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. An important result is the spectral theorem for symmetric $n \times n$-matrices with real elements, and its counterpart for complex Hermitian matrices which was alreay discovere around 1810 by Cauchy in the case when eigenvalues are distinct. Recall that eigenvalues are found by regarding maxima and minima of associated quadratic forms. Weierstrass' collected work contains a wealth of results related to the spectral theorem for hermitian matrices and their interplay with quadratic forms. Here is an elementary result from Weierstrass' studies: Let $N \geq 2$ and $\{c_p q\} : 1 \leq p, q \leq N$ is 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 it follows that

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

for every N-tuple $\{x_p\}$ of non-negative real numbers. The reader is invited to supply a proof via the spectral theorem for symmetric matrices.

Using Lagrange's interpolation formula Sylvester exhibited extensive classes of matrix-valued functions by residue calculus and further results were achieved by Frobenius who treated the general case when a characteristic polynomial of a matrix has multiple roots. The usefulness of matrices and their determinants in analysis 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. Carleman's inequality for norms of resolvents in § 6 serves as a veritable cornerstone for the spectral theory of linear operators since it extends to the infinite dimensional case. See for example Chapter XX in [Dunford-Schwartz].

Hadamard's theorem. A high-light in this section appears in Theorem 7.1 whose proof relies heavily upon calculus with determinants. Therefore we shall give a rather extensive account of these. Of special relevance is the construction of Hankel determinants in § B and the Gram-Fredholm formula in § C. The less experienced reader may prefer to begin with the material in § 2-5 which is of a "soft-ware nature" compared to the more involved results which rely upon computations of determinants in § 1.

Integral operators and their entire functions. A result which motivates the material in this section in regard to functional analysis goes as follows: 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 $k(x,y) \neq k(y,x)$. If $f(x)$ is a continuous function on $[0,1]$ we get a sequence $\{f_n\}$ where $f_0(x) = f(x)$ and

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

If $M = \max_x |f(x)|$ is the maximum norm of f and $C = ||k||_\infty$ it is clear that the maximum norms $|f_n|_\infty \leq (CM)^n$. Hence there exists a power series:

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

which converges for every $|\lambda| < 1/CM$ and yields an analytic function With these notations the following result was established by Carleman in the article *xxx.*

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

Remark. So in particular the series

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

yields and entire function of λ for each f and every $0 \leq x \leq 1$. The proof is given in § xx and relies upon properties of the Hankel determinants

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

defined for all pairs $p, n \geq 1$ and each continuous function f. The crucial result is the following a priori inequality:

Proposition. *If the maximum norms of k and f both are ≤ 1 then one has the inequality*

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

for every $p \geq 2$ and $0 \leq x \leq 1$.

§ 1. Determinants.

Introduction. The material in this section is inspired by the contents in the excellent text-book *Determinantenheorie* from 1909 by Gerhard Kovalevski. We give a rather detailed account about some classic results since determinants are not always treated in detail in contemporary text-books devoted to matrices and more "abstract linear algebra".

A.0 The Hilbert-Schmidt norm

Let A be a matrix whose elements $\{a_{pq}\}$ are complex numbers. Its Hilbert-Schmidt norm is defined by

$$||A|| = \sqrt{\sum\sum |a_{pq}|^2}$$

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

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

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

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

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

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

This gives in particular the inequality

$$(3) \qquad \operatorname{Norm}(A) \leq ||A||$$

Example. Given a matrix A we replace each element by its absolute value $|a_{pq}|$ and get a matrix A_* whose elements are non-negative real numbers where

$$(i) \qquad \operatorname{Norm}(A_*) = \max_{x_1,\ldots x_n} \sqrt{\sum_{p=1}^{p=n} |\sum_{q=1}^{q=n} |a_{pq}| x_q|^2}$$

Above it suffices to compete with real n-vectors for which $\sum x_p^2 = 1$ and every $x_p \geq 0$. The triangle inequality gives $\operatorname{Norm}(A) \leq \operatorname{Norm}(A_*)$. The A_*-norm is found via Lagrange's multiplier i.e. one employs Lagrange's criterion for extremals of quadratic forms which entails that (i) is maximized by a real non-negative n-vector x which satsfies a linear system of equations

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

Introducing the double indexed numbers

$$\beta_{jq} = \sum_{p=1}^{p=n} a_{pj} a_{pq}$$

Lagrange's equations corresponds to the system

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

In the "generic case" the $n \times n$-matrix $\{\beta_{jq}\}$ is non-singular, i.e. its determinant is $\neq 0$ and (iii) has a unique solution x^* with every $x_j^* \geq 0$ for a uniquely determined multiplier $\lambda > 0$. As an example we consider an $n \times n$-matrix of the form

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

Thus, the diagonal elements are all units and T is upper triangular with $t_{ij} = s$ for pair $i < j$ while the elements below the diagonal are zero. In spite of the explicit expression for T the computation of its operator norm is a bit involved. The case $s = 2$ is of special interest and here one has the classic formula

(*) $$\text{Norm}(T_2) = \cot \frac{\pi}{4n}$$

Exercise. Prove (*) and find the x^*-vector which maximizes (iii), If necessary, consult the literature and let us remark that one can use numerical experminents with a computer to settle (*).

A.1 The Sylvester-Franke theorem.

Let A be some $n \times n$-matrix with elements $\{a_{ik}\}$. Put

$$b_{rs} = a_{11} a_{rs} - a_{r1} a_{1s} \quad : \quad 2 \leq r, s \leq n$$

These b-numbers give an $(n-1) \times (n-1)$-matrix where b_{22} is put in position $(1,1)$ and so on. The matrix is denoted by $\mathcal{S}^1(A)$ and called the first order Sylvester matrix. If $a_{11} \neq 0$ one has the equality

(A.1.1) $$a_{11}^{n-2} \cdot \det(A) = \det(\mathcal{S}^1(A))$$

Exercise. Prove this result or consult a text-book which apart from "soft abstract notions" does not ignore to treat determinants.

A.1.2 Sylvester's equation. For every $1 \leq h \leq n-1$ 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:

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

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

A.1.3 The Sylvester-Franke theorem. Let $n \geq 2$ and $A = \{a_{ik}\}$ an $n \times n$-matrix. Let $m < n$ and consider the family of minors of size m, i.e. one picks m columns and m rows which give an $m \times m$-matrix whose determinant is called a minor of size m of the given matrix A. The total number of such minors is equal to

$$N^2 \quad \text{where} \quad N = \binom{n}{m}$$

We have N many strictly increasing sequences $1 \leq \gamma_1 < \dots \gamma_m \leq n$ where a γ-sequence corresponds to preserved columns when a minor is constructed. Similarly we have 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 similarly s corresponds to the enumerated sequence of rows. Now we obtain the $N \times N$-matrix

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

We refer to $\mathcal{A}_m$ as the Franke-Sylvester matrix of order m. They are defined for each $1 \leq m \leq n-1$.

A.1.4 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 the proof of Theorem 0.A.1 the reader can consult [Kovalevski: page102-105].

§ B. Hankel determinants.

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

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

Let $\mathcal{D}_n^{(p)}$ denote the determinant. One refers to $\{\mathcal{D}_n^{(p)}\}$ as the recursive Hankel determinants. They are used to establish various 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

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

We say that $\{c_n\}$ has finite rank if there exists a number r^* such that r^* many ξ-vectors are linearly independent and the rest are linear combinations of these.

B.1 Rational series expansions. The sequence $\{c_n\}$ gives the formal power series

$$f(x) = \sum_{\nu=0}^{\infty} c_\nu x^\nu \tag{B.1.1}$$

If $n \geq 1$ we set

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

It is clear that $\{c_\nu\}$ has finite rank if and only if the sequence $\{\phi_\nu(x)\}$ 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 $(a_0, \dots, a_p)$ of complex numbers such that the power series

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

Multiplying this equation with x^p it follows that

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

where $q(x)$ is a polynomial. Hence the finite rank entails that the power series (B.1.1) represents a rational function.

Exercise. Conversely, assume that

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

for some pair of polynomials. Show that $\{c_n\}$ has finite rank. The next result is also left as an exercise to the reader.

B.2 Proposition. *A sequence $\{c_n\}$ has a finite rank if and only if there exists an integer p such that*

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

Moreover, one has the equality p is equal to the rank of $\{c_n\}$.

B.3 A specific example. Suppose that the degree of q is strictly less than that of g in the Exercise above and that the rational function $\frac{q}{g}$ is expressed by a sum of simple fractions:

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

where $\alpha_1, \dots, \alpha_p$ are distinct and every $d_k \neq 0$. Then we see that

$$c_n = \sum_{k=1}^{k=p} d_k \cdot \alpha_k^n \quad \text{where we have put} \quad \alpha_k^0 = 1 \quad \text{so that} \quad c_0 = \sum d_k$$

B.4 The reduced rank. Assume that $\{c_n\}$ has a finite rank r^*. 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 and we find a non-negative integer r_* such that $r_k = r_*$ for large k and refer to r_* as the reduced rank. By the construction $r_* \leq r^*$. The relation between r^* and r_* 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 discussion any further but refer to the literature. See in particular the exercises in [Polya-Szegö : Chapter VII:problems 17-34].

B.5 Hankel's formula for Laurent series. Consider a rational function of the form

$$R(z) = \frac{q(z)}{z^p - [a_1 z^{p-1} + \ldots + a_{p-1} z + a_p]}$$

where the polynomial q has degree $\leq p-1$. At ∞ we have a Laurent series expansion

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

Consider the $p \times p$-matrix

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

B.5.1 Theorem. *Let $\mathcal{D}_n^{(p)}$ be the Hankel determinants of $\{c_n\}$. Then the following hold for every $n \geq 1$:*

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

Exercise. Prove this result.

B.6 The Hadamard-Kronecker identity. For all pairs of positive integers p and n one has the equality:

$$\mathcal{D}_n^{(p+1)} \cdot \mathcal{D}_{n+2}^{(p-2)} = \mathcal{D}_n^{(p+1)} \mathcal{D}_{n+2}^{(p-1)} - \big[\mathcal{D}_{n+1}^{(p)}\big]^2 \tag{5.6.1}$$

Proof. The equality (5.6.1) is s special case of a determinant formula for symmetric matrices which is due to Sylvester. Namely, let $N \geq 2$ and consider a symmetric matrix

$$S = \begin{pmatrix} s_{11} & s_{12} & \ldots & s_{1N} \\ s_{21} & s_{22} & \ldots & s_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ s_{N1} & a_{N2} & \ldots & s_{NN} \end{pmatrix}$$

Now we consider the $(N-1) \times (N-1)$-matrices

$$S_1 = \begin{pmatrix} s_{22} & s_{23} & \ldots & s_{2N} \\ s_{32} & s_{33} & \ldots & s_{3N} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ s_{N2} & s_{N3} & \ldots & s_{NN} \end{pmatrix} \quad : \quad S_2 = \begin{pmatrix} s_{12} & s_{13} & \ldots & s_{1N} \\ s_{22} & s_{23} & \ldots & s_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ s_{N-1,2} & s_{N-1,3} & \ldots & s_{N-1,N} \end{pmatrix}$$

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

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

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

B.7 Sylvester's identity. *One has the equation*

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

Exercise. Prove this result and deduce the Hadamard-Kronecker equation.

§ C. 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 $[0,1]^p$:

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

Product rules for determinants give the Gram-Fredholm equation

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

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

§ D. 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 no symmetry condition is imposed. 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 less regular kernel functions, the original determinants used by Fredholm were modified by Hilbert which avoid the singularities and lead to quite general formulas for resolvents of the integral operator $\mathcal{K}$ defined by

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

One studies foremost the case when K is square integrable, i.e. when

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

In this case one refers to a Hilbert-Schmidt operator. An eigenvalue 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 $\{\lambda_n\}$. In the article [Schur: 1909] Schur proved the inequality

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

Notice that one does not assume that the kernel function is symmetric, i.e. in general $K(x,y) \neq K(y,x)$.

D.1 Hilbert's determinants. Let K be a kernel function for which (*) is finite where K is singular on the diagonal subset of $\Omega \times \Omega$ but continuous in its complement. 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:

(i) $$D_m^*(\xi, \eta) = \int_{\Omega^m} C_m^*(\xi, \eta : x_1, \ldots, x_m) \cdot dx_1 \cdots dx_m$$

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

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

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

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

§ E. Results by Carleman

Using the 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.

(*) $$\iint |K(x,y)|^2 \, dxdy < \infty$$

From § D we have the linear operator $\mathcal{K}$ on $L^2(\Omega$. When λ is outside $\sigma(\mathcal{K})$ we get the resolvent operator

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

It is represented by a kernel function $\Gamma(\xi, \eta; \lambda)$, i.e.

$$R(\lambda)(\phi)(x) = \int_{\Omega} \Gamma(\xi, \eta; \lambda) \cdot \phi(\xi) \, d\xi$$

E.1 Theorem. *Outside the spectrum one has the equation*

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

E.2 Remark. This equation is due to Fredholm and Hilbert and a detailed proof appears in Hilbert's text-book *Integralgleichungen.* The reader may try to discover a proof via the Gram-Fredholm formula. Next, since $\mathcal{K}$ is of Hiblert-Schmidt type it is a compact operator on $L^(\Omega)$ wuth a discrete spectrum $\{\lambda_\nu\}$ where multiple eigenvalues are repeated when the corresponding eigenspaces have dimension ≥ 2. Theorem E.1 shows that the spectrum correspond to zeros of the entire function $\mathcal{D}(\lambda)$.

E.3 Exercise. Apply inequalities of the Fredholm-Hadamard type for determinants to show that

$$\int_\Omega \Gamma(\xi,\xi;\lambda)\cdot d\xi = -\lambda\cdot\sum_{\nu=1}^\infty \frac{1}{\lambda_\nu(\lambda-\lambda_\nu)} \tag{E.3}$$

Now we announce a major result from Carleman's cited article.

E.4 Theorem. *$\mathcal{D}(\lambda)$ is an entire function of the complex parameter λ given by a Hadamard product*

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

where $\{\lambda_n\}$ satisfy

$$\sum |\lambda_n|^{-2} < \infty \tag{2}$$

Remark. The crucial step in Carleman's proof is based upon 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:

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

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

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

and finally

$$N = \sum_{j=1}^{j=q}\sum_{\nu=1}^{\nu=q} |a_{p+j,p+\nu}|^2$$

E.5 Theorem. *Let D be the determinant of the matrix (*). Then*

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

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

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

The absolute value of the last determinant is majorized by Hadamard's inequality in § F.XX and the requested inequality in Theorem E.5 follows.

2. Resolvent matrices

Let A be some matrix $n \times n$-matrix. Its characteristic polynomial is defined by

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

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:

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

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

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

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

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

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

2.1 Exercise. Prove the equality

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

Thes local series expansions show 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:

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

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:

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

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

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

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

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

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

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

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

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

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

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

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

(***) $$Q(A) = 0$$

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

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

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.

Before we enter Jordan's Theorem we discuss some consequences of the material in the previous section. The Cayley-Hamilton decomposition from 2.7. shows that an arbitrary $n \times n$-matrix A has a similar matrix $B = S^{-1}AS$ which is represented in a block form. More precisely, to every root α_ν of some multiplicity $e(\alpha_\nu)$ there occurs a square matrix B_ν of size $e(\alpha_\nu)$ 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 description of these nilpotent S-matrices which therefore yields a refinement of the Cayley-Hamilton decomposition.

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 and replacing A by $A - \alpha$ there remains only to consider the nilpotent case, i.e when $P_A(\lambda) = \lambda^n$ so that $A^n = 0$. In this nilpotent case we must find a basis where A is represented in Jordan's normal form. This is done below.

3.2 The case of nilpotent operators. Let S be a nilpotent **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

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

The unique integer m is denoted by $\text{ord}(S, v)$. The case $m = 1$ occurs if $S(v) = 0$. If v has order $m \geq 2$ one verifies 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 $\mathcal{C}(v)$ and called a *cyclic* subspace of V. When $S(v) =$ the corresponding cyclic space $\mathcal{C}(v)$ is reduced to $\mathbf{C} \cdot v$. With these notations 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

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

Choose $v^* \in V$ such that $\text{ord}(S, v^*) = m^*$ and construct the quotient space $W = \frac{V}{\mathcal{C}(v^*)}$ on which S induces a linear operator denoted 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

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

To each α we have the integer $k_\alpha = \text{ord}(\bar{S}, \bar{v}_\alpha)$. The construction of a quotient space means that

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

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

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

Next, put

$$k_\alpha^* = \text{ord}(S, v_\alpha) \tag{4}$$

It is obvious that $k_\alpha^* \geq k_\alpha$. If this inequality is strict we use (3) and obtain

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

The maximal choice of m^* entails that $k_\alpha^* \leq m^*$ which gives

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

Hence (3) enable us to find $w_\alpha \in \mathcal{C}(v^*)$ such that

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

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

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

Moreover, the construction of the ξ-vectors entail that

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

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

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

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 using an implemented program which for example can be found in *Mathematica*.

4. Hermitian and Normal 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. 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

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

4.0.1 Adjoint operators. Let A be a linear operator. Its adjoint A^* is the linear operator for which

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

4.0.2 Exercise. Show that if $e_1, \ldots, e_n$ is an arbitrary orthonormal basis in the inner product space $\mathbf{C}^n$ where A is represented by a matrix with elements $\{a_{p,q}\}$, then A^* is represented by the matrix whose elements are

$$a^*_{pq} = \bar{a}_{qp}$$

4.0.3 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.

4.0.4 Self-adjoint projections. Let V be a subspace of $\mathbf{C}^n$ of some dimension $1 \le k \le n-1$. 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.

4.0.5 Exercise. Show that if E is some $n\times n$-matrix which is idempotent in $M_n(\mathbf{C})$ and Hermitian in the sense of 4.0.3 then E is the self-adjoint projection attached to the subspace $V = E(\mathbf{C}^n)$.

4.0.6 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

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

4.1 The spectral theorem.

This important result asserts the following:

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.

4.2 Normal operators.

An $n \times n$-matrix A is normal if it commutes with its adjoint, i.e.

$$(*) \qquad A^* A = A A^* \quad \text{holds in} \quad M_n(\mathbf{C})$$

4.2.0 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 from this that we can refer to normal linear operators on $\mathbf{C}^n$.

4.2.1 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.

4.3 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

(i) $$R(R^*(\xi_1)) = R^*(R(\xi_1)) = \lambda_1 \cdot R^*(\xi_1)$$

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:

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

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

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.

4.4 Remark. Proposition 4.3 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.

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

4.6 Theorem. *Let R be an invertible normal operator with distinct eigenvalues. Then there exists a unique pair of Hermitian operators A, B such that $AB = BA$ and*

$$R = A + iB$$

4.7 Exercise. Prove Theorem 4.6.

4.8 The operator R^*R. Let R as above be an invertible normal operator with eigenvalues $\lambda_1, \ldots, \lambda_n$. From Remark 4.4 it is clear that R^*R is a Hermitian operator whose eigenvalues all are given by the positive numbers $\{|\lambda_\nu|^2\}$ and if $A.B$ are the Hermitian operators in Theorem 4.6 then we have

$$R^*R = A^2 + B^2$$

Thus, R^*R is represented as a sum of squares of two pairwise commuting Hermitian operators.

4.9 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 . Let us denote them by $r_1, \ldots, r_n$. The spectral theorem gives a unitary matrix U such that U^*AU is diagonal with elements $\{r_\nu\}$. It follows that the matrix $A + iE_n$ is invertible and its 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.

4.11 Exercise. Prove the assertion above. The hint is to establish the following which has independent interest:

4.12 Proposition. *Let R be normal and nilpotent. Then $R = 0$*

Proof. By Jordan's Theorem it suffices to prove this when R is a single Jordan block represented by a special S-matrix whose elements below the diagonal, are 1 while all the other elements are zero. If $n = 2$ we have for example

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

The reader verifies that $S^*S \neq SS^*$ and a similar calculation gives Proposition 4.12 for every $n \geq 3$.

4.13 Remark. The result above means that if R is normal then there never appear Jordan blocks of size > 1 and hence there exists an invertible matrix S such that SRS^{-1} is diagonal.

6. Carleman's inequality

Introduction Theorem 6.1 below was proved 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. It is defined for an $n \times n$-matrix $A = \{a_{ik}\}$ by:

$$||A|| = \sqrt{\sum \sum |a_{ik}|^2}$$

where the double sum extends over all pairs $1 \leq i, k \leq n$. Notice that this norm is the same as

$$||A||^2 = \sum_{i=1}^{i=n} ||A(e_i)||^2$$

where $e_1, \ldots, e_n$ can be taken as an arbitrary orthogonal basis in $\mathbf{C}^n$. Next, for a linear operator S on $\mathbf{C}^n$ its *operator norm* is defined by

$$\text{Norm}[S] = \max_x ||S(x)|| \quad \text{with the maximum taken over unit vectors.}$$

6.1 Theorem. *Let $\lambda_1, \ldots, \lambda_n$ be the roots of $P_A(\lambda)$ and $\lambda \neq 0$ is outside $\sigma(A)$. Then one has the inequality:*

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

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*

$$\big|\det(A)\big| \cdot \text{Norm}(A^{-1}) \leq \frac{||A||^{n-1}}{(n-1)^{n-1)/2}}$$

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.

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

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

Recall that $-\text{Tr}(A)$ 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:

6.4 Theorem. *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

$$m_* = \min_{\phi} \big|\langle A(\phi), \phi\rangle\big|$$

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.

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

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

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

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

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

Since $n \geq 3$ the last number is $< m_*$ which contradicts the minimal choice of m_*. Hence we must have $m_* = 0$ which proves lemma 6.5

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 Theorem 6.4.

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

$$||B||^2=\frac{||A||^2||}{|\lambda|^2}\quad\text{and}\quad \big|\lambda\big|\cdot\text{Norm}[R_A(\lambda)]=\text{Norm}[(E-B)^{-1}]$$

Hence Theorem 6.1 follows if we prove the inequality

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

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 $\mathbf{C}^{N+n}$:

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

It follows that

$$\text{Norm}\big[(E-B)^{-1})\big]\leq\text{Norm}\big[(E-B_N)^{-1}\big]\implies$$

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

Hadarmard's inequality estimates the hand side in (2) by:

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

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:

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

Hence, (1) and the inequalities from (2-3) give:

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

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

This inequality holds for arbitrary large N. Passing to the limit as $N \to \infty$ the definition of Neper's constant e give

$$\lim_{N\to\infty} \big(1 + \frac{\mathrm{Tr}(B)}{N}\big)^N = e^{\mathrm{Tr}(B)}$$

and the reader may also verify that the limit of the last term above is equal to $\exp\big[\frac{1+||B||^2}{2}\big]$ which finishes the proof of (*) above and hence also of Theorem 6.1.

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 Hadarmard's special and very wellknowen formula for the radius of convergence of a powrer series gives the equality:

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

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:

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

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

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

If p is the smallest integer where (2) holds we get a number $\rho_* > \rho$ such that

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

7.1 Theorem. *With p and ρ_* as in (3), 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:*

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

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:

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

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

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

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

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

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 $\rho \leq |z| < \rho_*$.

10. 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 $[0,1]$.

10.1 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 ontridice 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 10.2 *For every $p \geq 2$ and $0 \leq x \leq 1$ one has the inequality*

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

10.3 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 10.1

Proof of Proposition 10.2

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 $n \geq 0$.

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

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

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

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

Lemma. *One has the equality*

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

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

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

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

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 10.2 holds.

Continued fractions and there limit discs.

Cnosioder a pair of sequence of real numbers $\{a_p\}$ and $\{b_p\}$ where each b_p is positive. For the a-sequence it can occur that some - for even all -$a_p = 0$ and no conditions are imposed upon their signs. For each postivre integer n we have the symmetric $n \times n$-matrix

$$A_n = xxxx$$

If μ is a non-real complex number there exists a unique n-vector $\phi_1(\mu), \ldots, \phi_n(\mu)$ siuch that

$$A_n(\Phi) = \mu \cdot \Phi + e_1$$

It means that the n-vector satisfies the system of inear equations

$$xxxx$$

Set

$$g_n(\mu) = \phi_1(\mu)$$

Form the system above we see thst

$$g_n(\mu) = \frac{P_n(\mu)}{Q_n(\mu)}$$

Where P and Q are polynomials without a common zero: bottom line 192 gives formula. and here Q_n has degree n, while that of P_n is $n-1$. By an induction over n one easily verifes the equations

$$xxxx = xxxx$$

Now we can regard the Lureiunt series expnasion at $\mu = \infty$ for each rational g-function. Since the degree of $Q_n(\mu) \cdot Q_{n+}(\mu) = 2n+1$ for each n, it follows from (xx) that the first $2n$ terms in the Laurent series expansion of g_{n+1} are equal to those of g_n. This gives an asympttic series

$$-\sum \frac{c_n}{\mu^{n+1}}$$

hos truncated series up to degree $2n$ conicides with that of $g_n(\mu)$. Next, for each n we find a postive measure ρ_N on the real t-line such that

$$g_n(\mu) = |int \frac{d\rho_n(t)}{t-\mu}$$

and from the above we get

$$c_\nu = \int t^\nu \cdot d\rho_n(t)$$

when $0 \leq \nu \leq 2n-1$. After this one can pass to subsequences and get solutions to the moment problem.

We get an associated quadratic form in an inifinite number of variables $x = (x_1, x_2, \ldots)$ defined by

$$J(x) = \sum a_p x_p +$$

Consider also the symmetric matrix

$$W = xxx = xxx$$

1. The ϕ-vectors. For each pair of complex numbers ϕ_1 and μ there exists a unique vector $\phi(\mu) = (\phi_1, \phi_2(\mu), \ldots$ such that

$$W - \mu \cdot E)(\phi(\mu)) = 0 \tag{1.1}$$

In fact, (1.1) holds if and only if

$$b_1\phi_2(\mu) = (a_1 - \mu)\phi_1 \quad : \quad b_p\phi_{p+1}(\mu) = (a_p - \mu)\phi_p(\mu) - b_{p-1}\phi_{p-1}(\mu) \colon \ p \geq 3$$

Since $\{b_p\}$ are postive this determines the solution uniquely and it depends on ϕ_1 in a linear fashion. So we can write

$$\phi_n(\mu) = \psi_n(\mu) \cdot \phi_1 \quad : n \geq 2$$

2. Exercise. Show that when μ varies then $\mu \mapsto \psi_n(\mu)$ is a polynomial of degree $n-1$ for every $n \geq 2$. Moreover, $\psi_1(\mu) = 1$ and

$$b_1\psi_2(\mu) = (a_1 - \mu)\psi_1(\mu) \quad : \quad b_n\psi_{n+1}(\mu) = (a_n - \mu)\psi_n(\mu) - b_{n-1}\psi_{n-1}(\mu) \colon \; n \geq 3$$

3. Exercise. Multiply each of the equations above by the complex conjugates $\overline{\psi_1}(\mu), \ldots, \overline{\psi_n}(\mu)\}$ and take the sum to show that

$$0 = \mathfrak{Im}(\mu) \cdot \sum_{p=1}^{p=n} |\psi_1(\mu)|^2 + b_n \cdot \mathfrak{Im}(\overline{\psi_n}(\mu) \cdot \psi_{n+1}(\mu)) \tag{3.1}$$

Suppose that $\mathfrak{Im}(\mu) > 0$ and show via the Cauchy-Schwarz inequslity that this gives

$$\frac{1}{b_n} \leq \frac{1}{2 \cdot \mathfrak{Im}(\mu)} \cdot |\psi_n(\mu)|^2 + |\psi_{n+1}(\mu)|^2$$

THiese inequalities give:

4. Proposition. *If the series* $\sum \psi_n(\mu)|^2 < \infty$ *for some non-real* μ, *then* $\sum \frac{1}{b_n} < \infty$.

5. A conclusion. Propostion 4 shows that the quadratic form $J(x)$ is of Case II in the sense of § xx, then we must have

$$\sum \frac{1}{b_n} < \infty$$

So conversely, if we have a divergent series

$$\sum \frac{1}{b_n} = +\infty$$

then $J(x)$ is of Case I.

6. The kernel function $K(x,y)$. It is defined as zero outside the positive quadrant in $|bfR^2$, and if n is a positive integer we have $K(x,y) = a_p$ in the square $\{p-1 \leq x, y \leq p\}$, while $K(x,y) = b_p$ in the two squares

$$\{p-1 \leq x, y \leq p\} \times \{p < y \leq p+1\} : \; \{p-1 \leq x, y \leq p\} \times \{p < y \leq p+1\}$$

Next, For each integer n we denote by $K_n(x,y)$ the kernel which is equal to K when both x and y are $\leq n$ and is otherwise zero. Let $f_*(x)$ be the function which is 1 if $0 < x < 1$ and zero when $x > 1$. Since K_n is symmetric the integral equation

$$\int K_n(x,y) \cdot G(y)\, dy = \mu \cdot G(x) + f_*(x) \tag{6.1}$$

has a unique solution for each non-real μ and the reader can check that the G-function is constant over integer interval and vanishes when $x > n$.

Excercise. Let g_p denote the constant value on the interval $(p-1, p)$ of the G-function above. Show that one has the system of equations

$$xxxxxpage133 \tag{6.2}$$

In particular g_1 is eqaual to the value of the finite continued fraction of degree n as explained in § xx. This clarifies the relation between expansions in continued fractions and the quadratic form $J(x)$ above.

7. Some rational functions. From (6.2) it is clear that g_1 as a function of μ is a rational function of the form

$$\frac{P_n(\mu)}{Q_n(\mu)}$$

where the polynomials have no common factor. Moreover, (xx) gives the system of equations

$$xxxxpage193 \tag{7.1}$$

Exercise. Show that (7.1) above and the equations for the ψ-polynomials in (§ xx) give the equations:

$$\psi_{n+1}(\mu) = \frac{Q_n(\mu)}{b_1 \cdots b_n} \quad : \ n \geq 1 \tag{7.2}$$

8. The Ψ-function. For each μ we get the function $x \mapsto \Psi(x;\mu)$ defined on $x \geq 0$ by

$$\Psi(x;\mu) = \psi_p(\mu) \quad : p-1 < x < p)$$

If $n \geq 1$ we get the truncated function $\Psi_n(x;\mu)$ which is equal to $\Psi(x;\mu)$ when $x < n$ and zero if $x > n$.

Exercise. Let $n \geq 1$ and suppose that $\psi_{n+1}(\mu) = 0$ for some complex number μ. Apply the equation (xx) and show that this entails that $\Psi_n(x;\mu)$ solves the equation

$$\int_0^n K_n(x,y) \cdot \Psi_n(x;\mu)\, dx = \mu \cdot \Psi_n(x;\mu) \tag{8.1}$$

9. Reality of zeros. Since the kernel function K_n is symmetric we cannot hve a non-trivial solution in (8.1) above unless μ is real. Hence we conclude that the zeros of all the ψ-polynomials are real.

10. Some Cauchy transforms.

Consider the finite continued fraction in (x) for each pair n and μ. Keeping n fixed we obtain a unique probability measure s on the real t-line such that

$$g_n^*(\mu) = \int \frac{d\rho_n(t)}{t-\mu}$$

Now $\{g_n^*(\mu)\}$ is a normal family of analytic functions in the upper half-plane $\mathfrak{Im}(\mu) > 0$. Montel's theorem entails that we can extract convergent subsequences. Every analytic function $g^*(\mu)$ obtained as a limit in the Frechet space $\mathcal{O}(U_+)$ from a subsequence $\{g_{n_1}^*(\mu), g_{n_2}^*(\mu)\ldots\}$ is called a limit function of the infinite continued fraction from (xx).

Exercise. Use weak-star convergence in the space of probability measures on the real t-line to prove that each limit function g^* above is the Cauchy transform of a unique probability measure ρ, i.e.

$$g^*(\mu) = \int \frac{d\rho^*(t)}{t-\mu}$$

Moreover, check that the weak-star convergence and the orthogonal equations in (xx) give

$$\int \psi_n(t) \cdot \psi_m(t)\, d\rho^*(t) = 0 \quad : n \neq m$$

On the range of limit functions.