis this eigenspace is 1-dimensional which gives

$$R^*(\xi_1) = \mu \cdot \xi_1 \implies$$

$$\lambda_1 \cdot \langle \xi_1, \xi_1 \rangle = \langle R(\xi_1), \xi_1 \rangle = \langle \xi_1), R^*(\xi_1) \rangle = \bar{\mu} \cdot \langle \xi_1, \xi_1 \rangle$$

Hence $\mu = \bar{\lambda}_1$ which shows that the eigenvalues of R^* are the complex conjugates of the eigenvalues of R. There remains to show that the ξ-vectors are orthogonal. Consider two eigenvectors, say ξ_1, ξ_2. Then we obtain:

$$\bar{\lambda}_2 \lambda_1 \cdot \langle \xi_1, \xi_2 \rangle = \langle R\xi_1, R\xi_2 \rangle = \langle \xi_1, R^* R\xi_2 \rangle \langle \xi_1, RR^* \xi_2 \rangle =$$

$$\text{(ii)} \qquad \langle R^* \xi_1, R^* \xi_2 \rangle = \bar{\lambda}_1 \cdot \lambda_2 \cdot \langle \xi_1, \xi_2 \rangle \implies (\bar{\lambda}_2 \lambda_1 - \lambda_2 \bar{\lambda}_1) \cdot \langle \xi_1, \xi_2 \rangle = 0$$

By assumption $\lambda_1 \neq \lambda_2$ and both are $\neq 0$. It follows that $\bar{\lambda}_2 \lambda_1 - \lambda_2 \bar{\lambda}_1 \neq 0$ and then (ii) gives $\langle \xi_1, \xi_2 \rangle = 0$ as required.

5.4 Remark. Proposition 5.3 shows that if R is an invertible normal operator with n distinct eigenvalues then there exists a unitary matrix U such that U^*RU is a diagonal matrix. But in contrast to the Hermitian case the eigenvalues can be complex.

5.5 Exercise. Let R be an invertible normal operator with distinct eigenvalues. Show that R is a Hermitian matrix if and only if the eigenvalues are real numbers.

5.6 A consequence. Let R as above be an invertible normal operator with distinct eigenvalues. From the above there exists a unitary matrix U so that U^*RU is a diagonal matrix. Let $\{a_k + ib_k\}$ be the diagonal elements where the a- and the b-mumners are real. Now we get the diagonalmatrix A with $a_{kk} = a_k$ and the diagonal matrix B with $b_{kk} = b_k$. it follows that

$$U^*RU = A + iB$$

5.7 Exercise. Conclude from the above that if R is an invertible normal operator woith distinct eigenvalues then there exists *a unique pair of pairwise commuting Hermitian operators* A, B such that

$$R = A + iB$$

5.8 The operator R^*R. Let R be an invertible normal operator. Here we do not assume that the charactersitic polynomial $P_R(\lambda)$ has simple roots. Since R commutes with R^* the reader can check that R^*R is a Hermitian operator whose eigenvalues are all real and positive.

5.9 The normal operator $(A + iE_n)^{-1}$. Let A be a arbitrary Hermitian $n \times n$-matrix. We have already seen that its eigenvalues are real and denote them by $r_1, \ldots, r_n$. It follows that the matrix $A + iE_n$ is invertible. The reader should check that the inverse

$$R = (A + iE_N)^{-1}$$

is a normal operator with eigenvalues $\{\frac{1}{r_\nu + i}\}$.

4.10 The case of multiple roots The assumption that the eigenvalues of a normal operator are all distinct can be relaxed. Thus, for every normal and invertible operator R there exists a unitary operator U such that U^*RU is diagonal.

5.11 Exercise. Prove the assertion above. The hint is to show that if R is normal and nilpotent, i.e.$R^k = 0$ for some $k \geq 1$, then R must be the zero operator. To prove this one employs Jordan's theorem and the reader should verify that normality of an operator prevents Jordan blocks of size ≥ 2. For example, with I$n = 2$ we take

$$S = \begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix} \Longrightarrow S^* = \begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix}$$

and the reader can observe that $S^*S \neq SS^*$, i.e. the Jordan matrix S is not normal.

5. Fundamental solutions to ODE:s.

Recall from Calculus that every ordinary differential equation can be expressed as a system of first order equations. The fundamental issue is therefore to consider a matrix valued function $A(t)$, i.e. an $n \times n$-matrix whose elements $\{a_{ik}(t)\}$ are functions of t. Given $A(t)$ there exists at least locally close to $t = 0$, a unique $n \times n$-matrix $\Phi(t)$ such that

$$\frac{d\Phi}{dt} = A(t) \cdot \Phi(t)$$

with the initial condition $\Phi(0) = E_n$. One refers to Φ as a fundamental solution. The columns of the Φ-matrix give solutions to the homogenous system defined by $A(t)$. Moreover, the determinant of $\Phi(t)$ is $\neq 0$ for every t. In fact his follows from the equality (*) below:

Exercise. The trace function of A is defined by:

$$\mathrm{Tr}(A)(t) = a_{11}(t) + \ldots + a_{nn}(t)$$

Show that the function $t \mapsto \det(\,\Phi(t))$ satisfies the ODE.equation

$$\frac{d}{dt}(\det \Phi(t)) = \det \Phi(t) \cdot \mathrm{Tr}(A)(t)$$

Hence we have the formula

$$(*) \qquad \det \Phi(t) = e^{\int_0^t \mathrm{Tr}(A)(s) \cdot ds} \quad : t \geq 0$$

For example, if the trace function is identically zero then $\det \Phi(t) = 1$ for all t.

5.1 Inhomogeneous equations. From (*) it follows that the matrix $\Phi(t)$ is invertible for all t. This gives a formula to solve a inhomogeneous equation:

$$(1) \qquad \frac{d\mathbf{x}}{dt} = A(t)(\mathbf{x}(t)) + \mathbf{u}(t)$$

Here $\mathbf{u}(t) = (u_1(t), \ldots, u_n(t)$ is a given vector-valued function and one seeks a vector-valued function $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t)$ such that (1) holds and in addition satisfies the initial condition:

$$(2) \qquad \mathbf{x}(0) = \mathbf{b} \quad \text{where } \mathbf{b} \text{ is some vector}$$

Exercise. Show that the unique solution to (1) is given by

$$(**) \qquad \mathbf{x}(t) = \Phi(t)(\mathbf{b}) + \Phi(t)\big(\int_0^t \Phi^{-1}(s)\big(\mathbf{u}(s)\big) \cdot ds\big)$$

In other words, for every t we first evaluate the matrix $\Phi(t)$ on the n-vector $\mathbf{b}$ which gives the first time dependent vector in the right hand side. In the second term the inverse matrix $\Phi^{-1}(s)$ is applied to $\mathbf{u}(s)$ for every $0 \leq s \leq t$. After integration over $[0, t]$ we get a time-dependent n-vector on which $\Phi(t)$ is applied.

6. Carleman's inequality

Introduction Theorem 6.1 below was proved 1917 by Carleman in the article *Sur le genre du denominateur* $D(\lambda)$ *de Fredholm* from 1917. At that time the result was used to study non-singular integral equations of the Fredholm type. For more recent applications of Theorem 6.1 we refer to Chapter XI in [Dunford-Schwartz]. The Hilbert-Schmidt norm, respectively the operator norm were defined in § xx.

6.1 Theorem. *Let A be an $n\times n$-matrix and $\lambda_1,\ldots,\lambda_n$ are the roots of $P_A(\lambda)$. For every $\lambda\neq 0$ is outside $\sigma(A)$ one has the inequality:*

$$|\prod_{i=1}^{i=n}[1-\frac{\lambda_i}{\lambda}]e^{\lambda_i/\lambda}|\cdot||R_A(\lambda)||\leq|\lambda|\cdot\exp(\frac{1}{2}+\frac{\mathbf{HS}(A)^2}{2\cdot|\lambda|^2})$$

The proof requires several steps.First we shall need some preliminsries about traces of matrices.

6.2 Traceless matrices. Let A be an $n\times n$-matrix. The trace is given by:

(i) $$\mathrm{Tr}(A)=b_{11}+\ldots+b_{nn}$$

Recall that $-\mathrm{Tr}(A)$ is equal to the sum of the roots of $P_A(\lambda)$. In particular the trace of two equivalent matrices are equal. This will be used to prove the following:

6.3 Proposition. *Let A be an $n\times n$-matrix whose trace is zero. Then there exists a unitary matrix U such that the diagonal elements of U^*AU all are zero.*

Proof. Consider first consider the case $n=2$. By Theorem 4.0.7 it suffices to consider the case when the 2×2-matrix A is upper diagonal and since the trace is zero it has the form

$$A=\begin{pmatrix}a & b\\ 0 & -a\end{pmatrix}$$

where a,b is a pair of complex numbers. If $a=0$ then the two diagonal elements are zero and wee can take $U=E_2$ to be the identity in Lemma 6.5. If $a\neq 0$ we consider a vector $\phi=(1,z)$ in $\mathbf{C}^2$. Then $A(\phi)$ is the vector $(a+bz,-az)$ and hence the inner product becomes:

(i) $$\langle A(\phi),\phi\rangle=a+bz-a|z|^2$$

We can write

$$\frac{b}{a}=re^{i\theta}$$

where $r>0$ and then (i) is zero if

(ii) $$|z|^2=1+se^{i\theta}\cdot z$$

With $z=se^{-i\theta}$ it amounts to find a positive real number s such that $s^2=1+s$ which clearly exists. Now we get the vector

$$\phi_*=\frac{1}{1+s^2}(1,se^{-i\theta})$$

which has unit length and

(ii) $$\langle A(\phi_*),\phi_*\rangle=0$$

By 4.0.6 we find another unit vector ψ_* so that ϕ_*,ψ_* is an orthonormal base in $\mathbf{C}^2$ and hence there exists a unitary matrix U such that $U(e_1)=\phi_*$ and $U(e_2)=\psi_*$. If $B=U^*AB$ the vanishing in (ii) gives $b_{11}=0$. At the same time the trace is unchanged, i.e. $\mathrm{tr}(B)=0$ holds and hence we also get $b_{22}=0$. This means that the diagonal elements of U^*AU are both zero as required.

The case $n\geq 3$. For the induction the following is needed:

Sublemma. *Let $n\geq 3$ and assume as above that $Tr(A)=0$. Then there exists some non-zero vector $\phi\in\mathbf{C}^n$ such that*

(*) $$\langle A(\phi),\phi\rangle=0$$

Proof. If (*) does not hold we get the positive number

$$m_* = \min_\phi \left|\langle A(\phi), \phi\rangle\right|$$

where the minimum is taken over unit vectors in $\mathbf{C}^n$. The minimum is achieved by some unit vector ϕ_*. Let $\phi_*^\perp$ be its orthonormal complement and E the self-adjoint projection from $\mathbf{C}^n$ onto $\phi_*^\perp$. On the $(n-1)$-dimensional inner product space $\phi_*^\perp$ we get the linear operator $B = EA$, i.e.

$$\text{(i)} \qquad B(\xi) = E(A(\xi)) \quad : \quad \xi \in \phi_*^\perp$$

If $\psi_1, \ldots, \psi_{n-1}$ is an orthonormal basis in $\phi_*^\perp$ then the n-tuple $\phi_*, \psi_1, \ldots, \psi_{n-1}$ is an orthonormal basis in $\mathbf{C}^n$ and since the trace of A is zero we get

$$\text{(ii)} \qquad 0 = \langle A(\phi_*), \phi_*\rangle + \sum_{\nu=1}^{\nu=n-1} \langle A(\psi_\nu), \psi_\nu\rangle = m + \sum_{\nu=1}^{\nu=n-1} \langle B(\psi_\nu), \psi_\nu\rangle$$

where we used that $E(\psi_\nu) = \psi_\nu$ for each ν and that E is self-adjoint so that

$$\langle A(\psi_\nu), \psi_\nu\rangle = \langle A(\psi_\nu), E(\psi_\nu)\rangle = \langle E(A(\psi_\nu)), \psi_\nu\rangle = \langle B(\psi_\nu), \psi_\nu\rangle$$

Now (ii) gives

$$\mathrm{Tr}(B) = -m$$

Hence the $(n-1) \times (n-1)$-matrix which represents $B + \frac{m}{n-1} \cdot E$ has trace zero. By an induction over n we find a unit vector $\psi \in \phi_*^\perp$ such that

$$\langle B(\psi_*), \psi_*\rangle = -\frac{m}{n-1}$$

Finally, since E is self-adjoint we have already seen that

$$\langle A(\psi_*), \psi_*\rangle = \langle B(\psi_*), \psi_*\rangle \implies \left|\langle A(\psi_*), \psi_*\rangle\right| = \left|\frac{m}{n-1}\right| = \frac{m_*}{n-1}$$

Since $n \geq 3$ the last number is $< m_*$ which contradicts the minimal choice of m_*. Hence we must have $m_* = 0$ which proves the sublemma.

Final part of the proof. Let $n \geq 3$. The Sublemma gives unit vector ϕ such that $\langle A(\phi), \phi\rangle = 0$. Consider the hyperplane $\phi^\perp$ and the operator B from the Sublemma which now has trace zero on this $(n-1)$-dimensional space. So by an induction over n there exists an orthonormal basis $\psi_1, \ldots, \psi_{n-1}$ in $\phi^\perp$ such that $\langle B(\psi_\nu), \psi_n u\rangle = 0$ for every ν. Now $\phi, \psi_1, \ldots, \psi_{n-1}$ is an orthonormal basis in $\mathbf{C}^n$ and if U is the unitary matrix which has this n-tuple as column vectors it follows that the diagonal elements of U^*AU all vanish. This finishes the proof of Proposition 6.3.

Proof Theorem 6.1

.

Set $B=\lambda^{-1}A$ so that $\sigma(B)=\{\lambda_i/\lambda\}$ and $\text{Tr}(B)=\sum\frac{\lambda_i}{\lambda}$. We also have

$$\mathbf{HS}(B)^2=\frac{\mathbf{HS}(A)^2||}{|\lambda|^2}\quad \& \quad \big|\lambda\big|\cdot||R_A(\lambda)||=||(E-B)^{-1}||$$

Hence Theorem 6.1 follows if we prove the inequality

$$(*)\qquad \big|e^{\text{Tr}(B)}\big|\cdot\big|\prod_{i=1}^{i=n}\big[1-\frac{\lambda}{\lambda_i}\big]\cdot||(E-B)^{-1}||\leq \exp\big[\frac{1+\mathbf{HS}(B)^2}{2}\big]$$

To prove (*) we choose an arbitrary integer N such that $N>\big|\text{Tr}(B)\big|$ and for each such N we define the linear operator B_N on the $n+N$-dimensional complex space with points denoted by (x,y) with $y\in\mathbf{C}^N$ as follows:

$$(**)\qquad B_N(x,y)=(Bx\,,\,-\frac{\text{Tr}(B)}{N}\cdot y)$$

The eigenvalues of the linear operator $E-B_N$ is the union of the n-tuple $\{1-\frac{\lambda_i}{\lambda}\}$ and the N-tuple of equal eigenvalues given by $1+\frac{\text{Tr}(B)}{N}$. This gives the determinant formula

$$(1)\qquad \det(E-B_N)=\big(1+\frac{\text{Tr}(B)}{N}\big)^N\cdot\prod_{i=1}^{i=n}(1-\frac{\lambda_i}{\lambda})$$

The choice of N implies that (1) is $\neq 0$ so the inverse $(E-B_N)^{-1}$ exists. Moreover, the construction of B_N gives for any pair (x,y) in $\mathbf{C}^{N+n}$:

$$(E-B_N)^{-1}(x,y)=\big(E-B)^{-1}(x),\frac{y}{1+\frac{1}{N}\cdot\text{Tr}(B)}\big)\implies$$

$$||(E-B)^{-1}||\leq||(E-B_N)^{-1}||$$

Multiply both soide with $\det(E-B_N)$ which gives

$$(2)\qquad \big|\det(E-B_N)\big|\cdot||(E-B)^{-1}||\leq\big|\det(E-B_N)\big|\cdot||(E-B_N)^{-1}||$$

Hadarmard's inequality from § xx majorises the right hand side in (2) by:

$$(3)\qquad \frac{\mathbf{HS}(E-B_N)^{N+n-1}}{(N+n-1)^{N+n-1)/2}}$$

Next, the construction of B_N implies that its trace is zero. So by the result in 6.3 we can find an orthonormal basis $\xi_1,\ldots,\xi_{n+N}$ in $\mathbf{C}^{n+N}$ such that

$$\langle B_N(\xi_k),\xi_k\rangle=0\quad :1\leq k\leq n+N$$

Relative to this basis the matrix of $E-B_N$ has 1 along the diagonal and the negative of the elements of B_N elsewhere. It follows that the Hilbert-Schmidt norm satisfies the equality:

$$(4)\qquad \mathbf{HS}(E-B_N)^2=N+n+\mathbf{HS}(B_N)^2=N+n+\mathbf{HS}(B)^2+N^{-1}\cdot|\text{Tr}(B)|^2$$

To simplify notations we set $\mathfrak{B}=\mathbf{HS}(B)$ Hence, (1) and the inequalities from (2-3) give:

$$\big(1+\frac{\text{Tr}(B)}{N}\big)^N\cdot\prod_{i=1}^{i=n}(1-\frac{\lambda_i}{\lambda})\cdot||(E-B)^{-1}||\leq$$

$$\frac{\big(N+n+\mathfrak{B}^2+N^{-1}\cdot|\text{Tr}(B)|^2\big)^{(N+n-1)(2}}{\big(N+n-1\big)^{N+n-1/2}}=\frac{\big(1+\frac{\mathfrak{B}^2}{N+n}+\frac{|\text{Tr}(B)|^2}{N(N+n)}\big)^{(N+n-1)/2}}{(1-\frac{1}{N+n}\big)^{N+n-1/2}}$$

This inequality holds for arbitrary large N. Passing to the limit as $N \to \infty$ the definition of Neper's constant e shpws that the last term above converges to $\exp\left[\frac{1+\mathfrak{B}^2}{2}\right]$ which gives (*) above.

0.C.2 Hadamard's inequality.

The following result is due Hadamard whose proof is left as an exercise.

0.C.3 Theorem. *Let $A = \{a_{\nu k}\}$ be some $p \times p$-matrix whose elements are complex numbers. To each $1 \leq k \leq p$ we set*

$$\ell_p = \sqrt{|a_{1k}|^2 + \ldots + |a_{pk}|^2}$$

Then

$$\big|\det(A)\big| \leq \ell_1 \cdots \ell_p$$

7. Hadamard's radius theorem.

Hadamard's thesis *Essais sur l'études des fonctions donnés par leur dévelopment d Taylor* contains many interesting results. Here we expose material from Section 2 in [ibid]. Consider a power series

$$f(z) = \sum c_n z^n$$

whose radius is a positive number ρ. So f is analytic in the open disc $\{|z| < \rho\}$ and has at least one singular point on the circle $\{|z| = \rho\}$. Hadamard found a condition in order that these singularities consists of a finite set of poles only so that f extends to be meromorphic in some disc $\{|z| < \rho_*\}$ with $\rho_* > \rho$. The condition is expressed via properties of the Hankel determinants $\{\mathcal{D}_n^{(p)}\}$ from § 0.B. For each $p \geq 1$ we set

$$\delta(p) = \limsup_{n\to\infty} [\mathcal{D}_n^{(p)}]^{\frac{1}{n}}$$

In the special case $p = 0$ we have $\{\mathcal{D}_n^{(0)}\} = \{c_n\}$ and hence

$$\delta(0) = \frac{1}{\rho} = \limsup_{n\to\infty} |c_n|^{\frac{1}{n}}$$

This entails that for every $\epsilon > 0$ there exists a constant C_ϵ such that

$$|c_n| \leq C \cdot (\rho - \epsilon)^{-n} \quad \text{hold for every} \quad n$$

It follows trivially that

$$|\mathcal{D}_n^{(p)}| \leq (p+1)! \cdot C^{p+1}(\rho - \epsilon)^{-(p+1)n}$$

Passing to limes superior where high n:th roots are taken we conclude that:

$$\delta(p) = \limsup_{n\to\infty} \left[\mathcal{D}_n^{(p)}\right]^{\frac{1}{n}} \leq \rho^{-(p+1)} \tag{1}$$

Suppose there exists some $p \geq 1$ where a strict inequality occurs:

$$\delta(p) < \rho^{-(p+1)} \tag{2}$$

Let p be the smallest integer ≥ 1 where the strict inequality holds. This gives a number $\rho_* > \rho$ such that

$$\delta(p) = \rho_*^{-1} \cdot \rho^{-p} \tag{3}$$

7.1 Theorem. *With p chosen minimal as above, it follows that $f(z)$ extends to a meromorphic function in the disc of radius ρ_* where the number of poles counted with multiplicity is at most p.*

The proof requires several steps. To begin with one has

7.2 Lemma. *When p as above is minimal one has the unrestricted limit formula:*

$$\lim_{n\to\infty} \left[\mathcal{D}_n^{(p-1)}\right]^{\frac{1}{n}} = \rho^{-p} \tag{*}$$

TO BE GIVEN: Exercise power series+ Sylvesters equation.

7.3 The meromorphic extension to $\{|z| < \rho_*\}$. Lemma 7.2 entails that if n is large $\{\mathcal{D}_n^{(p-1)}\}$ are $\neq 0$. So there exists some n_* such that every $n \geq n_*$ gives a unique p-vector $(A_n^{(1)}, \ldots, A_n^{(p)})$ which solves the inhomogeneous system

$$\sum_{k=0}^{k=p-1} c_{n+k+j} \cdot A_n^{(p-k)} = -c_{n+p+j} \quad : \quad 0 \leq j \leq p-1$$

Or expressed in matrix notation:

$$(*)\qquad \begin{pmatrix} c_n & c_{n+1} & \ldots & c_{n+p-1}\\ c_{n+1} & c_{n+2} & \ldots & c_{n+p}\\ \ldots & \ldots & \ldots & \ldots\\ c_{n+p-1} & c_{n+p} & \ldots & c_{n+2p-2} \end{pmatrix} \begin{pmatrix} A_n^{(p)}\\ \ldots\\ \ldots\\ \ldots\\ A_n^{(1)} \end{pmatrix} = - \begin{pmatrix} c_{n+p}\\ \ldots\\ \ldots\\ \ldots\\ c_{n+2p-1} \end{pmatrix}$$

7.4 Exercise. Put

$$H_n = c_{n+2p} + A_n^{(1)} \cdot c_{n+2p-1} + \ldots + A_n^{[(p)} \cdot c_{n+p}$$

Show that the evaluation of $\mathcal{D}_n^{(p)}$ via an expansion of the last column gives the equality:

$$\text{(i)}\qquad H_n = \frac{\mathcal{D}_n^{(p)}}{\mathcal{D}_n^{(p-1)}}$$

Next, the limit formula (3) above Theorem 7.1 together with Lemma 7.2 give for every $\epsilon > 0$ a constant C_ϵ such that the following hold for all sufficiently large n:

$$\text{(ii)}\qquad |H_n| \leq C_\epsilon \cdot \big(\frac{\rho+\epsilon}{\rho_* - \epsilon}\big)^n$$

Next, put

$$\text{(iii)}\qquad \delta_n^k = A_{n+1}^{(k)} - A_n^{(k)} \quad : \quad 1 \leq k \leq p$$

Solving (*) above for n and $n+1$ a computation shows that the δ-numbers satisfy the system

$$\sum_{k=0}^{k=p-1} c_{n+j+k+1} \cdot \delta_n^{(p-k)} = 0 \quad : \quad 0 \leq j \leq p-2$$

$$\text{(iv)}\qquad \sum_{k=0}^{k=p-1} c_{n+p+k} \cdot \delta_n^{(p-k)} = -(c_{n+2p} + A_n^{(1)} \cdot c_{n+2p-1} + \ldots + A_n^{[(p)} \cdot c_{n+p})$$

The δ-numbers in the linear system (iv) are found via Cramer's rule. The minors of degree $p-1$ in the Hankel matrices $\mathcal{C}_{n+1}^{(p-1)}$ have elements from the given c-sequence and (7.0) implies that every such minor has an absolute value majorized by

$$C \cdot (\rho - \epsilon)^{-(p-1)n}$$

where C is a constant which is independent of n. We conclude that the δ-numbers satisfy

$$\text{(v)}\qquad |\delta_n^{(k)}| \leq |\mathcal{D}_n^{(p-1)}|^{-1} \cdot C \cdot (\rho-\epsilon)^{-(p-1)n} \cdot |H_n|$$

The unrestricted limit in Lemma 7.2 give upper bounds for $|\mathcal{D}_n^{(p-1)}|^{-1}$ so that (iii) and (v) give:

7.5 Lemma *To each $\epsilon > 0$ there is a constant C_ϵ such that*

$$|\delta_n^{(k)}| \leq C_\epsilon \cdot \big(\frac{\rho+\epsilon}{\rho_* - \epsilon}\big)^n \quad : \quad 1 \leq k \leq p$$

7.6 The polynomial $Q(z)$. Lemma 7.5 and (iii) entail that the sequence $\{A_n^{(k)} : n = 1, 2, \ldots\}$ converges for every k and we set

$$A_*^{(k)} = \lim_{n\to\infty} A_n^{(k)}$$

Notice that Lemma 7.5 after summations of geometric series gives a constant C_1 such that

$$\text{(7.6.i)}\qquad |A_*^{(k)} - A_n^{(k)}| \leq C_1 \cdot \big(\frac{\rho+\epsilon}{\rho_* - \epsilon}\big)^n$$

hold for every $1 \leq k \leq p$ and every n.

Now we consider the sequence

$$b_n = c_{n+p} + A_*^{(1)} \cdot c_{n+p-1} + \ldots A_*^{(p)} \cdot c_n \tag{7.6.ii}$$

Equation (*) applied to $j = 0$ gives

$$b_n = (A_*^{(1)} - A_n^{(1)}) \cdot c_{n+p-1} + \ldots + (A_*^{(p)} - A_n^{(p)}) \cdot c_n \tag{7.6.iii}$$

Next, we have already seen that $|c_n| \leq C \cdot (\rho - \epsilon)^{-n}$ hold for some constant C which together with (7.6.i) gives:

7.7 Lemma. *For every $\epsilon > 0$ there exists a constant C such that*

$$|b_n| \leq C \cdot \big(\frac{1+\epsilon}{\rho_*}\big)^n$$

Finally, consider the polynomial

$$Q(z) = 1 + A_*^{(1)} \cdot z + \ldots A_*^{(p)} \cdot z^p$$

Set $g(z) = Q(z)f(z)$ which has a power series $\sum d_\nu z^\nu$ where

$$b_n = c_n \cdot A_*^{(p)} + \ldots c_{n+p-1} A_*^{(1)} + c_{n+p} = d_{n+p}$$

Above p is fixed so Lemma 7.7 and the trivial spectral radius formula show that $g(z)$ is analytic in the disc $|z| < \rho_*$. This proves that f extends and the poles are contained in the zeros of the polynomial Q which occur in the annulus $\rho \leq |z| < \rho_*$.

8. On positive definite quadratic forms

In many situations one is asking when a given a bi-linear form is positive definite. We prove a result which has a geometric interpretation. Let $m \geq 2$ and denote m-vectors in $\mathbf{R}^m$ with capital letters, i.e. $X = (x_1, \ldots, x_m)$. Let $N \geq 2$ be some positive integer and $X_1, \ldots, X_N$ an N-tuple of real m-vectors. To each pair $j \neq k$ we set

$$b_{ij} = ||X_j|| + X_k|| - ||X_j - X_k||$$

where $||\cdot||$ is the usual euclidian length in $\mathbf{R}^m$. We get the symmetric $N \times N$-matrix with elements $\{b_{ij}\}$ and the associated quadratic form

$$H(\xi_1, \ldots, \xi_N) = \sum \sum b_{ij} \cdot \xi_i \cdot \xi_j$$

8.1 Theorem. *If the X-vectors are all different then H is positive definite.*

The proof relies upon a useful formula to express the length of a vector in $\mathbf{R}^m$.

8.2 Lemma There exists a constant C_m such that for every m-vector X one has

(*)
$$||X|| = C_m \cdot \int_{\mathbf{R}^m} \frac{1 - \cos\langle X, Y\rangle}{||Y||^{m+1}} \cdot dY$$

Proof. We use polar coordinates and denote by dA the area measure on the unit sphere S^{m-1} and $\omega = (\omega_1, \ldots, \omega_m)$ denote points on the unit sphere S^{m-1}. Notice that the integrals

$$\int_{S^{m-1}} (1 - \cos\langle X, \omega\rangle) \cdot dA$$

only depend upon $||X||$. Hence it suffices to prove Lemma 8.2 when $X = (R, \ldots, 0)$ where $R = ||X||$ and here the integral in (*) becomes:

$$\int_0^\infty \Big[\int_{S^{m-2}} (1 - \cos Rr\omega_1) \cdot dA_{m-1} \Big] \cdot \frac{dr}{r^2}$$

where dA_{m-1} is the area measure on S^{m-2}. Set

$$B(R, \omega_1) = \int_0^\infty (1 - \cos Rr\omega_1) \cdot \frac{dr}{r^2}$$

for each $-1 < \omega_1 < 1$. The variable substitution $r \to s/R$ gives

$$B(R, \omega_1) = R \cdot \int_0^\infty \frac{1 - \cos s\omega_1}{s^2} \cdot ds = R \cdot B_*(\omega_1)$$

With these notations the integral in (*) becomes

(1)
$$R \cdot \int_{S^{m-2}} B_*(\omega_1) \cdot dA_{m-2}$$

Hence Lemma 8.2 follows where C_m^{-1} is equal to (1) above.

Proof of Theorem 8.1. For a given pair i, j the addition formula for the cosine-function gives:

$$1 - \cos\langle X_i, Y\rangle + 1 - \cos\langle X_j, Y\rangle + \cos\langle (X_i - X_j), Y\rangle =$$

(1)
$$(1 - \cos\langle X_i, Y\rangle) \cdot (1 - \cos\langle X_i, Y\rangle) + \sin\langle X_i, Y\rangle \cdot \sin\langle X_j, Y\rangle$$

It follows that the matrix element b_{ij} is given by

$$C_m \cdot \int_{\mathbf{R}^m} \frac{(1 - \cos\langle X_i, Y\rangle) \cdot (1 - \cos\langle X_j, Y\rangle) + \sin\langle X_i, Y\rangle \cdot \sin\langle X_j, Y\rangle}{||Y||^{m+1}} \cdot dY$$

From this we see that

$$H(\xi) = C_m \cdot \int_{\mathbf{R}^m} \big([\sum (\xi_k \cdot (1-\cos\langle X_k, Y\rangle)\,]^2 + [\sum (\xi_k \cdot (\sin\langle X_k, Y\rangle)\,]^2\big) \cdot \frac{dY}{||Y||^{m+1}}$$

This shows that H is positive definite as requested.

8.3 Exercise. Prove more generally that for every $1 < p < 2$ a similar result as above holds when the elements of the matrix are:

$$b_{ij} = ||X_j||^p + X_k||^p - ||X_j - X_k||^p$$

Hint. Employ a similar formula as in (*) where a new constant $C_{p,m}$ appears and $||Y||^{m+1}$ is replaced by $||Y||^{m+p}$.

8.4 A class of Hermitian matrices. *Let $z_1, \ldots, z_N$ be an n-tuple of distinct and non-zero complex numbers. Set*

$$b_{ij} = \{\frac{z_i}{z_j}\}$$

Then the matrix $B = \{b_{ij}\}$ is Hermitian and positive definite.

Again the proof is left as an exercise to the reader.

8.5 Remark. Theorem 8.1 has several applications. For example, Beurling used it to prove the existence of certain spectral measures which arise in ergodic processes. Another application from [Beurling: Notes Uppsala 1935] goes as follows: Let f and g be a pair of continuous and absolutely integrable functions on the real line. Define the function on the real t-line by

$$\phi(t) = \int_{-\infty}^{\infty} [f(t+s) - g(s)| \cdot ds$$

8.6 Theorem. *There exists a measure μ on the ξ-line of total variation $\leq 2\sqrt{||f||_1 \cdot ||g||_1}$ such that*

$$\phi(t) = ||f||_1 + ||g||_1 + \int_{-\infty}^{\infty} e^{i\xi t} \cdot d\mu(\xi)$$

The reader is invited to try to prove this theorem using Theorem 8.1 and the observation that the a similar result as above holds for L^2-functions f and g, i.e. this time we set

$$\psi(t) = \int_{-\infty}^{\infty} [f(t+s) - g(s)|^2 \cdot ds$$

and one shows that there exists a measure γ whose total variation is $\leq 2\sqrt{||f||_2 \cdot ||g||_2}$ and

$$\psi(t) = ||f||_2 + ||g||_2 + \int_{-\infty}^{\infty} e^{i\xi t} \cdot d\gamma(\xi)$$

9. The Davies-Simon inequality.

Introduction. Every $n \times n$-matrix A can be regared as a $\mathbf{C}$-linear operator on the hermitian complex n-space which yields the operator norm $\text{Norm}(A)$. Just as in Theorem 6.1 we shall exhibit an inequality for the operator norm but this time another feature appears. Namely, Theorem 9.1 yields an upper bound expressed by the euclidian distance from λ to $\sigma(A)$ which is better than the product which appears in the left hand side of Theorem 6.1. On the other hand, the inequality below is restricted to special λ-values whose absolute values are larger than the operator norm of A. Hence the results in 6.1 and 9.1 supplement each other.

9.1 Theorem. *For every $n \times n$-matrix A whose operator norm is ≤ 1 the inequality below holds for every $0 \leq \theta \leq 2\pi$ outside $\sigma(A)$*

$$\text{Norm}(R_A(e^{i\theta})) \leq \cot \frac{\pi}{4n} \cdot \text{dist}(e^{i\theta}, \sigma(A))^{-1}$$

Proof. Schur's result in Theorem 4.0.7 reduces the proof to the case when A is upper triangular and replacing A by $e^{i\theta}A$ we may take $\theta = 0$. Set $B = (E - A)^{-1}$ and let B^* be the adjoint operator. The equations $B - BA = E$ and $A^*B^* - B^* = -E$ give

$$B(E - AA^*)B^* = BB^* - (B - E)A^*B^* = BB^* - (B - E)(B^* - E) = B + B^* - E$$

Set $C = B + B^* - E$ and notice that the diagonal elements

(1) $$c_{kk} = \frac{1}{1 - \lambda_k i} + \frac{1}{1 - \bar{\lambda}_k i} - 1 = \frac{1 - |\lambda_k|^2}{|1 - \lambda_k|^2}$$

where $\{\lambda_k\}$ are the diagonal elements of A which give points in $\sigma(A)$. Now we shall we prove the inequality:

(2) $$|b_{ij}|^2 \leq \frac{(1 - |\lambda_i|^2) \cdot (1 - |\lambda_j|^2)}{(1 - \lambda_i|^2 \cdot |1 - \lambda_j|^2}$$

To get (2) we consider a vector x and obtain

(3) $$\langle Cx, x \rangle = \langle B(E - AA^*)B^*x, x \rangle = \langle (E - AA^*)B^*x, B^*x \rangle \geq 0$$

where the last equality holds since the self-adjoint matrix $E - AA^*$ is non-negative because A by assumption has operator norm ≤ 1. From (3) and the Cauchy-Schwarz inequality applied to the symmetric matrix we get

(4) $$|c_{ij}|^2 \leq |c_{ii}| \cdot |c_{jj}| \quad : \quad i < j$$

for each pair $i \neq j$. Since $c_{ij} = b_{ij}$ when $i < j$ we get (2). Next, put $\delta = \text{dist}(1, \sigma(A))$ which means that $|1 - \lambda_i| \geq \delta$ for every i. From this it is clear that (2) and the triangle inequality give

(5) $$|b_{ij}|^2 \leq \frac{4}{\delta^2} \quad : \quad i < j$$

At the same time the diagonal elements satisfy:

(6) $$|b_{ii}|^2 = \frac{1}{|1 - \lambda_i|^2} \leq \frac{1}{\delta^2}$$

Let T be the upper triangular matrix where $t_{ij} = 2$ when $i < j$ and $t_{ii} = 1$ for each i. Then the elements in $\frac{1}{\delta} \cdot T$ majorize the absolute values of the B-matrix. The observation from § xx implies that

$$\text{Norm}(B) \leq \frac{1}{\delta} \cdot \text{Norm}(T)$$

Now Theorem 9.1 follows from the formula in § xx for the operator norm of T.

10. An equality by Schur.

Let A be an $n \times n$-matrix with operator norm ≤ 1, i.e., A is a contraction. For each polynomisl $p(z) = a_0 + a_1 z + \ldots + a_N z^N$ with complex coefficients we get the matrix $p(A)$.

10.1 Theorem. *One has the inequality*

$$||p(A)|| \leq \max_{z \in D} |, |p(z)|$$

To prove this we first consider an analytic function $g(z)$ in the unit disc which extends continuously to the boundary and an $n \times n$-matrix A whose spectrum is contained in the open unit disc. Here we do not assume that A is a contraction, i.e. the sole assumption is that $\sigma(A)$ is a compact subset of the open unit disc. Now $g(z)$ has a series expansion $\sum c_k z^k$ we know from § xx that the matrix-valued series $\sum c_k A^k$ converges and gives a matrix $g(A)$. Hence there also exists the exponential matrix

$$B = e^{g(A)}$$

The adjoint B^* is found as follows. Consider the analytic function $g^*(z) = \sum \overline{c_k} z^k$. The reader can check that

$$B^* = e^{g^*(A^*)}$$

Put

$$C = e^{g^*(A^*) + g(A)} = B^* B$$

The result in § xx gives

10.2 The Schur-Weierstrass inequality. *For each pair A and g as above one has*

$$(*) \qquad ||e^{g(A)}|| = \max_{\lambda \in \sigma(A)} e^{\mathfrak{Re}(g(\lambda))}$$

10.3. Another norm inequality. Let α be a point in the open unit disc and suppose now that A is a contraction. It follows that

$$\text{(i)} \qquad (1 - |\alpha|^2) \cdot \langle (E - A^* A)(y), y \rangle \geq 0$$

hold for every vector y. The reader can check that (i) gives

$$\text{(ii)} \qquad ||\alpha E - A)(y)||^2 \leq ||y - \bar{\alpha} \cdot A(y)||^2$$

Next, there exists the matrix

$$g_\alpha(A) = (\alpha E - A) \cdot (E - \bar{\alpha} A)^{-1}$$

Consider a vector x of unit norm and set

$$y = (E - \bar{\alpha} A)^{-1}(x)$$

Then

$$\text{(iii)} \qquad ||g_\alpha(A)(x)||^2 = ||\alpha E - A)(y)||^2 \leq ||y - \bar{\alpha} \cdot A(y)||^2$$

where the last inequality used (ii). Now $y - \bar{\alpha} \cdot A(y) = x$ and hence (iii) gives

$$||g_\alpha(A)(x)||^2|| \leq ||x||^2$$

Since the vector x was arbitrary we conclude that

$$\text{(iv)} \qquad ||g_\alpha(A)|| \leq 1$$

Proof of Theorem 10.1.

By scaling we can assume that the maximum norm $|p|_D = 1$. Construct the Blaschke product $B(z9$ taken over the zeros of p in the open unit disc which gives a factorisation

$$p(z) = B(z) \cdot e^{g(z)}$$

where the zero-free analytic function $e^{g(z)}$ has maximum norm one, and hence

$$\mathfrak{Re}(g)(z) \leq 0 \quad : z \in D$$

Now

$$p(A) = B(A) \cdot e^{g(A)}$$

Here $B(A)$ is the product of operators of the form $-g_\alpha(A)$ where α are zeros of p in D. By (iv) from (10.3) each of these operators have norm ≤ 1 and (*) in (*) in (10.2) entails that the same holds for $e^{g(A)}$. So $p(A)$ is the product of operators of norm ≤ 1 and Theorem 10.1 follows.

Remark. The interested reader should consult the text-book [Davies] for further extensions of Theorem 10.1 which appear in [ibid: Chapter 10].

11. An application to integral equations.

Let $k(x,y)$ be a complex-valued continuous function on the unit square $\{0 \leq x, y \leq 1\}$. We do not assume that k is symmetric, i.e, in general $k(x,y) \neq k(y,x)$. Let $f(x)$ be another continuous-function on $[0,1]$. Assume that the maximum norms of k and f both are < 1. By induction over n starting with $f_0(x) = f(x)$ we get a sequence $\{f_n\}$ where

$$f_n(x) = \int_0^1 k(x,y) \cdot f_{n-1}(y) \cdot dy \quad : \quad n \geq 1$$

The hypothesis entails that each f_n has maximum norm < 1 and hence there exists a power series:

$$u_\lambda(x) = \sum_{n=0}^{\infty} f_n(x) \cdot \lambda^n$$

which converges for every $|\lambda| < 1$ and yields a continuous function $u_\lambda(x)$ on $[0,1]$.

11.1 Theorem. *The function $\lambda \mapsto u_\lambda(x)$ with values in the Banach space $B = C^0[0,1]$ extends to a meromorphic B-valued function in the whole λ-plane.*

To prove this we introduce the recursive Hankel determinants for each $0 \leq x \leq 1$:

$$\mathcal{D}_n^{(p)}(x) = \det \begin{pmatrix} f_{n+1}(x) & f_{n+2}(x) & \dots & \dots & f_{n+p}(x) \\ f_{n+2}(x) & f_{n+3}(x) & \dots & \dots & f_{n+p+1}(x) \\ \dots & \dots & \dots & \dots & \\ \dots & \dots & \dots & \dots & \\ f_{n+p}(x) & f_{n+p+1}(x) & \dots & \dots & f_{n+2p-1}(x) \end{pmatrix}$$

Proposition 11.2 *For every $p \geq 2$ and $0 \leq x \leq 1$ one has the inequality*

$$\big|\, \mathcal{D}_n^{(p)}(x)) \,\big| \leq (p\,!)^{-n} \cdot \big(p^{\frac{p}{2}}\big)^n \cdot \frac{p^p}{p\,!}$$

11.3 Conclusion. The inequality above entails that

$$\limsup_{n \to \infty} \big|\mathcal{D}_n^{(p)}(x))\,\big|^{1/n} \leq \frac{p^{p/2}}{p\,!}$$

Next, Stirling's formula gives:

$$\lim_{p \to \infty} \big[\frac{p^{1/2}}{p\,!}\,\big]^{-1/p} = 0$$

Hence Hadamard's theorem gives Theorem 11.1

Proof of Proposition 11.2

The proof requires several steps. First, we get the sequence $\{k^{(m)}(x)\}$ which starts with $k = k^{(1)}$ and:

$$k^{(m)}(x) = \int_0^1 k^{(m-1)}(x,s)\ddot{k}(s) \cdot ds \quad : \quad m \geq 2$$

It is easily seen that

$$f_{n+m}(x) = \int_0^1 k^{m)}(x,s) \cdot f_n(s) \cdot ds$$

hold for all pairs $m \geq 1$ and $n \geq 0$.

11.4 Determinant formulas. Let $\phi_1(x), \dots, \phi_p(x)$ and $\psi_x), \dots, \psi_p(x)$ be a pair of p-tuples of continuous functions on $[0,1]$. For each point $(x_1, \dots, x_p)$ in $[0,1]^p$ we put

$$D_{\phi_1,\ldots,\phi_p}(x_1,\ldots,x_p) = \det\begin{pmatrix} \phi_1(x_1) & \cdots & \phi_1(x_p) \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ \phi_p(x_1) & \cdots & \phi_1(x_p) \end{pmatrix} \quad :$$

In the same way we define $D_{\psi_1,\ldots,\psi_p}(x_1,\ldots,x_p)$. Next, define the $p \times p$-matrix with elements

$$a_{jk} = \int_0^1 \phi_j(s) \cdot \psi_k(s)\, ds$$

11.5 Lemma. *One has the equality*

$$\det(a_{jk}) = \frac{1}{p\,!} \int_{[0,1]^p} \Phi(s_1,\ldots,s_p) \cdot \Psi(s_1,\ldots,s_p) \cdot ds_1 \cdots ds_p$$

11.6 Exercise. Prove this result using standard formulas for determinants.

Next, for each $0 \leq x \leq 1$ and every pair n, p of positive integers we consider the $p \times p$-matrix

$$\begin{pmatrix} \int_0^1 k(x,s) f_n(s)\, ds & \int_0^1 k^{(2)}(x,s) f_n(s)\, ds & \ldots & \ldots & \int_0^1 k^{(p)}(x,s) f_n(s) \\ \int_0^1 k^{(2)}(x,s) f_{n+1}(s)\, ds & \int_0^1 k^{(2)}(x,s) f_{n+1}(s)\, ds & \ldots & \ldots & \int_0^1 k^{(2)}(x,s) f_{n+1}(s) \\ \ldots & \ldots & \ldots & \ldots & \\ \ldots & \ldots & \ldots & \ldots & \\ \int_0^1 k^{(p)}(x,s) f_{n+p-1}(s)\, ds & \int_0^1 k^{(p)}(x,s) f_{n+p-1}(s)\, ds & \ldots & \ldots & \int_0^1 k^{(p)}(x,s) f_{n+p-1}(s) \end{pmatrix} \quad :$$

We also get the two determinant functions

$$\mathcal{K}^{(p)}(x, s_1,\ldots,s_p) = \det\begin{pmatrix} k^{(1)}(x,s_1) & k^{(1)}(x,s_2) & \ldots & \ldots & k^{(1)}(x,s_p) \\ k^{(2)}(x,s_1) & k^{(2)}(x,s_2) & \ldots & \ldots & k^{(2)}(x,s_p) \\ \ldots & \ldots & \ldots & \ldots & \\ \ldots & \ldots & \ldots & \ldots & \\ k^{(p)}(x,s_1) & k^{(p)}(x,s_2) & \ldots & \ldots & k^{(p)}(x,s_p) \end{pmatrix}$$

$$\mathcal{F}_n^{(p)}(s_1,\ldots,s_p) = \det\begin{pmatrix} f_n(s_1) & f_n(s_2) & \ldots & \ldots & f_n(s_p) \\ f_{n+1}(s_1) & f_{n+1}(s_2) & \ldots & \ldots & f_{n+1}(s_p) \\ \ldots & \ldots & \ldots & \ldots & \\ \ldots & \ldots & \ldots & \ldots & \\ f_{n+p-1}(s_1) & f_{n+p-1}(s_2) & \ldots & \ldots & f_{n+p-1}(s_p) \end{pmatrix}$$

11.7 Lemma. Let $\mathcal{D}_n^{(p)}(x)$ denote the determinant of the matrix (x). Then one has the equation

$$\mathcal{D}_n^{(p)}(x) = \frac{1}{p!} \cdot \int_{[0,1]^p} \mathcal{K}^{(p)}(x, s_1 \ldots, s_p) \cdot \mathcal{F}_n^{(p)}(s_1,\ldots,s_p)\, ds_1 \cdots ds_p$$

PROOF: Apply previous lemma ....

Next, using (xx) we have the equality

Exercise. Use the formulas above to conclude that the requested intequality in Proposition 11.2 holds.