2. Resolvents

Let $n \geq 2$ and A is some $n \times n$-matrix. Its characteristic polynomial is defined by

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

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:

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

One refers to $R(\lambda)$ as a resolvent of A. When λ is outside the specttum of A we notice that

$$\lambda \cdot R(\lambda) = (\lambda E_n - A) \cdot R(\lambda) + A \cdot R(\lambda) = E_n + A \cdot R(\lambda)$$

Since $\lambda \cdot R(\lambda) = R(\lambda) \cdot \lambda$ we conclude thst

$$A \cdot R(\lambda) = R(\lambda) \cdot A$$

Thus, the resolvent matrices commute with A.

Exercise 1. Use Cramér's rule to show thst

(1.1) $$R(\lambda) = \frac{1}{P(\lambda} \cdot [Q_0 + \lambda \cdot Q_1 + \cdot + \lambda^{n-1} \cdot Q_{n-1}]$$

where $\{Q_\nu\}$ is sn n-tuple of mstriced.

In psrticulsr the matrix-vslued function $R(\lambda)$ is snslytic in $\mathbf{C} \setminus \mathbf{C}$.

2. The Neumsnn series. Show thst if R is strictly lsrger than the absolute value of ectery spectrasl vslue of A, thrn

$$R(\lambda) = \sum -\nu = 1^\infty$$

3. Hamilton's equation Let $Q(\lambda)$ be s polynomial. Show via residue calculus that

(3.1) $$Q(A) =$$

and derive via (1.1) thst

(3.2) $$P(A) = 0$$

One refers to (3.2) as Hamilton's vanishing theorem.

4. Cayley-s formula. In (1.1) we introduced the Q-mstriced. Thry can be found via residue cslculus. Firdt, since $P(\lambda 9$ is s pllymimsl of drgree $n$whose higrthedt term is λn one has för large r:

$$\int_x xxx$$

Togehrter eith xxxx we grt

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

Nect, muyltiöying $R(\lambda < 9$ eith λ snd uysiung (xx¡9 we grt

$$A = Q_{n-1} + \int xxxx$$

and the reader can check that the last intgral is eqaul to $c_1 \cdot E - n$. ¡hence

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

we leave to find the redt.

5. The Cayley-Hamilton decomposition.

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

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

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

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 $\mathbf{C}^n$.

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.

2. Resolvents

Let $n \geq 2$ and A some matrix in $M_n(\mathbf{C})$. Its characteristic polynomial is defined by

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

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:

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

The wellknown construction of inverse matrices via Cramer's rule gives an n-tuple of matrices $\{Q_\nu\}$ such that

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

It turns out that $Q_{n-1} = E_n$ is the identity matrix and if $0 \leq j \leq n-2$ then

$$Q_j = q_j(A) \tag{0.4}$$

where $\{q_j\}$ are polynomials ls of degree $\leq n-j-1$. To prove this one regards the Neumann series

$$\frac{E_n}{\lambda} + \sum_{\nu=1}^{\infty} \lambda^{-\nu-1} \cdot A^\nu \tag{0.5}$$

which converges in an exterior disc $|\lambda| > R$, and as explained in § xx (0.5) is equal to $R_(\lambda)$ in this exterior disc. We can construct line integrals over circles $|\lambda| = w$ where w is strictly larger than the absolute value of every root of $P_A(\lambda)$. Then (0.5) gives

$$A^k = \frac{1}{2\pi i} \int_{|\lambda|=w} \lambda^k \cdot R_A(\lambda) \cdot d\lambda \quad : \quad k = 1, 2, \ldots \tag{0.6}$$

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

$$q(A) = \frac{1}{2\pi i} \int_{|\lambda|=w} q(\lambda) \cdot R_A(\lambda) \cdot d\lambda \tag{0.7}$$

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

$$E_n = \frac{1}{2\pi i} \cdot \int_{|\lambda|=w} R_A(\lambda) \cdot d\lambda \tag{0.8}$$

If e_ν is the multiplicity of $P_A(\lambda)$ at a zero $\alpha_n u$ and if $q(\lambda)$ a polynomial with a zero of order $\geq e_\nu)$ at every root, then the reader can check that

$$q(A) = 0 \tag{0.9}$$

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

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

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

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

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

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

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 than$n$. 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

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

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.

Chapter I: Basic material for complex analysis

Content

I:A Complex numbers, polynomials and Möbius transforms.

0. Introduction

1.A: The field of complex numbers

1.B: The fundamental theorem of algebra

1:C. Interpolation by polynomials

1:D Tchebysheff polynomials and transfinite diameters.

1:E Exercises

2. Möbius functions

3. The Laplace operator

4. Some complex mappings

5. The stereographic projection

Chapter I:B. Series

0. Introduction

1. Additive series

1:B Counting functions

2. Power series

2:B Radial limits

2:C A theorem by Landau

3. Product series

4. Blascke products

5. Estimates using the counting function

6. Convergence on the boundary

7. An example by Hardy

8. Convergence under substitution

9. The series $\sum (a_1 \cdots a_\nu)^{\frac{1}{\nu}}$

10. Thorin's convexity theorem.

11. Cesaro and Hölder limits

12. Power series and arithmetic means

13. Taylor series and quasi-analytic functions

I:C Complex vector spaces

0. Introduction

0.A The Sylvester-Franke theorem

0.B Hankel determinants

0.C Hadamard's theorem

0.D The Gram-Schur formula

0.E Hadamard's inequality

1. Wedderburn's Theorem
2. Resolvent marices
3. Jordan's theorem
4. Hermitian and normal operators
5. Fundamental solutions to ODE-equations
6. Carleman's inequality for resolvents
7. Hadamard's radius formula
8. Positive quadratic forms

D.I Zeros of polynomials

A: Preliminary results

B: The Laguerre operators A^n_ζ

C: Properties of Laguerre forms

D: Proof of Theorem 0.1

E: Proof of Theorem 0.2 and 0.1

F: Legendre polynomials

I:E. Fourier series

A: The kernels of Dini, Fejer and Jackson

B: Legendre polynomials

C. The space $\mathcal{T}_n$

D. Tchebysheff polynomials and transfinite diameters

E. The kernels of Dini, Fejer, Jackson and Gibbs phenomenon

F. Partial Fourier sums and convergence in the mean

G. Best approximation by trigonometric polynomials.

Introduction.

Background to study analytic functions are covered by subsections 1.A-1.B and 2-5 in § I:A together with subsections 1-4 in § I:B. Let us emphasise that the material in subsections 3-5 in § 1.A is instructive. To learn about geometric properties of Möbius transforms on the unit disc is not only exciting for its own sake but gives insight about the geometry when one studies complex analytic mappings. A high-light appears in § I:A.5 where Theorem 5.6 asserts that Möbius transformations preserve the hyperbolic δ-distance in the unit disc which to each pair z_1, z_2 in D assigns the distance

$$\delta(z_1, z_2) = |\frac{z_1 - z_2}{1 - \bar{z}_1 \cdot z_2}|$$

Matrices and their determinants. Complex vector spaces and linear operators expressed via matrices are treated in § C. The fundamental theorem of algebra gives complex roots of the characteristic polynomial $\det(\lambda \cdot E - n - A)$ when A is an $n \times n$-matrix whose elements are complex numbers. This is a fundamental result with far-reaching applications for the theory about linear operators on complex vector spaces. We have included a rather extensive discussion about determinants in § I. C: 0.A-0.E which which apart from results of independent interest are used to prove some important results in analytic function theory, especially Hadamard's theorem in in § I. C:7 which is describes the absolute values of poles outside the origin of a meromorphic function whose local Taylor series at $z = 0$ is given.

Remark about series. Subsections 1-5 in § I:B contain material whose proofs are easy to while § 6 treats more advanced results due to Hardy and Littlewood which lead to Tauberian Theorems. Here the proofs are involved but it is rewarding to pursue the details since the methods employed by Hardy and Littlewood can be extended to general context to study linear operators on Banach spaces. Let us mention one such result which appears in Ergodic Theory and goes as follows: Let X be a reflexive complex Banach space and $T\colon X \to X$ a bounded linear operator. To each $n \geq 1$ we define the averaging operator

$$A_n = \frac{E + T \ldots + T^{n-1}}{n}$$

where E is the identity map on X. Suppose there exists a constant M such that the operator norms $||A_n|| \leq M$ for all n. Then a vector $x \in X$ has an averaged limit vector $A_*(X)$ in the sense that

$$\lim_{n\to\infty} ||A_n(x) - A_*(x)|| = 0$$

if and only if the appearantly much weaker condition holds:

$$\lim_{n\to\infty} \frac{T^n x}{n} = 0$$

In § XX [Functional Appendix] we demonstrate this result using methods which are similar to those employed by Hardy and Littlewood. This illustrates that it pays to study "pure series" since here many techniques are appear which are applicable in a more general context.

Zeros of polynomials. They are studied in in § I:D. We have also included a special section about Fourier series since the interplay betwen these and analytic functions play an essential role in more advanced studies. In this first chapter we are content to expose basic facts about Fourier series but remark that some results are not so standard in text-books. See in particular § I:E. F where we prove a theorem due to Carleman which is seldom mentioned in the literature,

I:A Complex numbers

Content

0. Introduction

1.A: The field of complex numbers

1.B: The fundamental theorem of algebra

1:C. Interpolation by polynomials

1:D Tchebysheff polynomials and transfinite diameters.

1:E Exercises

2. Möbius functions

3. The Laplace operator

4. Some complex mappings

5. The stereographic projection

Introduction.

Bieberbach's text-book [Bi:1] starts with a historic account about the origin of complex numbers and reflections upon how they are introduced to the beginner. Here follows an excerpt:

Every school pupil who learns about complex numbers and how to compute with them follows the same path as mathematical science did in the past. One first becomes familiar with the new and unpleasant concepts which surround complex numbers and by the natural inertia of the human mind it is not obvious why one should learn about all formal rules from the start. It is only later that one learns about the usefulness of complex numbers which makes it possible to settle previously unsolved problems. The miracle is that many problems which are phrased in the real number system can be solved by a detour over complex numbers. Once such examples have been understood the strength and beauty of complex numbers becomes clear.

The active role of complex numbers in algebra and analysis appeared quite late in the history of mathematics. An explanation might be that *conceptual thinking* (Begriffliches Denken) was rather remote to most mathematicians until the end of 1700. Even in his thesis from 1799, Gauss still did not fully break to the traditions in using complex numbers. Not until 1831 "würde Volle Klarheit nachbeweisbar" in his mathematical work where the Gaussian plane gives a geometric description of complex numbers. Here one must also give credit to the Norwegian mathematician Caspar Wessel who already in 1799 presented a work at the Danish Academy where "Eine ausfürliche Theorie der Komplexen Zahlen auf Geometrischen Grundlage ist enwickelt".

Remark. Wessel's article became most likely familiar to Niels Henrik Abel (1801-1829) when he visited in Copenhagen as a student in 1822. Two years later he demonstrated that the general algebraic equation of degree ≥ 5 cannot be solved by roots and radicals. Abel's proof laid the foundations for the modern theory of algebraic number fields. Complex analysis appears in Abel's famous article [Ab.2] from 1827 where several pioneering methods were introduced. He proved for example that if $f(z)$ is a doubly-periodic meromorphic function in $\mathbf{C}$, i.e. $f(z) = f(z+1) = f(z+i)$ hold for all $z = x + iy$, then the sum of its zeros in the open unit square $\{0 < x, y < 1\}$ minus the sum of its poles is equal to $p + qi$ where p, q are integers, and conversely there exist a doubly periodic meromorphic function with these zeros and poles when this condition holds.

The argument principle. Calculus using complex numbers was performed at an early stage by mathematicans such as Cauchy, Laguerre and Legendre. Here is a result due to Laguerre from 1820: Consider a monic polynomial of some even degree $2m$:

$$P(z) = z^{2m} + c_{2m-1}z^{2m-1} + \ldots + c_1 z + c_0$$

Separating real and imaginary parts of the complex coefficients we write $c_k = a_k + ib_k$ and get the polynomial

$$R(z) = z^{2m} + a_{2m-1}z^{2m-1} + \ldots + a_1 z + a_0$$

Suppose that $R(z)$ has some real zeros with odd multiplicity, i.e. if a is such a real zero then the signs of $P(a-\epsilon)$ and $P(a+\epsilon)$ differ for small ϵ. Let $\alpha_1 < \ldots < \alpha_k$ be this set of real zeros. Eventual real zeros of where R vanishes with an even order are not included. We have also the polynomial:

$$S(z) = b_{2m-1}x^{2m-1} + \ldots + b_1 z + b_0$$

Under the hypothesis that $S(\alpha_\nu) \neq 0$ for each $1 \leq \nu \leq k$ we shall learn in § XX that the number of zeros of $P(z)$ counted with multiplicities in the upper half-plane $\mathfrak{Im} z > 0$ is equal to

$$(*) \qquad m + \frac{1}{2} \cdot \sum_{\nu=1}^{\nu=k} (-1)^{\nu-1} \cdot \text{sign}(S(\alpha_\nu))$$

From (*) it is obvious that if $P(z)$ has all zeros in the open upper half-plane, then $k = 2m$, i.e. $R(x)$ has $2m$ simple real zeros $\alpha_1 < \ldots < \alpha_{2m}$. Moreover, (*) entails that $x \mapsto S(x)$ must change signs at these simple zeros and then calculus shows that the polynomial $S(z)$ has degree $2m-1$ with interlacing simple real zeros $\beta_1 < \ldots < \beta_{2n-1}$, i.e.

$$\alpha_1 < \beta_1 < \alpha_2 < \ldots < \alpha_{2m-1} < \beta_{2m-1} < \alpha_{2m}$$

Finally, the leading coefficient b_{2m-1} of the S-polynomial must be strictly negative. Hence Laguerre's theorem gives a necessary and sufficient condition in order that all zeros of $P(z)$ are located in the upper half-plane.

Fundamental solutions to differential operators. Bieberbach's claim that "many problems which are phrased in the real number system can be solved by a detour over complex numbers" is illustrated with the following result: Let $m \geq 2$ and consider an ODE-operator with polynomial coefficients

$$Q(x, \partial) = q_m(x)\partial^m + q_{m-1}(x)\partial^{m-1} + \ldots + q_1(x)\partial + q_0(x)$$

The q-polynomials have in general complex coefficients. Suppose that a is a real number where the leading polynomial q_m has a simple real zero. and assume also that

$$\frac{q_{m-1}(a)}{q'_m(a)} \neq \{0, 1, -2 \ldots\}$$

i.e. this quotient is not a non-positive integer. Under this assumption there exists a distribution μ on the real line supported by the closed half-line $(-\infty, a]$ such that $Q(x, \partial)(\mu) = 0$ with the property that the restriction of μ to the open half-line $(-\infty, a)$ is a real-analytic density which extends to a complex analytic function in a subdomain of the complex z-plane defined by

$$\{|\mathfrak{Im} z| < \delta\} \cap (\mathbf{C} \setminus [a, +\infty))$$

for some $\delta > 0$. In §§ xx we prove this result which employs a detour to the complex domain even if the problem at start deals with an ordinary differential equation on the real line.

About the contents. Subsection A treats basic material about complex numbers and § B is devoted to the fundamental theorem of algebra where Theorem B.2 gives a proof due to Augustine Cauchy in 1815. Interpolation formulas for polynomials appear in subsection 1:C and 1:D contains results about extremal Tchebyscheff polynomials.

Möbius functions are studied in § 2 which give examples of conformal mappings. The Laplace operator and the complex logarithm are introduced in § 3. Here some of the results are expressed in the real (x, y)-coordinates which enable us to apply ordinary calculus. Examples of complex mappings occur in § 4 and the stereographic projection is constructed in § 5 together with the spherical and the hyperbolic metrics. The reader is often asked to supplement the text with examples and here computers are helpful since plots give insight about the geometry.

1.A. The field of complex numbers

A complex number is expressed by $x+iy$ where (x,y) is a point in $\mathbf{R}^2$. This identifies the complex plane $\mathbf{C}$ with $\mathbf{R}^2$. When $z=x+iy$ we set

$$\mathfrak{Re}(z)=x \quad : \mathfrak{Im}(z)=y$$

where x is the real part and y the imaginary part of z. The sum of two complex numbers is defined by

(i) $$z_1+z_2=(x_1+x_2)+i(y_1+y_2)$$

Complex multiplication is defined by:

(ii) $$(x_1+iy_1)(x_2+iy_2)=(x_1x_1-y_1y_2)+i(x_1y_2+x_2y_1)$$

One verifies that the product satisfies the associative law. If $z=x+iy$ is non-zero we define the complex number:

(*) $$z^{-1}=\frac{x-iy}{x^2+y^2}\,.$$

One verifies that this gives a multiplicative inverse and conclude that the set of all complex number with zero included is a commutative field.

1.1 Conjugation and absolute value. If $z=x+iy$ its complex conjugate is $x-iy$ and denoted by $\bar{z}$. The absolute value of z is defined as $\sqrt{x^2+y^2}$ and denoted by $|z|$. The map $z\mapsto\bar{z}$ corresponds to reflection of plane vectors with respect to the x-axis and (*) gives

(**) $$z^{-1}=\frac{\bar{z}}{|z|^2}$$

1.2. The complex argument. In $\mathbf{R}^2$ we have polar coordinates (r,ϕ). If z is non-zero we write:

(1.2) $$z=r\cdot\cos\phi+i\cdot r\cdot\sin\phi \quad : \quad r=|z|\,.$$

The angle ϕ is denoted by $\arg(z)$ and called the argument of z. Since trigonometric functions are periodic, $\arg(z)$ is determined up to an integer multiple of 2π. Specific choices of $\arg(z)$ appear in different situations. As an example we consider the upper half-plane $\mathfrak{Im}(z)>0$ where one usually takes $0<\phi<\pi$ for $\arg(z)$. In the right half plane $\mathfrak{Re}(z)>0$ one takes $-\pi/2<\phi<\pi/2$. Another case occurs when we consider the polar representation of complex numbers z outside the *negative* real axis $(-\infty,0]$. Then every z has a unique polar representation in (1.2) with $-\pi<\phi<\pi$.

1.3. The complex number $e^{i\phi}$. This is a notstion for the complex number with absolute value one and argument ϕ. Thus

(1.3) $$e^{i\phi}=\cos\phi+i\cdot r\sin\phi\,,$$

where e as *Neper's constant* defined by

$$e=\lim_{n\to\infty}(1+\frac{1}{n})^n$$

The notation (1.3) stems from the Taylor series expansions of the sine- and the cosine functions. Recall from *Calculus* that

(i) $$\sin\phi=\sum_{\nu=0}^{\infty}(-1)^\nu\cdot\frac{\phi^{2\nu+1}}{(2\nu+1)!} \quad : \quad \cos\phi=\sum_{\nu=0}^{\infty}(-1)^\nu\cdot\frac{\phi^{2\nu}}{(2\nu!)}$$

Adding these series and using that $i^2=-1$ which gives $i^4=1$ and so on, we get:

(ii) $$\cos\phi+i\cdot\sin\phi=\sum_{\nu=0}^{\infty}\frac{(i\phi)^\nu}{\nu\,!}$$

The last series resembles the series of the real exponential function from *Calculus*:

(iii) $$e^x=\sum_{\nu=0}^{\infty}\frac{x^\nu}{\nu\,!} \quad : \quad x\in\mathbf{R}$$

1.4 Addition formula for arg(z). Euclidian geometry gives addition formulas for the sine-and the cosine functions:

$$\sin(\phi_1 + \phi_2) = \sin(\phi_1)\cos(\phi_2) + \sin(\phi_2)\cos(\phi_1) \tag{1}$$

$$\cos(\phi_1 + \phi_2) = \cos(\phi_1)\cos(\phi_2) - \sin(\phi_1)\sin(\phi_2) \tag{2}$$

Since (1) and (2) are essential in complex analysis we recall the proof. Let Δ be a triangle with angles α, β, γ where and A, B, C denote the opposed sides and consider the case when both α and β are $< \pi/2$. Now $\sin\gamma = \sin(\pi - \alpha - \beta) = \sin(\alpha + \beta)$ and the sine-theorem in euclidian geometry gives:

$$\frac{\sin\alpha}{A} = \frac{\sin\beta}{B} = \frac{\sin\gamma}{C} \tag{i}$$

Draw the normal line from the corner where the angle which hits the opposed side at a point whose distance to the α-corner is x. So then $C - x$ is the distance to the β-corner. Looking at a figure the reader can recognize that

$$\cos\alpha = \frac{x}{B} \quad \cos\beta = \frac{C - x}{A} \implies A \cdot \cos\beta + B \cdot \cos\alpha = C \tag{ii}$$

Together with (i) we get

$$\sin(\alpha + \beta) = \frac{C}{A}\sin\alpha = \sin\alpha \cdot \cos\beta + \sin\alpha \cdot \frac{B}{A}\cos\alpha$$

Finally, the first equality in (i) gives $\sin\alpha \cdot \frac{B}{A}\cos\alpha = \sin\beta \cdot \cos\alpha$ and the requested addition formula for the sine-function follows. A similar proof gives the addition formula for the cosine-function.

1.4 Exercise. Show that the construction of complex multiplication and (1.3) yield the equality

$$r_1 \cdot e^{\phi_1} \cdot r_2 \cdot e^{\phi_2} = r_1 r_2 \cdot e^{\phi_1 + \phi_2}$$

for all pairs of positive numbers r_1, r_2 and a pair of ϕ-angles. So when complex arguments are identified up to integer multiples of 2π we get:

$$\arg(z_1) + \arg(z_2) = \arg(z_1 z_2) \tag{3}$$

for each pair of non-zero complex numbers. By an induction over k the following hold for every k-tuple of complex numbers:

$$\sum_{\nu=1}^{\nu=k} \arg(z_\nu) = \arg\Big(\prod_{\nu=1}^{\nu=k} z_\nu\Big)\,. \tag{*}$$

We refer to (*) as the addition formula for the argument function.

1.5 Associated matrices. Let $z = a + ib$ be a complex number. Identifying $\mathbf{C}$ with $\mathbf{R}^2$ the complex multiplication with z yields a linear operator represented by a matrix. More precisely, the euclidian basis vectors e_1, e_2 correspond to the complex numbers 1 and i. Since $z \cdot i = ia - b$ the 2×2-matrix M_z associated to multiplication with z becomes

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

Notice that the determinant of M_z is $a^2 + b^2$ and the inversion formula (*) from 1.1 corresponds to the matrix identity

$$M_z^{-1} = \frac{1}{a^2 + b^2} \cdot \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

1.6 A polynomial approximation. With $0 \leq \phi \leq 2\pi$ we consider the ϕ-polynomials

$$P_n(\phi) = \big(1 + \frac{i\phi}{n}\big)^n \tag{i}$$

Notice that

$$\arg(1+\frac{i\phi}{n})=\frac{\phi}{n}$$

The addition formula (*) in 1.4 therefore gives

$$\arg(P_n(\phi))=\phi$$

for every n. Next, regarding absolute values we have

$$|1+\frac{i\phi}{n}|^2=1+\frac{\phi^2}{n^2}\implies |(P_n(\phi)|^2=\bigl(1+\frac{\phi^2}{n^2}\bigr)^n$$

Recall from calculus that $\log(1+t)\leq t$ for each real $t>0$. With $0\leq\phi\leq 2\pi$ it follows that

$$0\leq\log|P_n(\phi)|\leq\frac{\phi}{\sqrt{n}}\leq\frac{2\pi}{\sqrt{n}}$$

It follows that

$$\lim_{n\to\infty}|P_n(\phi)|=1$$

holds uniformly on $[0,2\pi]$ and the reader may also notice that Neper's limit formula for e entails that

$$\lim_{n\to\infty}P_n(\phi)=e^{i\phi}$$

Remark. The function $\phi\mapsto e^{i\phi}$ is 2π-periodic which entails that

$$\lim_{n\to\infty}(1+\frac{2\pi i}{n})^n=1$$

It is instructive to check this limit formula numerically with a computer for large integers n.

1.7 Complex numbers and geometry. Many results in euclidian geometry can be proved in a neat fashion by complex numbers. Let us give an example. Consider a triangle Δ with sides of length a,b,c. Let α be the angle at the corner p point opposed to the side of length a. Then

$$(*)\qquad \cos\alpha=\frac{b^2+c^2-a^2}{2bc}$$

To prove this we may without loss of generality assume that the corner point p is the origin and the two other corner points of Δ are represented by a pair of complex vectors z and w. Here

$$a^2=|z-w|^2\quad : |z|^2=b^2\quad : |w|^2=c^2$$

The formula (*) is invariant under dilation with z replaced by rz and w by rw for some r, and also under a rotation. So without loss of generality we can take $w=1$ and $z=x+iy$ with $x>0$. In this case a figure - or rather the definition of the cosine-function gives

$$\cos\alpha=\frac{x}{|z|}$$

So (*) amounts to prove the equation

$$\text{(i)}\qquad \frac{x}{|z|}=\frac{|z|^2+1-|1-z|^2}{2|z|}$$

Above $|z|$ is cancelled and we have

$$|z|^2+1-|1-z|^2=x^2+y^2+1-(y^2+(1-x)^2)=2x$$

which gives (i) and (*) follows. This illustrates how complex numbers provide an efficient tool to establish geometric formulas.

Exercise. Let Δ be a triangle with corner points at the origin, $(1,0)$ and $z_0=x_0+iy_0$ where $|z|\leq\sqrt{2}$ and both x_0 and y_0 are positive. The line ℓ passing $(1,0)$ which is $\perp$ to the vector z_0 consists of complex numbers of the form $1+\mathbf{R}\cdot iz_0$ where we use that the vectors z_0 and $i\cdot z_0$ are $\perp$ to each other. The normal from the corner point z_0 stays on the line $\{x=x_0\}$ and to get the intersection point of ℓ and this normal we seek a real number a such that

$$\text{(i)}\qquad x_0=1+ai(x_0+iy_0)\implies a=\frac{1-x_0}{y_0}$$

Hence the intersection point becomes $(x_0 + iy_*$ where (i) gives

$$y_* = x_0 \cdot \frac{1 - x_0}{y_0} \tag{ii}$$

Next, we draw the line from the origin passing (x_0, y_*) and it turns out that it is $\perp$ to the vector $1 - z_0$. This amounts to show that there exists a *real* number b such that

$$x_0 + iy_* = bi(1 - z_0) = by_0 + ib(1 - x_0) \tag{iii}$$

But this is clear from (ii) which shows that (iii) holds with $b = \frac{x_0}{y_0}$. So these complex computations verify the wellknown fact that the three normals intersect at a point.

B. The fundamental theorem of algebra.

Introduction. The proof of Theorem B.2 below was given by Cauchy in 1815 who employed that the absolute value of a complex-valued continuous function on a compact disc achieves its minimum some point. In the article [Weierstrass] from 1868 Weierstrass gave another proof. Here follows a citation from the introduction in [ibid]: *Obgleich wir gegenwärtig von dem in Rede stehenden Fundamentaltheoreme der Algebra eine Reihe strengen Beweise besitzen, so dürfte doch die Mitteilung der nachstehenden Begründung desselben, deren Eigenthümlichkeit hauptsächlich darin besteht, dass sie ohne Heranziehung von Hilfsmitteln und begriffen die der Algebra fremd sind, rein arithmetisch durchgeführt wird, vielen Mathematikern nicht unwillkommen sein.* So Weierstrass points out that in spite of the already known existence proofs, a procedure which is not remote from algebra derives the fundamental theorem of algebra by arithematical methods, a fact that might be appreciated by many mathematicians.

To acieve this Weierstrass considered symmetric polynomials in n indeterminates $\alpha_1,\ldots,\alpha_n$. Namely, to every *unordered* n-tuple of complex numbers $\alpha_1,\ldots,\alpha_n$ there exists the polynomial of the complex variable z defined by:

$$P(z)=\prod_{\nu=1}^{\nu=n}(z-\alpha_\nu)$$

This is a monic z-polynomial of degree n whose coefficients depend on the given n-tuple $\alpha_\bullet$ and we can write

$$P(z)=z^n+s_1(\alpha)z^{n-1}+\ldots+s_n(\alpha)$$

An algebraic manipulation shows that $s_k(\alpha)$ is a symmetric polynomial of the n-tuple of α-numbers, which is homogeneous of degree n. Weierstrass observation is that the fundamental theorem of algebra amounts to prove that for every n-tuple $c_0,c_1,\ldots,c_{n-1}$ of complex numbers there exists unique unordered n-tuple $\alpha_1,\ldots,\alpha_n$ such that

(i) $$c_k=s_k(\alpha)\quad:\ 1\leq k\leq n$$

In other words the mapping of unordered n-tuples of α-numbers to their associated symmetric s-polynomials is injective and the range is equal to all ordered complex n-tuples $c_0,c_1,\ldots,c_{n-1}$. Or equivalently, for each n-tuple of complex numbers $w_1,\ldots,w_n$ there exists a unique unordered n-tuple $\{\alpha_\nu\}$ such that

(i) $$\sum_{\nu=1}^{\nu=n}\alpha_\nu^k=w_k\quad:1\leq k\leq n$$

A device by Weierstrass. Consider a polynomial

(ii) $$P(z)=z^n+c_1z^{n-1}+\ldots+c_n$$

It has simple zeros if and only if the ideal generated by $P(z)$ and its derivative $P'(z)$ is equal to $\mathbf{C}[z]$, i.e. if and only if there exists a unique pair of polynomials A,B such that

(iii) $$1=A(z)P(z)+B(z)P'(z)$$

where $\deg A\leq n-2$. The existence of such a pair A,B is equivalent to the existence of a solution of a linear system of equations in $2n-1$ many indeterminates corresponding to coefficients of A and B.

Exercsie. Use Cramer's rule to show that (iii) has a solution for a pair of polynomial A and B as above if and only if

(iv) $$\mathcal{D}_n(c_0,\ldots,c_{n-1})=0$$

where $\mathcal{D}_n$ is a polynomial in n variables with integer coefficients. The exact formula for $\mathcal{D}_n$ can be found in text-books devoted to algebra.

In the article [Weierstrass] it is proved that if one starts with an n-tuple $\{c_\nu\}$ for which $P(z)$ has simple zeros, then for each $\epsilon > 0$ one can perform a finite number of arithmetical operations which give an unordered n-tuple of complex numbers $a_1, \ldots, a_n$ such that the polynomial

$$Q(z) = \prod_{\nu=1}^{\nu=n} (z - a_\nu) = z^n + \sum c_\nu^* z^\nu$$

has coefficients for which

$$|c_\nu - c_\nu^*| < \epsilon \quad : 0 \leq \nu \leq n-1$$

Next, starting with a sufficiently small ϵ Weierstrass proved that the roots of the Q-polynomial can be used in a recursive formula to attain the true roots of $P(z)$. More precisely, put

$$a_\nu^{(1)} = a_\nu - \frac{P(a_\nu)}{\prod_{j \neq \nu}(a_\nu - a_j)}$$

Inductively we put:

$$a_\nu^{(k+1)} = a_\nu - \frac{P(a_\nu^{(k)})}{\prod_{j \neq \nu}(a_\nu^{(k)} - a_j^{k)})}$$

Then it is proved in [ibid] that the true roots of P are given by

$$a_\nu^* = \lim_{k \to \infty} a_\nu^{(k)}$$

Moreover, the rate of convergence is rapid in the sense that there is a constant C which depends on P and the choice of ϵ such that

$$|a_\nu^* - a_\nu^{(k)}| \leq C \cdot 2^{-k} \qquad \text{for every} \quad 1 \leq \nu \leq n$$

Weierstrass' constructions can be implemented into a computer which leads to to approximations of zeros polynomials with high accuracy. So readers interested in numerical investigations should consult the rich material in [Weierstrass].

Cauchy's proof

.

Here we admit the existence of extremal values taken by continuous functions on compact sets. Let $P(z)$ be given in (ii) above, If P has a zero α one gets a factorisation

$$P(z) = (z - \alpha)(z^{n-1} + d_{n-2}z^{n-2} + \ldots + d_1 z + d_0)$$

where the d-coefficients are found by algebraic identities. One has for example

$$d_{n-2} = c_{n-1} - \alpha \quad : \quad d_{n-3} = c_{n-2} - \alpha d_{n-2}$$

and so on. If the factor polynomial of degree $n-1$ also has a complex root we can continue and conclude

Proposition. *Assume that every polynomial $P(z)$ has at least one complex root. Then it has a factorisation*

$$P(z) = \prod_{\nu=1}^{\nu=k} (z - \alpha_\nu)$$

Here k is the degree of P and $\alpha_1, \ldots, \alpha_k$ is a k-tuple of complex numbers where repetitions occur when P has multiple roots.

Hence the fundamental theorem of algebra follows if we have proved:

B.1 Theorem *Every polynomial $P(z)$ has at least one root.*

Proof. We are given $P(z)$ as in (ii) above and put $M = |c_0| + \ldots + |c_{n-1}|$. If $|z| \geq 1$ the triangle inequality gives

(i) $$|P(z)| \geq |z|^n - M \cdot |z|^{n-1} \geq |z| - M$$

With $R = M + 2 \cdot |c_0| + 1$ it follows that

(ii) $$|z| \geq R \implies |P(z)| \geq R - M \geq 2 \cdot |c_0| = 2 \cdot |P(0)|$$

Next, the restriction of $P(z)$ to the closed disc $|z| \leq R$ is a continuous function and therefore the absolute value takes a minimum at some point z_0 which in particular gives $|P(z_0)| \leq |P(0)|$. Hence (ii) implies that we have a global minimum, i.e.

(iii) $$|P(z_0)| \leq |P(z)|$$

hold for all z. To show that (iii) entails $P(z_0) = 0$ we argue by contradiction, i.e suppose that $P(z_0) \neq 0$ and with a new variable ζ we get the polynomial

(iv) $$P(z_0 + \zeta) = P(z_0) + d_m\zeta^m + d_{m+1}\zeta^{m+1} + \ldots + d_n\zeta^n$$

where $1 \leq m \leq n$ and $d_m \neq 0$. We find real numbers α, β such that

(v) $$P(z_0) = |P(z_0)|e^{i\alpha} \quad \text{and} \quad d_m = |d_m|e^{i\beta}$$

Next, with $\epsilon > 0$ we set

(vi) $$\zeta = \epsilon \cdot e^{i \cdot \frac{\pi+\alpha-\beta}{m}}$$

Since $e^{i\pi} = -1$ this choice of ζ together with (v) gives

(vi) $$P(z_0) + d_m\zeta^m = (1 - |d_m| \cdot \epsilon^m)P(z_0)$$

Put $M^* = |d_{m+1}| + |d_{m+2}| + \ldots |d_n|$. When $\epsilon < 1$ the triangle inequality gives

(vii) $$|d_{m+1}\zeta^{m+1} + d_{m+2}\zeta^{m+2} + \ldots d_k z^k| \leq M \cdot \epsilon^{m+1}$$

Together with (vii) another application of the triangle inequality gives:

$$|P(z_0 + \epsilon \cdot e^{i \cdot \frac{\pi+\alpha-\beta}{m}})| \leq |P(z_0)|(1 - |d_m|\epsilon^m| + M \cdot \epsilon^{m+1} =$$

(viii) $$|P(z_0)| - \epsilon^m \big(|d_m| \cdot |P(z_0)| - M \cdot \epsilon\big)$$

Now we can take

$$0 < \epsilon < \frac{|d_m| \cdot |P(z_0)|}{M}$$

and then (viii) gives a strict inequality

$$|P(z_0 + \zeta)| < |P(z_0)|$$

This contradicts that z_0 gave a minimum for the absolute value of P and the proof is finished.

Remark. The proof below relies upon the fact that absolute values of complex polynomials cannot achieve local minima. Consider as an example some integer $k \geq 2$ and the function

$$g(z) = |1 + z^k|^2$$

Here $g(0) = 1$ but $z = 0$ is not a minimum for with a small $\epsilon > 0$ we can take $z = \epsilon \cdot e^{\pi i/k}$ which gives $z^k = \epsilon^k$ and hence

$$g(\epsilon \cdot e^{\pi i/k}) = (1 - \epsilon^k)^2 < 1$$

Notice the contrast to arbitrary real polynomials where a minimum can occur. For example, the polynomial $g(x, y) = 1 + x^4 + x^2y^2 + y^4$ has a minimum at the origin and no zeros in the (x, y)-plane.

Proof by residue calculus.

Later Cauchy gave other proofs using reside theory in his famous text-books published around 1830 which are devoted to analytic functions. For if the polynomial $P(z)$ in (ii) has no complex zeros then $P^{-1}(z)$ is an entire function and taking the complex derivative $P'(z)$ it follows that the complex line integrals

(*) $$\int_{|z|=R} \frac{P'(z)}{P(z)}\, dz = 0$$

for all $R > 0$. When the line integral is evaluated in polar coordinates it becomes

$$\int_0^{2\pi} \frac{n + (n-1)c_{n-1}R^{-1}e^{-i\theta} + \ldots c_1 R^{-n-1}e^{-i(n-1)\theta}}{1 + c_{n-1}R^{-1}e^{-i\theta} + \ldots + c_0 R^{-n}e^{-in\theta}}\, d\theta$$

Passing to the limit as $R \to +\infty$ the last integral converges to n which gives the contradiction. Cauchy concluded that P must have at least one zero and later on we wil show that residue calculus immediately entails that the number of zeros counting multiplicities is equal to the degree of the polynomial.

Abel's theorem.

If the polynomial $P(z)$ has degree ≤ 4 one can find the roots by *Cardano's formula*. See § B.3 for a details. But as soon as the degree is ≥ 5 it is in general not possible to find the zeros of a polynomial by roots and radicals even if the coefficients are integers. This was proved by Niels Henrik Abel in 1823. His article [Ab:1] published in the first volume of Crelle's Journal contains pioneering results about algebraic field extensions. Abel used his new discoveries in algebraic field theory to prove that the general algebraic equation of degree ≥ 5 cannot be solved by roots and radicals by investigating a system of 120 linear equations expressed by the coefficients of a polynomial in degree ≥ 5. An example from Abel's work where a Cardano solution fails is the equation

$$z^5 + z + 1 = 0$$

For an account about Abel's original work the reader should consult articles from *The Abel Legacy* published in 2004 on the occasion of the first Abel Prize. After Abel's decease in 1829, Everiste Galois constructed a group to every polynomial of arbitrary degree and arrived at another proof of Abels result. In this way one is led to *Galois theory* which brings the theory about field extensions together with group theory which has become a central topic in algebra. Many text-books treat Galois theory. Personally I recomend Emil Artin's lectures from 1948 where Galois theory is presented in a masterful manner.

An algebraic problem.

Consider a pair of polynomials

$$p(z) = z^n + a_1 z^{n-1} + \ldots + a_{n-1}z + a_n$$

$$q(z) = z^n + b_1 z^{n-1} + \ldots + b_{n-1}z + b_n$$

where $\{a_\nu\}$ and $\{b_\nu\}$ are rational numbers. Both polynomials are assumed to be irreducible in the unique factorization domain $Q[z]$ which entails that the roots of p and q are simple. By the fundamental theorem of algebra we can write

$$p(z) = \prod (z - \alpha_j) \quad \text{and} \quad q(z) = \prod (z - \beta_j)$$

Each root α_j of p generates a field $K = Q[\alpha_j]$ which as a vector space of Q has dimension n and a basis is given by $1, \alpha_j, \ldots, \alpha_j^{n-1}$. In fact, this holds since $p(z)$ was irreducible and we remark that the field K is isomorphic to the field $\frac{Q[z]}{(p)}$ where (p) denotes the pirnicipal ideal generated by p in the polynomial ring $Q[z]$. Similar conclusions hold for the roots of q. Now one may ask when there exists a pair of roots α_j, β_ν for p and q respectively, such that the fields $K[\alpha_j]$ and $K[\beta_\nu]$ are equal. By elementary field theory the necessary and sufficient condition for the equality $K[\alpha_j] = K[\beta_\nu]$ is that

$$\text{(i)} \qquad \beta_\nu = q_0 + q_1 \cdot \alpha_j + \ldots + q_{n-1}\alpha_j^{n-1}$$

holds for some n-tuple $\{q_\nu\}$ of rational numbers. The problem is to find equations satisfied by the pair of n-tuples $\{a_j\}$ and $\{b_k\}$ which appear as coefficients of the two polynomials in order that (i) holds for some pair of roots. This is a problem in algebraic elimination theory and solved as follows: Let $\lambda, \xi_0, \ldots, \xi_{n-1}$ be $n+1$ many new variables and set

$$S(\lambda, \xi_0, \ldots, \xi_{n-1}) = \prod_{j=1}^{j=n} \left(\lambda - (\xi_0 + \xi_1\alpha_j + \ldots + \xi_{n-1}\alpha_j^{n-1})\right)$$

This is a *symmetric* expression in the n-tuple of roots of p and text-books in elementary algebra teaches that every symmetric polynomial of the roots can be expressed as a polynomial of the coefficients with integer coefficients. It follows that

$$\text{(ii)} \qquad S(\lambda, \xi_0, \ldots, \xi_{n-1}) = \sum_{j=0}^{j=n-1} \phi_j(a_1, \ldots, a_n, \xi_0, \ldots, \xi_{n-1}) \cdot \lambda^j$$

where $\{\phi_j\}$ are polynomials with integer coefficients of the two n-tuples $\{a_\nu\}$ and $\{\xi_\nu\}$ expressed by explicit interpolatation formulas. With these notations, (1) is satisfied for a pair of roots if and only if the λ-polynomial in (ii) has at least one root in common with q. To check if this holds one employs a determinant of a certain $2n \times 2n$-matrix whose elements are determined explicitly by the coefficients $\{b_\nu\}$ and the n-tuple $\phi_j(a, \xi)$. See Exercise § xx from § I:C for this. The conclusion is that there exists a polynomial of the ξ-variables

$$\mathcal{S}(\xi_0, \ldots, \xi_{n-1}) = \sum \rho_\gamma(a_\bullet, b_\bullet) \cdot \xi^\gamma$$

where $\{\rho_\gamma\}$ are polynomials of the $2n$-tuple formed by the coefficients of the given polynomials and $\gamma = (\gamma_0, \ldots \gamma_{n-1})$ are multi-indices expressing the monomials

$$\xi^\gamma = \xi_0^{\gamma_0} \cdots \xi_{n-1}^{\gamma_{n-1}}$$

Now (i) has a solution with rational numbers $\{q_\nu\}$ if and only if $\mathcal{S}(q_0, \ldots, q_{n-1}) = 0$. In other words, the necessary and sufficient condition to obtain (i) for a pair of roots is that the $\mathcal{S}$-polynomial of n variables has at least one zero in the n-dimensional ξ-space given by an n-tuple of rational numbers. Using terminology from algebraic geometry it means that the algebraic hypersurface $\{\mathcal{S} = 0\}$ contains at least one rational point. This example gives a glimpse of elimination theory where the problems consist in finding various algorithmic formulas. Concerning the specific problem above we remark that calculations which lead to equations in order that (i) holds were carried out in work by Delannay and Tschebotaröw for polynomials of degree ≤ 4. The interested reader may consult the plenary talk by Tschebotaröw from the IMU-congress at Zürich in 1932 which describes the interplay between the problem above and Galois theory. It appears that a complete investigation for arbitrary large n remains unsettled.

B.2 Algebraic numbers. Of special interest are complex numbers which are *algebraic* over the field Q of rational numbers, i.e. complex numbers α which are roots to some polynomial whose coefficients are rational numbers. The set of all such complex numbers is a subfield of $\mathbf{C}$ denoted by A. Inside A there occur subfields generated by roots to a finite family of polynomials. These subfields are finite dimensional vector spaces over Q and called finite algebraic fields. Given such a field K one then gets a subring $\mathcal{D}(K)$ which consists of all $\alpha \in K$ such that α is a root of a monic polynomial with integer coefficients, i.e. α satisfies an equation

$$\alpha^m + c_{m-1}\alpha^{m-1} + \ldots + c_1\alpha + c_0 \quad : \quad c_0, \ldots, c_{m-1} \text{ are integers}$$

The ring $\mathcal{D}(K)$ is a Dedekind ring and enjoys nice properties which are exposed in text-books devoted to algebraic number fields. Analytic function theory is used to study approximations of algebraic numbers by rational numbers. Let ξ be a positive real number which satisfies an algebraic equation

$$\xi^n + c_{n-1}\xi^{n-1} + \ldots + c_1\xi + c_0 = 0$$

where $\{c_\nu\}$ are integers and the polynnomial $P(z) = z^n + \sum c_\nu z^\nu$ is irreducible in the unique factorisation domain $Q[z]$. In 1908 Thue proved a remarkable result in the article *Bemerkungen* *über gewisse Näherungsbrüche algebraishen Zahlen.* Namely, for every $\epsilon > 0$ the set of positive rational numbers $q = \frac{x}{y}$ such that

$$(*) \qquad \left|\xi - \frac{x}{y}\right| \leq \frac{1}{y^{\frac{n}{2}+1+\epsilon}}$$

is finite. Thue's result means that there are lower bounds for approximations of algebraic integers which are not rational numbers. In his thesis from 1921, Siegel proved that if ξ is as above then the set of rational numbers $\frac{x}{y}$ for which

$$(**) \qquad \left|\xi - \frac{x}{y}\right| \leq \frac{1}{y^{2\sqrt{n}}}$$

is finite. Notice that Sigel's result improves (*) as soon as $n \geq 16$. The proof of (**) is quite involved. The interested reader may consult Siegel's article *Über Näherungswerte algebraischen Zahlen*(Math. Zeitschrift 1921) for refined results about approximations of algebraic numbers by rationals. But let us give one of the minor steps from Siegel's impressive work.

An inequality by Siegel. Let $p(z) = z^n + c_{n-1}z^{n-1} + \ldots + c_1 z + c_0$ be a polynomial with integer coefficients. Suppose that $p(z)$ has a factorisation in the polynomial ring $Q[z]$:

$$p(z) = (k_0 z^m + \ldots + k_{m-1}z + k_m) \cdot q(z)$$

where $1 \leq m < n$ and $k_0, \ldots, k_m$ are integers with no common divisor ≥ 2 while $q(z)$ is a polynomial of degree $n - m$ in $Q[z]$. Set

$$\rho^* = \max\{|c_0|, \ldots, |c_n|\} \quad \text{and} \quad \rho_* = \max\{|k_0|, \ldots, |k_m|\}$$

Then one has the inequality

$$(*) \qquad \frac{\rho_*}{\rho^*} \leq (m+1) \cdots n$$

Proof. Consider first a polynomial $f(z) = a_0 z^k + \ldots + a_k$ of some degree $k \geq 1$ with arbitrary complex coefficients. Let $\lambda \neq 0$ be another complex numbers and set

$$g(z) = (z - \lambda) f(z) = d_0 z^{k+1} + \ldots + d_k z + d_{k+1}$$

Let $d^* = \max\{|d_\nu|\}$ and $c^* = \max\{|c_\nu\}$. Then one has the inequality

$$(1) \qquad \frac{a^*}{d^*} \leq k + 1$$

To prove (1) we notice that

$$a_\nu = d_0 \lambda^\nu + d_1 \lambda^{\nu-1} + \ldots + d_\nu \quad : \; 0 \leq \nu \leq k$$

If $|\lambda| \leq 1$ it follows that

$$|a_\nu| \leq |d_0| + |d_1| + + \ldots + |d_\nu| \leq (\nu + 1) \cdot d^*$$

Since this holds for every ν we get $a^* \leq (k+1)d^*$ as requested. Next, if $|\lambda| > 1$ we rewrite (1) so that

$$(z - \frac{1}{\lambda})(a_k z^k + \cdots + a_0) = -\frac{1}{\lambda} \cdot (d_0 + \ldots + d_{k+1} z^{k+1})$$

Since $\frac{1}{\lambda}$ has absolute value ≤ 1 the previous case entails that

$$a^* \leq (k+1)\frac{d^*}{|\lambda|} \leq (k+1)d^*$$

and hence (1) also holds when $|\lambda| \geq 1$. To prove (*) we consider the factorisation

$$q(z) = k_0^{-1}(z - \lambda_1) \cdots (z - \lambda_{n-m})$$

Hence the polynomial $p(z)$ arises from $p_*(z) = k_0^{-1}(k_0 z^m + \cdots + k_m)$ via an $(n-m)$-fold application of the case above and from this the reader can deduce that

$$\frac{\rho_*}{\rho^*} \leq (m+1) \cdots n$$

which proves (*).

B.3 Solutions by roots and radicals.

Here follows Cardano's construction of roots to the specific polynomial

(*) $$P(x) = x^3 + x + 1$$

The derivative $P'(x) = 3x^2 + 1 > 0$ for all x. Hence $P(x)$ is strictly increasing on the real x-axis and has one real root x_* which is < 0. To find x_* Cardano proceeded as follows. Let u and v be a pair of *independent* variables and regard the function

$$f(u,v) = (u+v)^3 + u + v + 1 = u^3 + v^3 + (u+v)(3uv+1) + 1$$

When $v = -\frac{1}{3u}$ we see that (*) is zero at $x = u + v$ if

(**) $$u^3 - \frac{1}{27u^3} + 1 = 0$$

With $\xi = u^3$ this yields the second order algebraic equation

$$27\xi^2 - 1 + 27\xi = 0$$

It is rewritten as

$$(\xi + \frac{1}{2})^2 = \frac{1}{27} + \frac{1}{4}$$

Here we find the positive root

$$\xi_* = \sqrt{\frac{1}{27} + \frac{1}{4}} - \frac{1}{2}$$

Hence $P(x)$ has the real root

(***) $$x_* = \xi_*^{\frac{1}{3}} - \frac{1}{3} \cdot \xi_*^{-\frac{1}{3}}$$

where the reader should confirm that $x_* < 0$. There remains to find the two complex roots. Since P has real coefficients they occur in a conjugate pair, i.e. the complex roots are of the form $a + ib$ and $a - ib$ where we may take $b > 0$. Since the real root ξ_* has been found one can determine a and b as follows: By a wellknown algebraic identity the sum of the three roots of P is zero, i.e. it holds in our specific example since the x^2-coefficient of P is zero. We conclude that

(1) $$a = -\frac{x_*}{2}$$

Next, the product of the three roots is equal to -1. This gives

$$(a^2 + b^2) \cdot x_* = -1 \implies b^2 = -\frac{1}{x_*} - \frac{x_*^2}{4}$$

The reader should verify that the last term is > 0 and hence

$$b = \sqrt{-\frac{1}{x_*} - \frac{x_*^2}{4}}$$

We refer to text-books in algebra for the procedure to solve an arbitrary equation of degree ≤ 4 by roots and radicals.

B.4 Real roots

Consider a polynomial $p(x)$ of some degree $k \geq 1$ with real coefficients:

$$p(x) = x^k + a_{k-1}x^{k-1} + \ldots + a_1 x + a_0$$

Assume that at least some $a_\nu < 0$. This gives a strictly decreasing sequence

$$n_1 > n_2 > \ldots$$

where n_1 is the largest integer such that $a_{n_1} < 0$ and then n_2 is the largest integer $< n_1$ such that $a_{n_2} > 0$ if such an integer exists. In the next stage $a_{n_3} < 0$ and so on. The result is an integer $1 \leq \rho_+(p) \leq k$ which counts the number of times when non-zero coefficients of p change signs. With $p(x) = x^4 - x^3 + x + 1$ we have $\rho_+(p) = 2$ while the ρ_+-number of $x^4 - x^3 + x - 1$ is 3. A result due to René Descartes which is exposed Newton's in text-book [Newton] from 1666 goes as follows: Assume that $p(x)$ has at least one real zero $\alpha > 0$ and denote by $N_+(p)$ the number of positive real zeros counted with multiplicity.

B.4.1 Rule of Cartesi. *The difference $\rho_+(p) - N_+(p)$ is an non-negative even integer.*

Example. Let

$$p(x) = x^5 - x^3 + ax - 1$$

where the constant $a > 0$. Since $p(0) = -1$ is is clear that p has at least one positive root and here $\rho_+(p) = 3$. So the Cartesi rule entails that $N_+(p)$ is 1 or 3. To analyze which case occurs we consider the derivative

$$p'(x) = 5x^4 - 3x^2 + a$$

here $\rho_+(p') = 2$ so p' can have two positive roots or be everywhere > 0 when $x > 0$. In the last case p is strictly increasing and has exactly one positive real root. Next, the reader can verify that p' has two real roots if and only if

$$a < \frac{9}{100}$$

whose corresponding roots become:

$$x_1 = \frac{3}{10} - \sqrt{\frac{9}{100} - a} \quad : \quad x_2 = \frac{3}{10} + \sqrt{\frac{9}{100} - a}$$

By drawing a picture one sees that p has 3 positive roots if and only if

(i) $$p(x_1) > 0 \quad \text{and} \quad p(x_2) < 0$$

1. Exercise. Determine all real $a < \frac{9}{100}$ for which (i) holds. The result can be checked numerically with a computer.

2. Exercise. Prove Descartes' the result in B.4.1. The hint is to use that if p is an arbitrary polynomial with real coefficients such that $p(0) \neq 0$ and $a > 0$ some real number then the difference

$$\rho_+((x-a)p(x)) - \rho_+(p)$$

is a non-negative even integer. After this an induction over the degree of p finishes the proof.

5.4.2 Sturm chains.

Let $p(x)$ be a polynomial with real coefficients. Set $f_0 = p$ and $f_1 = p'(x)$. Euclidian divisions give a sequence of polynomials $f_0, f_1, f_2, \ldots$ where

$$f_\nu = Q_\nu \cdot f_{\nu+1} - f_{\nu+2}$$

and $\deg(f_{\nu+2}) < \deg(f_{\nu+1})$ hold for each $\nu \geq 0$. After a finite number of steps the division stops and if the last non-vanishing polynomial f_m is a non-zero constant we say that p has a Sturm chain. Here $m \leq k$ holds and in general strict inequality occurs. For example, let $p = x^4 + 1$ which gives

$$x^4 + 1 = \frac{x}{4} \cdot 4x^3 + 1$$

So here $f_2 = -1$ and $m = 2$. Suppose that a polynomial p has a Sturm chain. Then every real zero must be simple. For if $p(a) = p'(a) = 0$ for some real a we get $f_2(a) =$ and then the equation

$$f_1 = Q_1 f_2 - f_3$$

entails that $f_3(a) = 0$. By induction we see that $f_\nu(a) = 0$ for every ν which is impossible since the last polynomial f_m by the hypothesis is a non-zero constant. Next, let α be some real number where each $f_\nu(\alpha) \neq 0$ and consider the signs of the $(m+1)$-tuple $\{f_\nu(\alpha)$. Denote by $\mathcal{N}_p(\alpha)$ the number of sign changes of this sequence. For example, with $p = x^4 + 1$ the sign sequence at $x = -1$ becomes

$$+ : -1; -1 \implies \mathcal{N}_p(-1) = 1$$

When $x = 1$ we get the sign sequence

$$+; +; - \quad \implies \quad \mathcal{N}_p(x) = 1$$

Next, let $p(x) = x^4 - 1$ which gives $f_2(x) = 1$. So $x < -1$ gives the sign sequence

$$+ : -; + \implies \mathcal{N}_p(-1) = 2$$

When $x > 1$ we get the sign sequence

$$+ : +; + \implies \mathcal{N}_p(x) = 0$$

So here the difference $\mathcal{N}_p(-1-\delta) - \mathcal{N}(1+\delta) = 2$ for each $\delta > 0$ and at the same time we notice that $x^4 - 1$ has simple zeros at $x = 1$ and $x = -1$ in the interval $(-1-\delta, 1+\delta)$. It turns out that this example is not special since the following general result holds:

B.4.2 Theorem. *Let p be a polynomial which has a Sturm chain. Then all real zeros of p are simple and if (a, b) is an interval where $p(a)$ and $p(b)$ are $\neq 0$, the number of zeros in (a, b) is equal to $\mathcal{N}(a) - \mathcal{N}(b)$.*

Proof. Consider a zero $a < x_* < b$ of p. Since p has a Sturm chain we know from the above that the zeros are simple and that $f_\nu(x_*) \neq 0$ for each $1 \leq \nu \leq m$. For the sign of $p'(x_*) = f_1(x_*)$ two cases may occur. If $f_1(x_*) > 0$ then $p(x)$ is strictly increasing close to x_* which entails that $p(x_* - \delta) < 0$ for small positive δ while $p(x_* + \delta) > 0$. From this the reader may verify that

$$\mathcal{N}(x_* - \delta) - \mathcal{N}(x_* + \delta) = 1$$

for small positive δ. Similarly, if $p'(x_*) < 0$ we have $p(x_* - \delta) > 0$ and $p(x_* + \delta) < 0$ and again the reader can verify that (1) holds. Hence the $\mathcal{N}$-function decreases with a unit whenever we pass a zero of p. Now Sturm's theorem follows if we prove that $\mathcal{N}$ is unchanged when we pass a zero of some f_ν with $1 \leq \nu \leq m-1$. So let us suppose that $f_\nu(x_*) = 0$ holds for some $a < x_* < b$ while $f_{\nu-1}(x_*) \neq 0$. The equation

$$f_{\nu-1} = Q_{|u-1} f_\nu - f_{\nu+1}$$

shows that the signs of $f_{\nu-1}(x_*)$ and $f_{\nu+1}(x_*)$ are different. The sign sequence at some $x = x_* - \delta$ therefore contains a triple at place $\nu - 1, \nu, \nu + 1$ of one of the following four strings:

$$+; +; - \quad : +; -; - \quad : -; -; + \quad : -; -; +$$

where the first two cases occur when $f_{\nu-1}(x_*) > 0$ while $f_{\nu+1}(x_*) < 0$, and the last two with reversed signs. Suppose for example that the first triple occurs above so that $f_\nu(x_* - \delta) > 0$ for small $\delta > 0$. If we pass to $x_* + \delta$ where $f_\nu(x_* + \delta) < 0$ the new triple becomes

$$+; -; -$$

The number of sign changes of this triple is equal to that of $+ : +, -$, i.e no change occurs for the number of sign changes. The reader may check the other possible cases and conclude that a zero of f_ν with $\nu \geq 1$ does not give rise to a jump of the $\mathcal{N}$-function as x increases and Sturm's theorem follows.

Example. Consider the cubic polynomial

$$p(x) = x^3 - 3x^2 + 1$$

We get

$$p(x) = (\frac{x}{3} - \frac{1}{3})(3x^2 - 6x) - 2x + 1$$

Hence $f_2(x) = 2x - 1$ and finally:

$$3x^2 - 6x = (\frac{3x}{2} - \frac{9}{4})(2x - 1) - \frac{9}{4}$$

which gives $f_3 = \frac{9}{4}$. If $x = -1$ we get the sign chain

$$-; +; -; +$$

Hence $\mathcal{N}(-1) = 3$. Next, if $x = -\delta$ for s small $\delta > 0$ we get the sign chain

$$+; +; -; + \implies \mathcal{N}(-\delta) = 2$$

It follows that p has some simple zero in $(-1, 0)$. We leave it to the reader to show that Sturm's theorem entails that p also has a simple zero in $(0, 1)$ and in $(2, 3)$. Thus, p has three simple zeros which can be checked by an approximative plot of its graph and confirms Sturm's theorem.

B.4.3 Graeffe's approximation.

A classic procedure to approximate real roots of polynomials goes back to Newton and is exposed in many text-books. A more recent method was introduced by Graeffe in 1826 which has the merit that it can be used to approximate complex zeros. Let us illustrate Graeffe's method for polynomials of degree 3. Consider a cubic polynomial:

$$f(x) = x^3 - ax^2 + bx - c$$

which is assumed to have three simple roots α, β, γ which in general are complex. Set

$$f_1(x) = x(x + b)^2 - (ax + c)^2 = x^3 - a_1x^2 + b_1x - c_1$$

A computation which is left to the reader shows that the roots of f_1 are $\alpha^2, \beta^2, \gamma^2$. Following Graeffe we construct for each $n \geq 1$ the cubic polynomial

$$f_n(x) = x^3 - a_nx^2 + b_nx - c_n$$

whose roots are the 2^n-powers of the roots of f.

Exercise. Verify Graeffe's recursive formulas:

$$a_n = a_{n-1}^2 - 2b_{n-1} \quad : \quad b_n = b_{n-1}^2 - 2a_{n-1}c_{n-1} \quad : \quad c_n = c_{n-1}^2$$

Suppose that $|\alpha| > |\beta| > |\gamma|$ and deduce the limit formula

$$|\alpha| = \lim_{n\to\infty} |a_n|^{\frac{1}{2^n}}$$

This gives a procedure to find the maximum of absolute values of complex roots to polynomials. We discuss this further in § B.5 below.

B.4.4 Semi-algebraic sets.

To study the collection of real roots in a finite family of polynomials certain sign-symbols have been introduced by Hörmander and leads to Theorem B.5 below. The combined sign of an m-tuple of polynomials is denoted by $\text{SIGN}(p_1, \ldots, p_m)$ and registers, in increasing order, all the zeros of these polynomials and the signs of all polynomials at each zero and every interval, including the intervals which stretch to $+\infty$ or $-\infty$. For a single polynomial $p(x)$ the sign is given by a finite ordered sequence of + or -and 0 expressing eventual zeros of p and signs of $p(x)$ just before or after one zero. For example, if $p(x) = x^2 - 1$ one writes:

$$+; 0; -; 0; +$$

which reflects that $p > 0$ for large negative x, has a zero at $x = -1$ and is < 0 in the interval $(-1, 1)$ and is > 0 after the zero at $x = 1$. The sign sequence of a polynomial p without any real zeros is reduced to a single + if it is > 0 or a single minus sign if $p(x) < 0$ holds on the x-line.

Next, if p and q is a pair of polynomials the sign chain is expressed by pairs at each stage. For example, if $p(x) = x^2 - 1$ and $q(x) = x$ then SIGN(p, q) is given by

$$+/-; 0/-; -/0; -/+; 0/+ : +/+$$

Above the first symbol $+/-$ indicates that $p(x) > 0$ when $x < -1$ while $q(x) < 0$. The second term $0/-$ is the zero of p at $x = -1$ and the extra minus sign indicates that $q(-1) < 0$. The third term $-/0$ is the zero of q at $x = 0$ where the minus sign above 0 appears since $p(0) < 0$.

Sign-chains become more involved when the number m of polynomials increases. But they can always be found in an algorithmic way by an induction over the maximum of the degrees in a family of polynomials. The induction relies upon the following:

B.4.5 Theorem. *Let $p(x)$ be a real polynomial and $r(x)$ the polynomial after an euclidian division*

$$p = A \cdot p' + r$$

where $p'(x)$ is the derivative and $deg(r) \leq deg(p) - 2$. Then SIGN(p', r) determines SIGN(p) .

Proof. Let $p(x) = a_n x^n + \ldots + a_0$ where $a_n \neq 0$. Replacing p by $-p$ reverse all signs for p but also for the pair (p', r). So without loss of generality we can assume that $a_n > 0$. Next, when $x << 0$ then SIGN(p) starts with + if n is even and at the same time $p'(x) < 0$ for $x << 0$ so the sign-chain of (p', r) decides if n is even or not.

Consider for example the case when n is even and let x_0 be the first zero of p' which must exist since $p'(x) < 0$ when $x << 0$ while $p'(x) > 0$ when $x >> 0$. Now $p'(x) < 0$ for all $x < x_0$ which means that that $x \to p(x)$ is strictly decreasing for $x < x_0$ and therefore the sign-sequence of p is determined on this interval, i.e there only occurs + if p has no zero or otherwise it attains a zero and its sign-sequence starts eith $+; 0; -$ prior to x_0. Next, at x_0 the sign of $r(x_0)$ determines that of $p(x_0)$ where one does not exclude the case when $r(x_0) = 0$ which would give a zero for p at x_0. Now we pass to the (eventual) next zero $x_1 > x_0$ and whatever is the sign of $p'(x)$ on $(x_0, x_1$ we know at least that $x \mapsto p(x)$ is strictly increasing or strictly decreasing on this interval and since the sign or an eventual zero of p is known at x_0, we see that the sign-sequence of p is determined on $(x_, x_1)$. Arriving at x_1 we use that the sign of $r(x_1)$ is known and hence the sign-sequence for p is determined on $(-\infty, x_1]$. One can continue in this way and conclude that SIGN(p) is determined on the whole line.

Remark. Theorem B.4.5 extends to arbitrary finite families of polynomials and leads to a proof of the fundamental result which asserts that the family of semi-algebraic sets is preserved under an arbitrary polynomial map from one euclidian space to another. This theorem is due to Tarski and Seidenberg and has a wide range of applications in PDE-theory and is also used to establish the existence of various asymptotic expansions. In addition to Seidenberg's article [Seidenberg 1954] we refer to the Appendix in [Hörmander: XX] for a further account about semi-algebraic sets under polynomial maps from one euclidian space to another.

B.5 Absolute values of complex roots.

The material below stems from the article *Recherches sur la méthode de Graeffe et les zeros des polynomes et des series de Laurent* by Ostrowski which covers 150 pages in vol. 72 in Acta Mathematica (1940) and contains many interesting results. Let $n \geq 2$ and

$$(*) \qquad p(z) = a_0 + a_1 z + \ldots + a_n z^n$$

is a polynomial whose coefficients are complex numbers where a_0 and a_n both are $\neq 0$. The roots are arranged with non-decreasing absolute values, i.e. $\{0 < |\zeta_1| \leq |\zeta_2| \leq \cdots\}$.

Newton's diagram. Let $p(z)$ be given by (*) and in the (x, y)-plane one associates points as follows: For each $0 \leq \nu \leq n$ such that $a_\nu \neq 0$ we put:

$$\xi_\nu = (\nu, \log \frac{1}{|a_\nu|})$$

Starting from ξ_0 we find the unique piecewise linear convex curve ℓ_* which joins ξ_0 with ξ_n and stays below all the ξ-points. We refer to ℓ_* as Newton's minorizing convex curve.

Remark. The reader should illustrate the construction of ℓ_* by a figure. Consider for example the case when $a_0 = 1$ which gives $\xi_0 = (0, 0)$ and suppose that $|a_n| < 1$ so that $\xi_n = (n, \eta_n)$ where

$$\eta_n = \log \frac{1}{|a_n|} > 0$$

If

$$\min_{1 \leq k \leq n} \frac{\eta_k}{k} = \eta_n$$

then ℓ_* is the line from ξ_0 to ξ_n. If strict inequality occurs we find the smalleat integer $1 \leq k_* < n$ where the minimum is attained and then ξ_k gives the first corner point to the right of ξ_0 on the convex ℓ_*-curve.

B.5.0. The numerical inclination numbers. To each integer $0 \leq \nu \leq n$ we denote by $\{\chi_\nu\}$ the y-coordinates of the points on ℓ_* whose x-coordinate is ν. This means that

(i) $$\chi_\nu \leq \log \frac{1}{|a_\nu|} \implies |a_\nu| \leq e^{-\chi_\nu} \quad \text{for every} \quad \nu$$

Set $T_\nu = e^{-\chi_\nu}$ and notice that (i) gives the inequality

(ii) $$|a_\nu| \leq T_\nu$$

The numerical inclination number at place ν is defined by

(*) $$R_\nu = \frac{T_{\nu-1}}{T_\nu}$$

These constructions give for each $\nu \geq 1$:

(iii) $$\frac{T_\nu}{T_0} = \frac{1}{R_1 \cdots R_\nu} \implies |a_\nu| \leq \frac{|a_0|}{R_1 \cdots R_\nu}$$

where we used (i) and the observation that $T_0 = |a_0|$ since we start with a corner point on ℓ_* when $x = 0$.

1. Exercise. Suppose there exists some $1 \leq k \leq n$ such that $(k-1, 0)$ and $(k, 0)$ both belong to ℓ_*. Show with the aid of the figure that this implies that $|a_\nu| \leq 1$ for all other ν and there exists some $0 \leq k_* \leq k-1$ such that $|a_{k_*}| = 1$ and $(k_*, 0)$ is a corner point of ℓ_* while $\nu < k_*$ entails that $|a_\nu| < 1$.

2. Exercise. Show that if one starts from an arbitrary polynomial $p(z)$ where a_0 and a_n both are $\neq 0$, then there exist positive numbers B and b such that an integer k as in Exercise 1 exists for the scaled polynomial

$$q(z) = B \cdot p(bz)$$

Now we establish a result from Ostrowski's article [ibid].

B.5.1 Theorem. *Let p be a polynomial of degree $n \geq 2$ where $p(0) \neq 0$. Then*

$$\frac{|\zeta_k|}{R_k} \geq 1 - 2^{-1/k} \quad \text{hold for each} \quad 1 \leq k \leq n$$

Proof. By scaling we may assume that the situation in Exercise 1 occurs so now $|a_\nu| \leq 1$ hold for every ν and there exists some $k_* < k$ such that $|a_{k_*}| = 1$ and there remains to prove the inequality

(*) $$|\zeta_k| \geq 1 - 2^{-1/k}$$

To show (*) we may assume from the start that $|\zeta_k| < 1$. Put

(1) $$F(z) = \frac{z^k}{(z - \zeta_1) \cdot (z - \zeta_k)} = \frac{1}{(1 - \zeta_1/z) \cdots 1 - \zeta_k/z)}$$

The last expression gives a Lauren series expansion

$$F(z) = 1 + \sum_{\nu=1}^{\infty} \sigma_\nu \cdot z^{-\nu} \tag{2}$$

Next, put

$$F^*(z) = \frac{1}{(1 - |\zeta_1|/z) \cdots 1 - |\zeta_k|/z)} = 1 + \sum_{\nu=1}^{\infty} \sigma_\nu^* \cdot z^{-\nu}$$

It is clear that one has the majorisations $|\sigma_\nu| \leq \sigma_\nu^*$ and taking the sum over all $\nu \geq 1$ we get

$$\sum |\sigma_\nu| \leq \sigma_\nu^* = F^*(1) - 1 = \frac{1}{(1 - |\zeta_1|) \cdots 1 - |\zeta_k|)} - 1 \tag{3}$$

The first expression of F in (1) shows that $p \cdot F$ is analytic and of the form $p \cdot F = z^k \cdot G(z)$ where G is analytic. Hence the coefficient of z^{k_*} is zero which entails that

$$a_{k_*} + \sum_{\nu \geq 1} a_{k_*+\nu} \cdot \sigma_\nu = 0$$

Since $|\alpha_\nu| \leq 1$ hold for all ν and $|a_{k_*}| = 1$ the triangle inequality gives $\sum_{\nu \geq 1} |\sigma_\nu|$. Together with (3) we obtain

$$2 \leq \frac{1}{(1 - |\zeta_1|) \cdots 1 - |\zeta_k|)} \implies (1 - |\zeta_1|) \cdots (1 - |\zeta_k|) \leq \frac{1}{2}$$

Finally we have $|\zeta_\nu| \leq |\zeta_k|$ when $\nu < k$ which gives:

$$(1 - |\zeta_k|)^k \leq \frac{1}{2} \implies |\zeta_k| \geq 1 - 2^{-k}$$

The next result is attributed to Polya in Ostrowski's article.

B.5.3 Theorem. *Let p be a polynomial of degree n as in (*). Then the following hold for each* $1 \leq k \leq n$

$$\frac{R_1 \cdot R_k}{|\zeta_1 \cdots \zeta_k|} \leq \sqrt{(k+1) \cdot (1 + \frac{1}{k})^k} \leq \sqrt{(k+1)e}$$

Proof. Write $p(z) = \sum a_\nu z^\nu$. Landau's inequality from Exercise §§ in chapter III gives the following inequality for every $r > 0$ and $k \geq 1$:

$$\frac{r^k}{|\zeta_1 \cdots \zeta_k|} \leq \frac{1}{|a_0|} \cdot \sqrt{\sum |a_\nu|^2 \cdot r^{2\nu}} \tag{*}$$

Next, the inequality (ii) from B.5.0 gives:

$$|a_\nu| \leq |a_0| \cdot \frac{T_\nu}{T_0} = \frac{|a_0|}{R_1 \cdots R_\nu} \quad : \quad \nu \geq 1 \tag{1}$$

Taking the square in Landau's inequality (1) gives

$$\frac{r^{2k}}{|\zeta_1 \cdots \zeta_k|^2} \leq 1 + \sum_{\nu \geq 1} \frac{r^2}{R_1^2} \cdots \frac{r^2}{R_\nu^2} \tag{2}$$

Keeping k fixed we set $\theta = \sqrt{\frac{k}{k+1}}$ and $r = \theta \cdot R_k$. Then (2) gives

$$\frac{R_k^{2k}}{|\zeta_1 \cdots \zeta_k|^2} \cdot \theta^{2k} \leq 1 + \sum_{\nu \geq 1} \theta^{2\nu} \cdot \frac{R_k^2}{R_1^2} \cdots \frac{R_k^2}{R_\nu^2} \implies \tag{3}$$

$$(\frac{R_1 \cdots R_k}{|\zeta_1 \cdots \zeta_k|})^2 \leq \theta^{-2k} \cdot [\sum_{\nu=0}^{\nu=k-1} \theta^{2\nu} \cdot (\frac{R_{\nu+1}}{R_k} \cdots \frac{R_{k-1}}{R_k})^2 + \theta^{2k} + \sum_{\nu > k} \theta^{2\nu} \cdot (\frac{R_k}{R_{k+1}} \cdots \frac{R_k}{R_\nu})^2] \tag{4}$$

Finally, since the sequence $\{R_\nu\}$ is increasing the right hand side in (4) is majorized by

$$\theta^{-2k} \cdot \sum_{\nu \geq 0} \theta^{2\nu} = \theta^{-2k} \cdot \frac{1}{1-\theta^2} = (k+1) \cdot (1+\frac{1}{k})^k$$

Taking square roots Polya's inequality follows.

1:C. Interpolation formulas

Consider a monic polynomial of degree k:

$$P(z) = z^k + c_{k-1}z^{k-1} + \ldots + c_1z + c_0$$

Let $\alpha_1, \ldots, \alpha_k$ be the roots where multiple zeros may occur. In contrast to sets of real numbers there is no procedure to order a set of complex numbers. Thus, the the roots should be regarded as an *unordered* k-tuple of complex numbers. But there exist *symmetric polynomials* of this unordered k-tuple. In particular we obtain the symmetric sums

(i) $$\sigma_m = \alpha_1^m + \ldots + \alpha_k^m \quad : 1 \leq m \leq k$$

C.1 Theorem. *For each $m \geq 1$ there exists a polynomial $Q_m(c_0, \ldots, c_{k-1})$ of the independent c-variables such that*

$$\sigma_m = Q_m(c_0, \ldots, c_{k-1})$$

Exercise. Residue calculus can be used to find the σ-numbers. First Euclidian division gives a unique pair of polynomials $A_m(z)$ and $\Gamma_m(z)$ such that

(i) $$z^m \cdot P'(z) = A_m(z) \cdot P(z) + \Gamma_m(z)$$

where Γ has degree $\leq k-1$. Residue calculus gives

$$\sigma_m = \frac{1}{2\pi i} \cdot \int_{|z|=R} \frac{z^m \cdot P'(z)\, dz}{P(z)} = \frac{1}{2\pi i} \cdot \int_{|z|=R} \frac{\Gamma_m(z))\, dz}{P(z)} = \gamma_{k-1}(m)$$

where $\gamma_{k-1}(m)$ is the coefficient of z^{k-1} in Γ_m. Assume that $k \geq 2$ and let $\gamma_{k-2}(m)$ be the coefficient of z^{k-2} in Γ_m. Now the reader can verify the recursion formula

$$\gamma_{k-1}(m+1) = c_{k-1}\gamma_{k-1}(m) + \gamma_{k-2}(m)$$

From this an obvious induction shows that σ_m is a polynomial in $c_0, \ldots, c_{k-1}$ with integer coefficients. The reader may consult some text-book in algebra for the explicit expression of these polynomials. We have for example

$$\sigma_2 = c_{k-1}^2 - 2 \cdot c_{k-2}$$

C.2 The discriminant. It is defined by:

(*) $$\mathfrak{D}_P = \prod_{i \neq \nu} (\alpha_i - \alpha_\nu)$$

In the product appears $k(k-1)/2$ many terms. Since the k-tuple of roots appear in a symmetric fashion Theorem C.1 gives a polynomial $Q^*(c_0, \ldots, c_{k-1})$ such that

(iii) $$\mathfrak{D}_P = Q^*(c_0, \ldots, c_{k-1})$$

The reader may consult a text-book in algebra for the expression of the Q^*-polynomial.

Example If $k = 2$ we have $P(z) = z^2 + c_1z + c_0$ and if α_1, α_2 are the roots, it follows that

$$\mathfrak{D}_P = -(\alpha_1 - \alpha_2)^2 = 2\alpha_1 \cdot \alpha_2 + (c_1\alpha_1 + c_0 + c_1\alpha_2 + c_0) = 4c_0 - c_1^2$$

Every k-tuple $(c_0, \ldots, c_{k-1})$ for which $Q^*(c_0, \ldots, c_{k-1}) \neq 0$ gives a polynomial with *distinct* roots. Since the Q^*-polynomial is not identically zero, it follows that the *generic* polynomial $P(z)$ of degree k has simple roots. The exception occurs when the point $(c_0, \ldots, c_{k-1})$ in $\mathbf{C}^k$ belongs to the algebraic hypersurface $\{Q^* = 0\}$ where Q^* is regarded as a complex-valued function of the k many complex variables $c_0, \ldots, c_{k-1}$. The detailed study of this algebraic hypersurface is a topic in algebraic geometry. Using euclidian divisions in the polynomial ring $\mathbf{C}[z]$ one can find an expression for Q^*. This is explained below.

C.3 The polynomial ring $\mathbf{C}[z]$. Given a monic polynomial $P(z)$ of some degree $k \geq 2$ its derivative $P'(z)$ is a polynomial of degree $k-1$. The condition that the roots of $P(z)$ are simple means that P and P' have no root in common. When this holds euclidian division gives a unique pair of polynomials $A(z)$ and $B(z)$ such that

$$(*) \qquad A(z)P(z) + B(z)P'(z) = 1 \quad : \quad \deg(B) \leq k-1$$

Above $B(z)$ is the unique polynomial of degree $k-1$ such that

$$B(\alpha_\nu) = \frac{1}{P'(\alpha_\nu)} \quad : \quad \alpha_1, \ldots, \alpha_k \text{ are the distinct roots of } P(z)$$

Exercise. Verify the formula below for $B(z)$ which already appeared in Newton's text-books in algebra and analysis from 1666:

$$B(z) = \sum_{\nu=1}^{\nu=k} \frac{1}{\prod_{i \neq \nu} (\alpha_\nu - \alpha_i) \cdot P'(\alpha_\nu)} \cdot \frac{P(z)}{z - \alpha_\nu}$$

C.4 Conditions for simple roots. Let $P(z)$ be a monic polynomial of degree $k \geq 2$ and assume that it has simple zeros so that the equation (*) above can be solved.

Exercise. Show that to every integer $0 \leq \nu \leq 2k-2$ there exists a unique pair of polynomials $A_\nu(z)$ of degree $\leq k-2$ and $B_\nu(z)$ of degree $\leq k-1$ such that

$$\text{(i)} \qquad A_\nu(z)P(z) + B_\nu(z)P'(z) = z^\nu$$

Next, the vector space of all polynomials of degree $\leq 2k-2$ has dimension $2k-1$. This criterion and the calculus with determinants implies that the polynomial $P(z)$ has simple roots if and only if a certain determinant of an $2k-1 \times 2k-1$-matrix is non-zero. If $k=3$ the condition for simple roots is:

$$\det \begin{pmatrix} c_0 & 0 & c_1 & 0 & 0 \\ c_1 & c_0 & 2c_2 & c_1 & 0 \\ c_2 & c_1 & 3 & 2c_2 & c_1 \\ 1 & c_2 & 0 & 3 & 2c_2 \\ 0 & 1 & 0 & 0 & 3 \end{pmatrix} \neq 0$$

The reader is invited to find matrices for higher k-values.

C.5 Newton's interpolation. Let $k \geq 2$ and consider a pair of k-tuples $w_1, \ldots, w_k$ and $z_1, \ldots, z_k$. Assume that the z-numbers are distinct, i.e. $z_j \neq z_\nu$ hold when $j \neq \nu$. The w-numbers are arbitrary and it may even occur that all w-numbers are equal. Then there exists a *unique* polynomial $P(z)$ of degree $k-1$ at most such that

$$\text{(i)} \qquad P(z_\nu) = w_\nu \quad : 1 \leq \nu \leq k$$

One refers to P as Newton's interpolating polynomial. One has the formula:

$$\text{(ii)} \qquad P(z) = \sum_{j=1}^{j=k} w_j \cdot \frac{\prod_{\nu \neq j} (z - z_\nu)}{\prod_{\nu \neq j} (z_j - z_\nu)}$$

Another procedure is to seek a polynomial

$$Q(z) = c_{k-1} z^{k-1} + \ldots + c_0$$

where (i) gives a system of equations:

$$\text{(iii)} \qquad c_0 + c_1 z_j + \ldots + c_{k-1} z_j^{k-1} = w_j \quad : 1 \leq j \leq k$$

Since $z_1, \ldots, z_k$ are distinct the *van der Monde determinant* of the $k \times k$-matrix whose rows are $(1, z_j, \ldots, z_j^{k-1})$ is non-zero. Hence (iii) has a unique solution $(c_0, \ldots, c_{k-1})$. As already predicted by Newton's formula above it follows that when the k-tuple $z_1, \ldots, z_k$ is kept fixed, thn the c-numbers are linear functions of $w_1, \ldots, w_k$ whose coefficients depend on the k-tuple $\{z_j\}$. and for each $0 \leq \nu \leq k-1$ we can write

$$\text{(iv)} \qquad c_\nu = \sum_{j=1}^{i=k} G_{\nu,j}(z_1, \ldots, z_k) \cdot \omega_j$$

Remark. Above we treat $z_1, \ldots, z_k$ as independent complex variables. The G-functions have been determined under the assumption that the k-tuple is distinct, i.e. $z_i \neq z_\nu$ hold when $i \neq \nu$. At the same time we recall that $c_0, \ldots, c_{k-1}$ can be solved via Cramer's rule. From this it follows that every doubly indexed G-function is a rational function of the k-many variables $z_1, \ldots, z_k$.

C.6 Exercise. Newton's formula (i) from C.5 and residue calculus enable us to express the G-functions. For each $0 \leq \nu \leq k-1$ one has

$$c_\nu = \frac{1}{2\pi i} \cdot \int_{|z|=R} \frac{P(z)}{z^{\nu+1}}\, dz$$

It follows that

$$G_{\nu j}(z_1, \ldots, z_k) = \frac{1}{2\pi i} \cdot \frac{1}{\prod_{\nu \neq j}(z_j - z_\nu)} \cdot \cdot \int_{|z|=R} \frac{\prod_{\nu \neq j}(z - z_\nu)}{z^{\nu+1}}\, dz$$

From this the reader can deduce an explicit expression of the rational G-funtions.

C.7 A question. Let $k \geq 2$ and consider the family of k-tuples $z_1, \ldots, z_k$ for which $\sum |z_\nu|^2 = 1$. When the k-tuple is distinct we get the positive number

$$\delta(z_\bullet) = \min_{j \neq \nu} |z_j - z_\nu|$$

Let us then consider a polynomial $Q(z)$ of degree $\leq k-1$ with coefficients $c_0, \ldots, c_{k-1}$. Now the c-coefficients can be estimated by the maximum norm

$$|Q|_{z_\bullet} = \max_\nu |Q(z_\nu)|$$

Newton's interpolation formula gives for each $0 \leq \nu \leq k-1$ a constant $C_\nu(k)$ which is independent of Q such that

$$(*) \qquad |c_\nu| \leq C_\nu(k) \cdot \delta(z_\bullet)^{-k+1} \cdot |Q|_{z_\bullet}|$$

The reader is invited to analyze the behavious or the numbers $C_\nu(k)$ as k increases.

1:D. Tchebyscheff polynomials and transfinite diameters

The construction of Tchebysheff numbers attached to arbitrary compact subsets in $\mathbf{C}$ is due to Faber whose article [Faber: 1920] treats various extremal problems in the complex domain. Let $N \geq 2$ and $E = (z_1, \ldots, z_N)$ an N-tuple of distinct complex numbers. For each integer $n \geq 0$ we denote by $\mathcal{P}(n)$ the set of polynomials of degree $\leq n$. If $p(z) \in \mathcal{P}(n)$ we define the maximum norm

$$|p|_E = \max_k |p(z_k)|$$

If $n \leq N-1$ the maximum norm must be positive since p has at most n distinct zeros and therefore cannot vanish on the N-tuple of points in E. Put

$$\mathfrak{Tch}_E(n) = \min_{q \in \mathcal{P}(n-1)} |z^n + q(z)|_E$$

D.1 Proposition. *For each $n \leq N-1$ there exists a unique $q_* \in \mathcal{P}(n-1)$ such that*

(*) $$\mathfrak{Tch}_E(n) = |z^n + q_*(z)|_E$$

Proof. A polynomial $q \in \mathcal{P}(n-1)$ is said to be extremal if equality holds in (*). By (*) in C.7 there is a uniform upper bound for the coefficients of competing extremal polynomials and since bounded sets of complex numbers are relatively compact there exists at least one extremal polynomial q. To show that q is unique. we t consider the set of points $z_k \in E$ such that

(i) $$\mathfrak{Tch}_E(n) = |z_k^n + q(z_k)|$$

Let $\mathcal{E}_q^*$ denote this subset of E. Suppose that $\mathcal{E}_q^*$ consists of $\leq n-1$ many points, say $z_1, \ldots, z_m$ for some $m \leq n-1$. Then we can find $\phi \in \mathcal{P}(n-1)$ such that

$$\phi(z_k) = z_k^n + q(z_k) \quad : \quad 1 \leq k \leq m$$

Now the reader can verify that if $\epsilon > 0$ is sufficiently small, then

$$|z^n + q(z) - \epsilon \cdot \phi(z)|_E = (1-\epsilon) \cdot \mathfrak{Tch}_E(n)$$

which cannot occur since q was extremal. So if q is extremal then $\mathcal{E}_q^*$ contains at least n many points. Suppose now that q_1 and q_2 are two extremal polynomials and set $q = \frac{1}{2}(q_1 + q_2)$ which gives

$$z^n + q = \frac{1}{2}(z^n + q_1) + \frac{1}{2}(z^n + q_2)$$

The triangle inequality for the maximum norm over E entails that q also is extremal and by the above $\mathcal{E}_q^*$ contains at least n points $z_1, \ldots, z_n$. Now

(ii) $$\mathfrak{Tch}_E(n) = |z_k^n + \frac{1}{2}(q_1(z_k) + q_2(z_k)|$$

Since q_1 and q_2 are extremal we also have

(iii) $$|z_k^n + q_\nu(z_k)| \leq \mathfrak{Tch}_E(n) \quad : \quad \nu = 1, 2$$

It follows from (ii-iiii) that we must have the equality $q_1(z_k) = q_2(z_k)$ for each k. Hence the polynomial $q_1 - q_2$ has at least n zeros. This can only can occur if they are identical which finishes the proof of uniqueness.

D.2 The case when E is infinite. Let E be an infinite compact set in $\mathbf{C}$. Let $\{z_\nu\}$ be a denumerable dense subset in E and for each N we put $E_N = \{z_1, \ldots, z_N\}$. Next, fix some positive integer n. Proposition D.1 gives for each $N \geq n+1$ a unique extremal $q_N \in \mathcal{P}(n-1)$ such that

(i) $$\mathfrak{Tch}_{E_N}(n) = |z^n + q_N(z)|_{E_N}$$

It is clear that

(ii) $$N \mapsto \mathfrak{Tch}_{E_N}(n)$$

increases with N. Since we can take q_N as the zero polynomial for every N one has the inequality

$$\mathfrak{Tch}_{E_N}(n) \leq \max_{z \in E} |z|^n$$

where the right hand side is finite because E is compact. Hence (ii) is bounded above and there exists a limit

(iii)
$$\lim_{N \to \infty} \mathfrak{Tch}_{E_N}(n)$$

At the same time we have the sequence $\{q_N\}$ in $\mathcal{P}(n-1)$. For each N we write

$$q_N(z) = c_0(N) + c_N(1) \cdot z + \ldots + c_N(n-1) \cdot z^{n-1}$$

From (C.7) the reader may verify that there is a constant M such that

$$\sum_{\nu=0}^{\nu=n=1} |c_N(\nu)| \leq M \quad : \ N \geq n+1$$

Exercise. Use the above to show that there always exist a subsequence $N_1 < N_2 < \ldots$ such that

$$\lim_{j \to \infty} c_{N_j}(\nu) = c_*(\nu) \quad : \quad 1 \leq \nu \leq n-1$$

From this extracted subsequence we obtain the polynomial

$$q_*(z) = \sum_{\nu=0}^{\nu=n-1} c_*(\nu) \cdot z^\nu$$

Use Proposition D.1 to show that this limit polynomial is the same for any chosen subsequence $\{N_k\}$ so there exist unrestricted limits

$$\lim_{j \to \infty} c_N(\nu) = c_*(\nu) \quad : \quad 1 \leq \nu \leq n-1$$

Finally, show that q_* is the unique extremal polynomial for which one has the equality

$$|z^n + q_*(z)|_E = \min_{q \in \mathcal{P}(n-1)} |z^n + q(z)|_E$$

D.3 Tchebyscheff norms. With q_* as the unique extremal in $\mathcal{P}(n-1)$ above we set

$$T_n^E(z) = z^n + q_*(z)$$

and refer to this monic polynomial as the Tchebyscheff polynomial of degree n attached to the compact set E. The Tchebyscheff norm of order n over E is defined by:

$$\mathfrak{Tch}_E(n) = |T_n^E|_E$$

For each $n \geq 1$ we put

$$\rho(n) = \log \mathfrak{Tch}_E(n)$$

We leave it as an exercise to verify that the function $n \mapsto \frac{\rho(n)}{n}$ is convex, i.e. that

$$\rho(n+m) \leq \frac{m}{n+m} \cdot \rho(m) + \frac{n}{n+m} \cdot \rho(n)$$

holds for each pair $m, n \geq 1$. The hint is to first verify via linear algebra that for every polynomial $q \in \mathcal{P}(n+m-1)$ there exist $q_1 \in \mathcal{P}(m-1)$ and $q_2 \in \mathcal{P}(n-1)$ such that

$$z^{n+m} + q(z) = (z^m + q_1(z))(z^n + q_2(z))$$

D.4 The Tchebyscheff diameter. The convexity entails by a general result about non-decreasing sequences of real numbers which is bounded above, that there exists the limit

$$\lim_{n \to \infty} \frac{\log \mathfrak{Tch}_E(n)}{n}$$

Passing to exponential functions we get the limit number

$$\mathfrak{DTch}(E) = \lim_{n\to\infty} \left[\mathfrak{Tch}_E(n)\right]^{\frac{1}{n}}$$

We refer to this number as the Tchebyscheff diameter of the compact set E.

D. 5 The transfinite diameter and Szegö's theorem

To each n-tuple of distinct points $z_1, \ldots, z_n$ in $\mathbf{C}$ we set

$$L_n(z_\bullet) = \frac{1}{n(n-1)} \cdot \sum_{k\neq j} \log \frac{1}{|z_j - z_k|}$$

If E is a compact and inifinite set we put

$$\mathcal{L}_n(E) = \min\, L_n(z_\bullet)$$

where the minimum is taken over all n-tuples in E. Since $\log \frac{1}{r}$ is large when $r \simeq 0$ this means that one tries to choose separated n-tuples in order to minimize the L_n-function. For example, when $n = 2$ the minimum is achieved for a pair of points in E whose distance is maximal, i.e. $\mathcal{L}_2$ is the usual diameter of E.

D.6 Proposition. *The sequence $\mathcal{L}_n\}$ is non-decreasing.*

Proof. Let $z_1^*, \ldots, z_{n+1}^*$ minimize the L_{n+1}-function. We get

$$\mathcal{L}_{n+1}(E) = \frac{1}{n(n+1)} \cdot \sum_{k\neq\nu}^{(1)} \log \frac{1}{|z_\nu^* - z_k^*|} + \frac{2}{n(n+1)} \cdot \sum_{k=2}^{k=n+1} \log \frac{1}{|z_1^* - z_k^*|}$$

where (1) above the first summation means that $2 \leq \nu \neq k$ holds. Here $z_2^*, \ldots, z_{n+1}^*$ is an n-tuple competing to maximize $\mathcal{L}_E(n)$ which gives the inequality:

$$\mathcal{L}_{n+1}(E) \geq \frac{1}{n(n+1)} \cdot n(n-1) \cdot \mathcal{L}_n(E) + \frac{2}{n(n+1)} \cdot \sum_{k=2}^{k=n+1} \log \frac{1}{z_1^* - z_k^*|}$$

The same inequaliiy holds when when we instead of z_1 delete some z_j for $2 \leq j \leq n+1$. Taking the sum of the resulting inequalities we obtain

$$(n+1)\mathcal{L}_{n+1}(E) \geq \frac{1}{n} \cdot n(n-1) \cdot \mathcal{L}_n(E) + \frac{2}{n(n+1)} \cdot \sum_{k\neq j} \log \frac{1}{|z_j - z_k|}$$

The last term is $2 \cdot \mathcal{L}_{n+1}$ which gives:

$$(n-1) \cdot \mathcal{L}_{n+1}(E) \geq \frac{1}{n} \cdot n(n-1) \cdot \mathcal{L}_n(E) = (n-1)\mathcal{L}_n(E)$$

A division by $n-1$ gives the requested inequality.

D.7 Definition. *The limit number defined by*

$$\mathfrak{D}(E) = \lim_{n\to\infty} e^{-\mathcal{L}_n(E)}$$

is called the transfinite diameter of E.

Remark. The definition means that $\mathfrak{D}(E) = 0$ if and only if $\mathcal{L}_n(E)$ tends to $+\infty$ as n increases. Intuitively this means that we are not able to choose large tuples in E separated enough to keep the sum of the log-terms bounded.

D.8 Example. Consider the interval $E = [-1, 1]$. Using the concavity of the log-function one shows that $L_n(z_\bullet)$ is minimized when $\{z_\nu\}$ are equi-distributed and from this a passage to the limit gives:

$$(*) \qquad \lim_{n\to\infty} \mathcal{L}_n(E) = \iint \log \frac{1}{|s-t|} \cdot ds dt$$

with the double integral taken over the square $-1 \leq s,t \leq 1$. Indeed, the reader may verify this by approximating the double integral by Riemann sums.

Exercise. Show that the double integral has the value $\log 2$ and conclude that

$$\mathfrak{D}([0,1]) = \frac{1}{2}$$

With $E = [0,1]$ original work by Tchebyscheff gives the unique extremal polynomials which determine $L_n(E)$ for each $n \geq 1$. More precisely the Tchebyscheff polynomials described in XX give the equalities:

$$\mathfrak{Tch}_E(n) = 2^{-n+1}$$

Passing to the limit and taking the n:th root we get

$$\mathfrak{DTch}_E = \frac{1}{2}$$

and hence this number is equal to $\mathfrak{D}(E)$. It turns out that this equality holds in general.

D.9 Theorem. *For every compact set E in $\mathbf{C}$ one has the equality*

$$\mathfrak{D}(E) = \mathfrak{DTch}(E)$$

Remark. Theorem D.9 is due to Szegö in [Szegö: 1924 Bemerkungen]. We give further comments on this result in § XX in Special Tpoics.

1:E. Further results and exercises.

E.1 Lagrange's identity. Let $n \geq 2$ and $z_1, \ldots, z_n$ and $w_1, \ldots, w_n$ are two n-tuples of complex numbers. Show that

$$\Big|\sum_{j=1}^{j=n} z_j w_j\Big|^2 = \sum_{j=1}^{j=n} |z_j|^2 \cdot \sum_{j=1}^{j=n} |w_j|^2 - \sum_{1 \leq j < k \leq n} |z_j \bar{w}_k - \bar{w}_j z_k|^2$$

Conclude that one has the inequality

$$\Big|\sum_{j=1}^{j=n} z_j w_j\Big| \leq \sqrt{\sum_{j=1}^{j=n} |z_j|^2 \cdot \sum_{j=1}^{j=n} |w_j|^2}$$

where equality holds if and only if here exists a complex number λ such that

$$w_\nu = \lambda \cdot z_\nu \quad : \quad 1 \leq \nu \leq n$$

E.2 Zeros of derivatives. Let $P(z)$ be a polynomial of degree n whose zeros are $\alpha_1, \ldots, \alpha_n$ where eventual multiple zeros are repeated. So in general the number of distinct zeros may be $< n$.

E.3 Theorem. *Let K be a convex set in $\mathbf{C}$ which contains all zeros of P. Then the zeros of P' belong to K.*

Proof. Suppose first that all zeros of P have real part ≤ 0. Newton's formula gives:

(i) $$\frac{P'(z)}{P(z)} = \sum \frac{1}{z - \alpha_\nu}$$

If $\mathfrak{Re}(z) = a > 0$ we get

$$\mathfrak{Re}\,\frac{P'(z)}{P(z)} = \sum \frac{a - \mathfrak{Re}(\alpha_\nu)}{|z - \alpha_\nu|^2}$$

By assumption $\mathfrak{Re}(\alpha_\nu) \leq 0$ for every ν so the right hand side is a sum of positive terms which entails that the derivative P' has no zeros in the open right half-plane $\mathfrak{Re}(z) > 0$. Theorem E.3 follows from this special case since every a convex set is the intersection of half-planes.

E.4 Kakeya's theorem. *Let $c_n > c_{n-1} > \ldots > c_0 > 0$ be a strictly decreasing sequence of positive real numbers. Then the zeros of the polynomial*

$$P(z) = c_n z^n + \ldots + c_0$$

all have absolute value < 1.

Proof. We have the equation

(i) $(z-1)P(z) = c_n \cdot z^{n+1} - \big[(c_n - c_{n-1}) \cdot z^n + (c_{n-1} - c_{n-2}) \cdot z^{n-1} + + \ldots + (c_1 - c_0) \cdot z + c_0\,\big]$

If $|z| \geq 1$ the absolute value of the last term above is majorized by

$$|z|^n \cdot \big[(c_n - c_{n-1} + (c_{n-1} - c_{n-2}) + \ldots + (c - 1c_0) + c_0\big] = c_n \cdot |z|^n$$

Then it is clear that (i) cannot hold if $|z| > 1$. Next, if $P(e^{i\theta}) = 0$ for some θ we set $a_k = c_k - c_{k-1}$ if $1 \leq k \leq n$ while $a_0 = c_0$. Then

$$\big|a_n \cdot e^{in\theta} + \ldots + a_1 e^{i\theta} + a_0\big| = c_n$$

where c_n is the sum of the positive a-numbers. This can only occur if $e^{i\theta} = 1$ and since it is obvious that $P(1) \neq 0$ Kakeya's result follows.

E.5 Exercise. Let $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ be a polynomial with real coefficients whose zeros belong to the left half-plane $\mathfrak{Re}\, z < 0$. Show that all a-coefficients are > 0. The hint is that the fundamental theorem of algebra gives a factorisation

$$P(z) = \prod (z + q_j) \cdot \prod (x + \alpha_\nu)^2 + \beta_\nu)$$

where $\{-q_j\}$ are the strictly negative real roots while the second product takes into the account that complex roots appear in conjugate pairs and by the hypothesis each α_ν above is real and > 0.

E.6 A second order differential equation Let $n \geq 2$. Apply linear algebra to the $(n+1)$-dimensional real vector space of polynomials with real coefficients to show that there exists a unique monic polynomial $p(z)$ of degree n with real coefficients which satisfies the second order differential equation

$$(*) \qquad (z^2-1)p''(z) + 2zp'(z) = n(n+1)p(z)$$

It turns out that the zeros of p are real. To prove this we argue by contradiction. If not all the zeros are real we find a zero z_0 where $\mathfrak{Im}\, z_0 > 0$ and $\mathfrak{Im}\, \beta \leq \mathfrak{Im}\, z_0$ for every other root of p. Consider the factorisation

$$(1) \qquad p(z) = (z-z_0)q(z)$$

where the polynomial $q(z)$ has degree $n-1$. Notice that

$$p'(z) = q(z) + (z-z_0)q'(z) \quad \text{and} \quad p''(z) = 2q'(z) + (z-z_0)q'(z) \implies$$
$$p'(z_0) = q(z_0) \quad \text{and} \quad p''(z_0) = 2q'(z_0)$$

So with $z = z_0$ in the differential equation we obtain

$$\text{(i)} \qquad (z_0^2-1)2q'(z_0) + 2z_0 \cdot q(z_0) = 0 \implies \frac{q'(z_0)}{q(z_0)} = \frac{z_0}{1-z_0^2}$$

Let $\beta_1, \ldots, \beta_{n-1}$ be the zeros of q. Then (i) and Newton's formula give

$$\text{(ii)} \qquad \frac{z_0}{1-z_0^2} = \frac{q'(z_0)}{q(z_0)} = \sum \frac{1}{z_0-\beta_\nu} = \sum \frac{\bar{z}_0 - \bar{\beta}_\nu}{|z_0-\beta_\nu|^2}$$

The choice of z_0 entails that $\mathfrak{Im}(\beta_\nu) \leq \mathfrak{Im}(z_0)$ hold for each ν which implies that the imaginary part in the right hand side is ≤ 0. This gives a contradiction since we the imaginary part in the left hand side was assumed to be > 0.

E.6.1 Exercise. Use the reality of the roots to show that they are all bounded and stay in the interval $[-1, 1]$. The hint is to use Rolle's mean-value theorem, or rather the classic rule of Descartes for real zeros of polynomials with real coefficients.

E.6.2 Asymptotic distribution of zeros. When n increases we study how the zeros of the polynomial of degree n which solves the differential equation above are distributed on $[-1, 1]$. Let $\alpha_1, \ldots, \alpha_n$ be the zeros of the polynomial p_n and set

$$(1) \qquad \phi_n(z) = \frac{1}{n} \cdot \sum \frac{1}{z-\alpha_\nu}$$

This is the Cauchy transform of the probability measure μ_n on $[-1, 1]$ which assign the mass $\frac{1}{n}$ and each α_ν. Newtons formula gives:

$$n \cdot \phi_n = \frac{p_n'}{p_n}$$

Taking the complex derivative on both sides it follows that

$$(1) \qquad n \cdot \phi_n' = \frac{p_n''}{p_n} - n^2 \cdot \phi_n^2$$

At the same time the differential equation satisfied by p_n entails that

$$(z^2-1)\frac{p_n''}{n(n+1)p_n} + 2z\frac{p'}{n(n+1) \cdot p} = 1 \implies$$

$$(2) \qquad (z^2-1) \cdot \frac{n^2}{n(n+1)} \cdot \phi_n^2 + (z^2-1) \cdot \frac{n}{n(n+1)} \cdot \phi_n' + z\frac{n}{n(n+1)}\phi_n = 1$$

Passage to a limit. The analytic functions $\{\phi_n(z)\}$ are defined outside the compact interval $[0,1]$ and from (1 it is easily seen that they form a normal sequence in $\mathbf{C}\setminus[0,1]$ in the sense of Montel. See § XX in chapter III. Passing to the limit one gets:

$$\text{(3)} \qquad \lim_{n\to\infty}\phi_n^2(z)=\frac{1}{z^2-1}$$

where the convergence holds uniformly over compact subsets of $\mathbf{C}\setminus[0,1]$. In Chapter IV we shall learn that $\sqrt{1-z^{-2}}$ is a well-defined analytic function outside $[0,1]$ and there exists a unique probability measure μ_* supported by $[0,1]$ such that

$$(*) \qquad \int_{-1}^{1}\frac{d\mu_*(s)}{z-s}=\frac{1}{z}\cdot\frac{1}{\sqrt{(1-z^{-2}}} \quad : \quad z\in\mathbf{C}\setminus[0,1]$$

From this one can conclude that the sequence $\{\mu_n\}$ converges weakly to μ_*, i.e there exists an asymptotic limit distribution for the zeros of the eigenpolynomials $\{p_n\}$. There remains to find μ_* which amounts to establish an inversion formula (*). We shall treat this in §§ XX and remark only that in the case above μ_* is the density function on the s-interval given by

$$s\mapsto XXX$$

E.6.3 Another example. Here we consider eigenpolynomials $\{p_n\}$ which solve the differential equation

$$xp_n''+xp_n'=np_n$$

This time the zeros are no longer bounded. Instead one employs the Bergquist scaling from [Bergquist] and if $\{\alpha_1,\ldots,\alpha_n\}$ are the roos of p_n we set

$$\phi_n(z)=\frac{1}{n}\sum\frac{1}{z-\frac{\alpha_\nu}{n}}$$

By a similar limit process as above one shows that the sequence $\{\phi_n(z)\}$ converges to the limit function $\phi(z)$ which satisfies the algebraic equation

$$\phi(z)^2+\phi(z)=\frac{1}{z}$$

From this one can deduce that the scaled sequences $\{\frac{\alpha_\nu}{n}\}$ become asymptotically distributed on the real interval $[-4,0]$ whose the density function $\rho(t)$ satisfies the equation

$$\int_{-4}^{0}\frac{\rho(t)\,dt}{z-t}=\frac{\sqrt{4+z}}{\sqrt{z}}$$

In § xx we shall learn hos one derives the real-valued and non-negative ρ-function from this equation. The examples above illustrates the interplay between zeros of eigenpolynomials to ordinary differential equations with polynomial coefficients and complex analytic formulas. It would bring us too far to discuss this further but the interested reader may consult the Pd. thesis [Bergquist] for further examples and results dealing with such asymptotic formulas for zeros of eigenpolynomials. We remark that many open problems remain to be settled which involve difficult questions about roots of algebraic equations which go beyond classic results about algebraic curves because one seeks analytic formulas whose existence require calculus of variation and cannot be handled by standard algebraic or geometric methods.

E.7 Fejer's orthogonal polynomials.

Let μ be a compactly supported probability measure in $\mathbf{C}$. In the appendix Measure we shall learn how to construct integrals with respect to such a measure which in general is a Riesz measure which can be singular, i.e. the support of μ has 2-dimensionmal Lebesgue measure zero. In particular there exists the Hilbert space $L^2(\mu)$ if complex valued and square integrable functions with respect to μ. Under the sole assumption that the support of μ is not confined to a finite set, the Gram-Schmidt process gives a unique infinite sequence of polynomials $\{p_k(z)\}$ where $p_k(z)$ has degree k whose leading coefficient for z^k is real and positive and

$$\int p_k(z)\cdot \bar{p}_m(z)\,d\mu(z) = \text{Kronecker's delta-function}$$

Here $p_0(z)=1$ and $p_1(z)=az+b$ where the constants a,b satisfy

$$\int |az+b|^2 d\mu(z) = 1 \quad \text{and} \quad \int (az+b)d\mu(z)=0$$

Let $n\geq 1$ and consider the polynomial $p_n(z)$. If α is a zero of p_n we can write

$$p_n(z)=(z-\alpha)q(z) \quad \text{where} \quad \deg(q)\leq n-1$$

Orthogonality entails that

$$(1) \qquad 0=\int p(z)q(z)\cdot \bar{q}(z)d\mu(z) = \int (z-\alpha)|q(z)|^2\,d\mu(z)$$

It is easily seen that (1) implies that α belongs to the convex hull of $\text{supp}(\mu)$. Hence all the Fejer polynomials attached to μ have zeros in the convex hull K of the support of μ.

E.7.1 Question. To each n we get the probability measure γ_n which assigns the point $\frac{1}{n}$ at every zero of p_n, where eventual mutliple zeros get the mass $\frac{e}{n}$ if the multiplicity is $e\geq 2$. Now $\{\gamma_n\}$ is a sequence of probability measures whose supports are confined to the compact convex set K. Hence there always exists weakly convergent subsequence. We ask if these weak limits are unique, i.e. if there exists a unique probability measure γ_* on K such that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{\nu=1}^{\nu=n} g(\alpha_\nu^{(n)}) = \int g\cdot d\gamma_*$$

hold for every continuous g-function on K where $\{\alpha_\nu^{(n)}\}$ denote the zeros of p_n.

E.8. Wronskian determinants.

Let $n\geq 1$ and $p_0(z),p_1(z)\dots p_n(z)$ is some $(n+1)$-tuple of polynomials which are $\mathbf{C}$-linearly independent, i.e. one regards $\mathbf{C}[z]$ as a complex vector space whose basis are the monomials $1,z,z^2,\dots$. The Wronskian $W(z)$ is the determinant of the $(n+1)\times(n+1)$-matrix with elements

$$w_{jk}(z)=p_j^{(k)}(x)$$

where $p_j^{(k)}(x)$ denote k:th order derivative for each $k\geq 1$. Under the sole assumption that the polynomials are $\mathbf{C}$-linearly independent the polynomial $W(z)$ is not identically zero. To prove this one argues by a contradiction. Regarding $\{w_{jk}(z)\}$ as a matrix with elements in the field $\mathbf{C}(x)$ the vanishing of $W(z)$ entails that there exist polynomials $q_0(z),\dots,q_n(z)$ which are not all identically zero while

$$(1) \qquad \sum_{k=0}^{k=n} q_k(z)\cdot p_j^{(k)}(z)=0 \quad : 0\leq j\leq n$$

This means that the differential operator

$$Q(z,\partial)=\sum q_k(z)\cdot\partial^k$$

annihilates the $(n+1)$-dimensional complex vector space generated by the p-polynomials. But this is impossible by general results about solutions to a differential operator with polynomial coefficients. In fact, since $Q(x,\partial)$ has order $\leq n$ its null space in $\mathbf{C}[z]$ is a vector space whose dimension cannot exceed n. For a detailed account we refer to § XX where one studies the Weyl algebra $A_1(\mathbf{C})$ of differential operators with polynomial coefficients.

E.8 Laguerre's theorem about zeros of polynomials.

Let $P(z) = z^n + c_{n-1}z^{n-1} + \ldots + c_0$ be a monic polynomial of some degree $n \geq 2$. The complex coefficients are written as $c_\nu = a_\nu + ib_\nu$ which yields a pair of polynomials with real coefficients:

$$R(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_0 \quad : S(z) = b_{n-1}z^{n-1} + \ldots + b_0$$

We shall determine the number of zeros of $P(z)$ located in ther upper half-plabne $\mathfrak{Im}\, z > 0$ where eventual mutliple zeros are repeated by their multiplicities. Let $\alpha_1 < \ldots < \alpha_k$ be the real zeros of R with odd multiplicities, i.e. at these zeros R changes sign. Under the assumption that $S(\alpha_\nu) \neq 0$ at these real zeros of R one has:

Theorem. *If $n = 2m$ is an even integer the number of zeros of $P(z)$ in the upper half-plane counted with multiplicities is equal to*

$$\text{(i)} \qquad m + \frac{1}{2} \cdot \sum_{\nu=1}^{\nu=k} (-1)^{\nu-1} \cdot \text{sign}(S(\alpha_\nu))$$

Proof. It relies on the argument principle to be exposed in Chapter XX. The strategy is to pursue the variation of $\arg P(x)$ along the real line. When $x << 0$ then $P(x) \simeq x^{2m}$ is real and positive which implies that the real part of P is positive when $-\infty < x < a_1$. Next, $P(a_1)$ is purely imaginary and equal to i or - i depending on the sign of $S(\alpha_1)$. If the sign is < 0 the argument has decreased $-\pi/2$ while it has increased with $\pi/2$ if $\arg S(a_1) > 0$. Next, on the interval (a_1, a_2) the real part of $P(z)$ is < 0 and signs are reversed while the variation of the argument along this interval is computed. Continuing in this way the reader can check that

$$\text{(i)} \qquad x \mapsto \arg(P(x))$$

yields a continuous function $\rho(x)$ on the real line which is $\simeq 0$ when $x = -R$ for large positive R while

$$\lim_{R \to +\infty} \rho(R) = -\pi \cdot \sum_{\nu=1}^{\nu=k} (-1)^{\nu-1} \cdot \text{sign}(S(\alpha_\nu))$$

Next, along the half-circle $\{z = R \cdot e^{i\theta} : 0 \leq \theta \leq \pi\}$ with R large, the polynomial $P(z) \simeq z^{2m}$ and here the reader can verify that its argument increases from zero to $2m\pi$. Along the closed curve bordered by a large real interval $[-R, R]$ and the half-curcle above, it follows that the total variation of $\arg(P)$ changes by $2m\pi$ minus the term from (xx). In §§ we shall learn that this gives the claim in Laguerre's theorem above.

Exercise. Show that if $P(z)$ has odd degree $n = 2m + 1$, then the number of zeros in the upper half-plane becomes

$$\text{(ii)} \qquad m + \frac{1}{2} + \frac{1}{2} \cdot \sum_{\nu=1}^{\nu=k} (-1)^{\nu} \cdot \text{sign}(S(\alpha_\nu))$$

Example. With $m = 1$ we consider a polynomial $P(z) = (z^2 - 1 + 2iaz$. So here $S(z) = 2az$ where Laguerre's assumption is that $a \neq 0$. Here

$$P(z) = (z + ai)^2 - a^2 - 1$$

If $a > 0$ we have $S(-1) = -a < 0$ and $S(1) = a > 0$ and see that (*) is zero which confirms that the zeros of P stay in the lower half-plane A more involved case occurs when $S(z) = az + b$ with $b \neq 0$. Suppose for example that $-a + b < 0$ while $a + b > 0$. Then P has two zeros in U_+. The reder is invited to confirm this by solving the second order equation.

Example. Let $m = 2$ and $a_1 < \ldots < a_4$ be the simple real zeros while $S(z) = z + 1$. So here $P(z) = r(z) + iz + i$. If $a_1 + 1 < 0$ while $a_k + 1 > 0$ for $2 \leq k \leq 4$ then Theorem xx gives the maximal number of four zeros in U_+. On the other hand, if $a_1 + 1$ and $a_2 + 1$ are < 0 while $a_3 + 1$ and $a - 4 + 1$ both are positive, then P has two zeros in U_+. The reader is invited to check Laguerre's result numerically by a computer where various polynomials can be tested.

Example. Let $n = 2m + 1$ be odd and $R(z) = z^{2m+1}$ while $S(0) \neq 0$. For a polynomial of the form

$$P(z) = z^{2m+1} + i(b_{2m}z^{2m} + \ldots + b_1 z + b_0)$$

where $\{b_\nu\}$ are real and $b_0 \neq 0$ the number of zeros in $\mathfrak{Im}\, z > 0$ is given by

$$m + \frac{1}{2} - \frac{1}{2} \cdot \mathrm{sign}(b_0)$$

It is instructive to check this formula numerically by choosing various S-polynomials with the sole constraint that $b_0 > 0$. Notice that m can be arbitrary large.

2. Möbius functions

Let a, z be a pair of complex numbers where $|a| < 1$ and $|z| \leq 1$. Set

(1) $$M_a(z) = \frac{z-a}{1-\bar{a}z}$$

In polar coordinates we write $a = se^{i\phi}$ and get

(2) $$M_a(z) = e^{i\phi} \cdot M_s(e^{-i\phi}z)$$

Thus, up to a rotations the study of M-functions is reduced to the case when a is real and positive. Let $0 \leq a < 1$ and set

(3) $$w = \frac{z-a}{1-az}$$

Notice that $z = re^{i\theta}$ gives

(i) $$|z-a|^2 = (r\cos\theta - a)^2 + r^2\sin^2\theta = r^2 + a^2 - 2ar\cos\theta$$

Similarly we find that

(ii) $$|1-az|^2 = 1 + a^2r^2 - 2ar\cos\theta$$

It follows that

$$|w|^2 = \frac{r^2 + a^2 - 2ar\cos\theta}{1 + a^2r^2 - 2ar\cos\theta} = 1 - \frac{(1-r^2)(1-a^2)}{1 + a^2r^2 - 2ar\cos\theta}$$

In particular $|w| < 1$ and we can solve out z in (3) and obtain:

(4) $$z = \frac{w+a}{1+aw}$$

This shows that $z \mapsto M_a(z)$ is a bijective map of the open unit disc D onto itself. Next, if $0 < a, b < 1$ we construct the composed map $M_b \circ M_a$. A calculation gives

(5) $$M_b(M_a(z)) = M_c(z) \quad \text{where} \quad c = \frac{a+b}{1+ab}$$

Remark. Hence the Möbius transforms M_b and M_a commute and the map

$$(a, b) \mapsto \frac{a+b}{1+ab} \quad : \quad 0 \leq a, b < 1$$

gives a product satisfies the associate law where $a = 0$ is the neutral element. To obtain a *commutative group* we need inverses. If a varies in the open interval $(-1, 1)$ we get a commutative group of bijective maps on D defined by $a \mapsto M_a : \ -1 < a < 1$ whose the group table satisfies

(*) $$M_b \circ M_a = \frac{a+b}{1+ab} \quad : \quad -1 < a, b < 1$$

Next we study the map M_a-map keeping $0 < a < 1$ fixed. When $|z| = 1$ one has $M_a[z)| = 1$, Hence there exists a map from the unit circle into itself defined by

$$e^{i\theta} \mapsto \frac{e^{i\theta} - a}{1 - ae^{i\theta}}$$

It means that we get a function $\phi(\theta)$ where $\phi(0) = 0$ and $\phi(2\pi) = 2\pi$ and

$$e^{i\phi(\theta)} = \frac{e^{i\theta} - a}{1 - ae^{i\theta}}$$

Exercise. Show that the derivative

$$\frac{d\phi}{d\theta} = \frac{1-a^2}{|1-ae^{i\theta}|^2}$$

Hence the ϕ-function is strictly increasing and since $\phi(2\pi) = 2\pi$ we have

(*) $$\int_0^{2\pi} \frac{1-a^2}{|1-ae^{i\theta}|^2} = 2\pi$$

2.1 The group $\mathcal{M}$. Above a was real. If we allow arbitrary $a \in D$ the family $\{M_a\}$ yields the full group of Möbius transformations denoted by $\mathcal{M}$. It contains the commutative subgroup $\mathcal{M}_*$ where we only use Möbius maps M_a with $-1 < a < 1$. In addition we have the rotation maps

$$z \mapsto e^{i\phi} z$$

which preserve the origin and in this way the unit circle T is identified with a commutative subgroup $\mathcal{M}$. However, the group $\mathcal{M}$ is not commutative. To see an example we study rotations given by $M_\phi(z) = e^{i\phi} \cdot z$ where $0 \leq \phi \leq 2\pi$. We get a map from the product set $T \times (-1, 1)$ which sends a pair (ϕ, a) to the Möbius transform

$$M_a \circ M_\phi(z) = \frac{e^{i\phi} z - a}{1 - e^{i\phi} a z} = e^{i\phi} \cdot M_\alpha(z) \quad \text{where} \quad \alpha = e^{-i\phi} a$$

At the same time we notice that

$$M_\phi \circ M_a(z) = e^{i\phi} \cdot M_a(z)$$

This shows that the pair M_ϕ and M_a do not commute when $a \neq 0$ and $0 < \phi < 2\pi$. It turns out that the group $\mathcal{M}$ is quite extensive. For example, given positive integer $n \geq 2$ it contains infinite subgroups which are free of rank n. We shall refer to § 5 for further comments about $\mathcal{M}$ and its subgroups after certain metrics have been introduced on the unit disc.

2.2 Some image curves.

Let $0 < a < 1$ and $0 < r < \frac{1-a^2}{a}$. Consider the image under $M_a(z)$ when $|z - a| = r$, i.e. when z moves on the circle of radius r centered at a. With $z = a + re^{i\theta}$ the image is a simple closed curve γ defined by

$$\text{(i)} \qquad \theta \mapsto \frac{re^{i\theta}}{1 - a^2 - are^{i\theta}}$$

2.2.1 Exercise. Prove that γ is a circle with center at the point

$$\text{(ii)} \qquad z_0 = \frac{r(1-a^2)}{(1-a^2)^2 - a^2 r^2}$$

and radius

$$\text{(iii)} \qquad \rho = \frac{r}{1 - a^2 - ar} - \frac{r}{1 - a^2 + ar} = \frac{2ar^2}{(1-a^2)^2 - a^2 r^2}$$

A hint is to observe that the absolute value in (i) takes its maximum when $\theta = 0$ and the minimum when $\theta = \pi$ so that γ stays inside the annulus

$$\{\frac{r}{1 - 2^2 + ar} \leq |z| \leq \frac{r}{1 - 2^2 + ar}\}$$

Then (ii-iii) can be deduced directly from the general fact to be proved in § XX that the Möbius transform above must map the circle $\{|z - a| = r\}$ onto another circle. Here it is instructive to check (ii-iii) using plots on a computer with some different choices of a and r.

Next, consider images of circles centered at the origin. With $0 < a < 1$ and $0 < r < \frac{1}{a}$ the image of $\{|z| = r\}$ yields a simple closed curve

$$\text{(vi)} \qquad w_r(\theta) = \frac{re^{i\theta} - a}{1 - are^{i\theta}} \quad : \; 0 \leq \theta \leq 2\pi$$

Notice that the curve is symmetric with respect to the real axis. Just as in the example above one calculates the maximum and the mimium of the absolute values in (vi) and the reader should verify that the image is a circle $\mathcal{C}$ with a center on the real axis which passes the two points

$$\text{(vii)} \qquad -\frac{r+a}{1+ar} \quad : \; \frac{r-a}{1-ar}$$

Moreover, the radius of $\mathcal{C}$ is equal to

$$\text{(viii)} \qquad \frac{2r(1-a^2)}{1-a^2r^2}$$

2.2.3 Images of general circles. Consider a circle $\mathcal{C} = \{|z - z_0| = r\}$ where $z_0 \in D$ and $r < 1 - |z_0|$. To find the image under $M_a(z)$ we write

$$(*) \qquad M_a(z) = \frac{z - \frac{1}{a}}{1 - a \cdot z} + \frac{\frac{1}{a} - a}{1 - a \cdot z} = -\frac{1}{a} + \frac{1-a^2}{a} \cdot \frac{1}{1 - a \cdot z}$$

Since translates of circles are circles and a dilation of the scale preserve circles it suffices to consider the image of $\mathcal{C}$ under the map

$$z \mapsto \frac{1}{1 - a \cdot z}$$

With $\zeta = 1 - az$ the given circle $\mathcal{C}$ is mapped to a circle $\mathcal{C}^*$ in the complex ζ-plane. Then there only remains to analyze the effect under an inversion map $\zeta \mapsto \frac{1}{\zeta}$.

Exercise. Let $\mathcal{C}$ be a circle in the complex ζ-plane which does not contain the origin. Show that its image under the inversion map is a new circle.

Example. Use a computer to plot images under the Möbius transform $M_a(z)$ of the circles defined by $|z - i/2| = r$ where $0 < r < 1/2$ and $0 < a < 1$.

2.2.4 The case $r = 1$. With $0 < a < 1$ we study how $M_a(z)$ maps the unit circle onto itself. Consider the map

$$\text{(iv)} \qquad \theta \mapsto \frac{e^{i\theta} - a}{1 - ae^{i\theta}} \quad : \ 0 \leq \theta \leq 2\pi$$

We already know that complex numbers of absolute value 1 appear which gives a function $\phi(\theta)$ such that

$$(*) \qquad e^{i\phi(\theta)} = \frac{e^{i\theta} - a}{1 - ae^{i\theta}} \quad : \ 0 \leq \theta \leq 2\pi$$

Identifying the imaginary parts we get:

$$\sin \phi(\theta) = \frac{(1-a^2) \cdot \sin\theta}{1 + a^2 - 2a\cos(\theta)} \quad : \ 0 \leq \theta \leq \pi$$

We see that $\phi(\theta)$ is *strictly increasing* on the interval $0 \leq \theta \leq \pi$ and the reader may check that the derivative at $\theta = 0$ becomes $\frac{1+a}{1-a}$.

2.2.5 The absolute value $|M_a(z)|$. Let $0 < a < 1$. If $z = re^{i\theta}$ we get

$$\text{(i)} \qquad |M_a(re^{i\theta})|^2 = \frac{|re^{i\theta} - a|^2}{|1 - \bar{a}re^{i\theta}|^2} = \frac{r^2 + a^2 - 2ar\cos\theta}{1 + a^2r^2 - 2ar\cos\theta}$$

Keeping θ fixed (i) is a function of r. Applying the real Log-function we get

$$\text{(ii)} \qquad \log|M_a(re^{i\theta})| = \frac{1}{2}\log\left[\frac{r^2 + a^2 - 2ar\cos\theta}{1 + a^2r^2 - 2ar\cos\theta}\right]$$

Using one-variable calculus we take the r-derivative of the right hand side. To find the result we apply the usual formula for the real Log-function so that (ii) becomes

$$\text{(iii)} \qquad \frac{1}{2} \cdot [\log(r^2 + a^2 - 2ar\cos\theta) - \log(1 + a^2r^2 - 2ar\cos\theta)]$$

Using (iii) an easy calculation gives:

2.2.6 Proposition. *The r-derivative evaluated at $r = 1$ becomes:*

$$\frac{1-a^2}{1 + a^2 - 2a \cdot \cos\theta}$$

Above a is positive and real. But recall from (2) in the beginning of § 2 that if $a = se^{i\phi}$ then we only have to rotate z to get a similar M-function which gives:

2.2.7 Proposition. *With $a = se^{i\phi}$, the r-derivative evaluated when $r = 1$ of $r \mapsto |M_a(re^{i\theta})|$ becomes*

$$(*) \qquad \frac{|re^{i\theta} - ae^{i\phi}|}{|1 - ae^{-i\phi}re^{i\theta}|} = \frac{1 - s^2}{1 + s^2 - 2s \cdot \cos(\theta - \phi)}$$

2.2.8 Curves of symmetry. Let D be the unit disc. A circular arc α in D is called a curve of symmetry if it intersects the unit circle at right angles. We also include diameteres, i.e. straight lines which passes the origin, in this family of curves to be denoted by $\mathcal{S}(D)$. By drawing a picture and using euclidian geometry one verifies that if a and b is a pair of points in D then there exists a unique curve $\alpha \in \mathcal{S}(D)$ which passes through these points. In the case when a is the origin this curve is the diameter passing b. Next, if p and q are end-points on an interval of the unit circle of length $< \pi$ there exists a unique $\alpha \in \mathcal{S}(D)$ which intersects T at these two points.

Example. Let $0 < \theta < \pi/2$ and consider the points $e^{i\theta}$ and $e^{-i\theta}$. With $0 < a < 1$ we consider the Möbius transform

$$M(z) = \frac{a - z}{1 - az}$$

It is quite special becauce the composed map $M \circ M$ is the identity. In fact

$$M^2(z) = \frac{a - \frac{a-z}{1-az}}{1 - a \cdot \frac{a-z}{1-az}} = z$$

where the reader can check the last equality. Now we seek a in order that

$$e^{-i\theta} = \frac{a - e^{i\theta}}{1 - ae^{i\theta}}$$

A computation gives

$$a = \cos\theta$$

2.2.9 Exercise. Let α be the curve in $\mathcal{S}(D)$ which intersects T at $e^{i\theta}$ and $e^{-i\theta}$. Show that M maps α into itself, i.e. this curve is M-invariant and plot M-images of points in $D \setminus \alpha$ to discover that M behaves like a reflection relative its invariant curve α. Notice also that M has a sole fixed point in D which is placed on the real x-line and determined by the equation

$$x = \frac{a - x}{1 - ax} \implies x = \sqrt{a^{-2} - 1} - a^{-1}$$

3. The Laplace operator.

The second order differential operator $\partial_x^2 + \partial_y^2$ is denoted by Δ and called the Laplace operator. Consider a real-valued function $u(x,y)$ of class C^2. At the origin u has a Taylor expansion

(i) $$u(x,y) = u(0,0) + ax + by + Ax^2 + By^2 + Cxy + O(x^2+y^2)$$

where the remainder term is small ordo of $x^2 + y^2$.

Exercise. Show that when $r > 0$ is small then the mean-value integral

$$M_u(r) = \frac{1}{2\pi} \cdot \int_0^{2\pi} u(re^{i\theta}) \cdot d\theta = u(0,0) + \frac{A+B}{2} \cdot r^2 + o(r^2)$$

Here $\Delta u(0,0) = 2A + 2B$. So if the Laplacian is > 0 at the origin then we locally obtain a strict mean-value inequality in the sense that

$$\lim_{r \to 0} \frac{M_u(r) - u(0,0)}{r^2} = \frac{\Delta u(0,0)}{4}$$

In general, let u be a C^2-function in an open set Ω. If $p \in \Omega$ and r is $<$ the distance from p to $\partial\Omega$ we set

$$M_u(r,p) = \frac{1}{2\pi} \cdot \int_0^{2\pi} u(p + re^{i\theta}) \cdot d\theta \quad : r < \text{dist}(p, \partial\Omega)$$

Taking the derivative with respect to r we obtain

(i) $$\frac{dM_u}{dr}(r,p) = \frac{1}{2\pi} \cdot \int_0^{2\pi} \big(\cos\theta \cdot u_x(p + re^{i\theta}) + \sin\theta \cdot u_y(p + re^{i\theta})\big)\, d\theta$$

If Ω is the disc of radius r centered at p we shall learn in Chapter II that (i) gives

(*) $$r \cdot \frac{dM_u}{dr}(r,p) = \frac{1}{2\pi} \cdot \int_{\partial\Omega} (u_x \cdot \mathbf{n}_x + (u_y \cdot \mathbf{n}_y)\, ds = \frac{1}{2\pi} \cdot \iint_\Omega \Delta(u)\, dxdy$$

A C^2-function $u(x,y)$ satisfying $\Delta(u) = 0$ is called a harmonic function. From (*) it follows that the vanishing of $\Delta(u)$ implies that the function $r \mapsto M_r(r,u)$ is constant, i.e. one has

3.1 Theorem. *Let u be a harmonic C^2-function defined in an open set Ω. Then*

$$u(p) = \frac{1}{2\pi} \cdot \int_0^{2\pi} u(p + re^{i\theta}) \cdot d\theta$$

hold for each point $p \in \Omega$ and every $r <$ dist$(p, \partial\Omega)$.

A converse result. Let u be a C^2-function satifying the mean-value equations in Theorem 3.1. Then u is harmonic. Indeed, this follows from (*) which shows that the integrals of $\Delta(u)$ are zero over all discs contained in Ω, which implies that the continuous function $\Delta(u)$ vanishes identically.

3.2 Poisson's harmonic function. Keeping θ fixed while $a = x + iy$ varies in the open unit disc D we get the function

$$(x,y) \mapsto \frac{1 - x^2 - y^2}{1 + x^2 + y^2 - 2s \cdot \cos(\theta - \phi)} \;:\; s = \sqrt{x^2 + y^2} \quad : x = s \cdot \cos(\phi) \;:\; y = s \cdot \sin(\phi)$$

The addition formula for the cosine function gives

(1) $$u(x,y) = \frac{1 - x^2 - y^2}{1 + x^2 + y^2 - 2\cos(\theta) \cdot x + 2\sin(\theta) \cdot y}$$

It turns out that u is harmonic. A direct verification by taking derivatives in x and y is a bit messy so we give a proof using invariance properties of the Laplace operator. Namely, Δ commutes with rotations, i.e. if $g(x,y)$ is a C^2-function we set

$$g_\theta(x,y) = g(x\cos\theta + y\sin\theta, x\sin\theta + y\cos\theta) \quad : \quad 0 \le \theta \le 2\pi$$

With this notation one has

$$(*) \qquad \Delta(g_\theta) = (\Delta(g))_\theta$$

The verification is left to the reader. So to prove that u is harmonic we can take a rotation and assume that $\theta = 0$ in (1). There remains to consider the function

$$(2) \qquad h(x, y) = \frac{1 - x^2 - y^2}{1 + x^2 + y^2 - 2x} = \frac{1 - x^2 - y^2}{(1 - x)^2 + y^2}$$

Here a further simplification is possible since Δ also commutes with *translations*. So by the linear map $x \to \xi + 1$ and $y \to \eta$ there remains to regard the function

$$(3) \qquad k(\xi, \eta) = \frac{-2\xi - \xi^2 - \eta^2}{\xi^2 + \eta^2} = -1 - 2 \cdot \frac{\xi}{\xi^2 + \eta^2}$$

Notice that we only regard h when $x^2 + y^2 < 1$ which means that we only consider the k-function when $\xi^2 + \eta^2 \neq 0$. Now it is obvious that:

$$(4) \qquad (\partial_\xi^2 + \partial_\eta^2)(\frac{\xi}{\xi^2 + \eta^2}) = 0 \quad : \xi^2 + \eta^2 > 0$$

Hence Poisson's u-function is harmonic.

3.3 The function $\log((x-a)^2 + (y-b)^2)$**.** Let a, b be two real numbers. In $\mathbf{R}^2 \setminus (a, b)$ we have $(x-a)^2 + (y-b)^2) > 0$ where the real-valued Log-function above is defined. We shall study its partial derivatives. First we get:

$$\text{(i)} \qquad \partial_x(\log((x-a)^2 + (y-b)^2)) = \frac{2(x-a)}{(x-a)^2 + (y-b)^2}$$

Taking the second order partial derivative we obtain

$$\text{(ii)} \qquad \partial_x^2(\log((x-a)^2 + (y-b)^2)) = \frac{2}{(x-a)^2 + (y-b)^2} - \frac{4(x-a)^2}{[(x-a)^2 + (y-b)^2]^2}$$

A similar result holds when we apply ∂_y^2. With $\Delta = \partial_x^2 + \partial_y^2$ we add up the result and obtain:

$$\text{(iii)} \qquad \Delta(\log{(x-a)^2} + (y-b)^2) = 0$$

Thus, the Log-function satisfies the Laplace equation in $\mathbf{R}^2 \setminus (a, b)$.

3.4 The $\mathcal{L}_\epsilon$-functions. There remains to understand what occurs at (a, b). Since the situation is invariant under translation we can take (a, b) as the origin and with $\epsilon > 0$ consider the function

$$\mathcal{L}_\epsilon(x, y) = \log[\ (x-a)^2 + (y-b)^2 + \epsilon\,]$$

Here derivatives exist in the whole of $\mathbf{R}^2$ and a calculation gives

$$(1) \qquad \Delta(\mathcal{L}_\epsilon)(x, y) = \frac{4\epsilon}{(x^2 + y^2 + \epsilon)^2}$$

Let us calculate the area integral over $\mathbf{R}^2$. Using polar coordinates it becomes

$$(2) \qquad 4\epsilon \cdot 2\pi \int_0^\infty \frac{rdr}{(r^2 + \epsilon)^2} = 8\epsilon \cdot \pi \cdot \frac{1}{2}\, \frac{1}{(r^2 + \epsilon)}\,|_0^\infty = 4 \cdot \pi$$

Taking one half of the $\mathcal{L}$-function which means that we take a square root of the Log-function we have therefore proved:

3.5 Proposition. *For every $\epsilon > 0$ one has*

$$\frac{1}{2\pi} \cdot \iint_{\mathbf{R}^2} \Delta(\log \sqrt{x^2 + y^2 + \epsilon}) \cdot dxdy = 1$$

Remark. Thus, for every $\epsilon > 0$ we have the function $\log\sqrt{x^2+y^2+\epsilon)}$ whose Laplacian is a positive function and its integral taken over $\mathbf{R}^2$ is equal to 2π. The passage to the limit as $\epsilon \to 0$ leads to an important conclusion. Namely, let $\phi(x,y)$ be an arbitrary C^2-function with compact support. In chapter II we shall learn that Green's formula entails that

$$\text{(i)} \qquad \iint \Delta(\phi)\cdot \mathcal{L}_\epsilon \cdot dxdy = \iint \phi \cdot \Delta(\mathcal{L}_\epsilon)\cdot dxdy$$

hold for each $\epsilon > 0$. Using (1-2) the reader may verify the limit formula

$$\text{(ii)} \qquad \lim_{\epsilon\to 0} \iint \phi \cdot \Delta(\mathcal{L}_\epsilon)\cdot dxdy = 4\pi\cdot\phi(0,0)$$

Hence the left hand side in (i) also has a limit. In the appendix about distributions we shall learn that the limit formula (ii) means that the Laplacian taken in the distribution sense of the locally integrable function

$$\log|z| = \log\sqrt{x^2+y^2}$$

is equal to 2π times the Dirac measure at the origin, i.e. one has the equation

$$(*) \qquad \Delta(\log|z|) = 2\pi\cdot\delta_{(0,0)}$$

3.6 Subharmonic functions. A C^2-function u is called subharmonic if $\Delta(u) > 0$. Using Green's formula and the log-function one gets a formula for the deviation between values of u at a point p and mean-values taken over discs centered at p. More precisely, suppose that p is the origin and u is defined in some disc $\{|z| < R\}$. When $0 < r < R$ the function $\log\frac{|z|}{r}$ is zero on $|z| = r$. In Chapter II we shall learn that one has the equation:

$$(*) \qquad u(0,0) = M_u(r) + \frac{1}{2\pi}\iint_{[z|<r} \log\frac{|z|}{r}\cdot\Delta(u)(x,y)\cdot dxdy$$

Since $\log\frac{|z|}{r} < 0$ when $|z| < r$ the integral above is negative which means that one has the mean-value inequality

$$u(0,0) \le M_u(r)$$

3.7 Radial functions. Outside the origin in $\mathbf{R}^2$ we can express the Lapålace operator in polar coordinates. Namely, with $u(x,y) = u(r\cos\theta, r\sin\theta)$ we get the equations

$$\partial u/\partial r = \cos\theta\cdot u_x + \sin\theta\cdot u_y$$

$$\partial u/\partial\theta = -r\sin\theta\cdot u_x + r\cos\theta\cdot u_y$$

Exercise. Verify from the above that the following hold for first order partial differential operators:

$$\text{(i)} \qquad \partial_x = \cos\theta\cdot\partial_r - \frac{1}{r}\sin\theta\cdot\partial_\theta \quad : \partial_y = \sin\theta\cdot\partial_r + \frac{1}{r}\cos\theta\cdot\partial_\theta$$

Show now that

$$\partial_x^2 = \cos^2\theta\cdot\partial_r^2 + \frac{1}{r^2}\sin^2\theta\cdot\partial_\theta^2 - \frac{2}{r}\sin\theta\cdot\cos\theta\cdot\partial_r\partial_\theta + \frac{1}{r^2}\cdot\sin\theta\cdot\cos\theta\cdot\partial_\theta + \sin^2\theta\frac{1}{r}\cdot\partial_r$$

and derive a similar formula for ∂_y^2. Finally, show that

$$(*) \qquad \Delta = \partial_r^2 + \frac{1}{r^2}\cdot\partial_\theta^2 + \frac{1}{r}\cdot\partial_r$$

Let $\phi(r)$ be a function defined for the positive real numbers where it has at least two derivatives. In the punctured complex plane with the origin is removed we get the function

$$z \mapsto \phi(|z|)$$

Using (*) above we obtain:

3.8 Proposition. *One has the equation*

$$\Delta(\phi(|z|) = \frac{1}{|z|} \cdot \phi'(|z|) + \phi''(|z|)$$

3.9 The inhomogeneous Laplace equation. If ϕ is a continuous and bounded function in the open uint disc D we define the function

$$f(z) = \frac{1}{2\pi} \cdot \iint_D \log \frac{|1 - \bar{\zeta}z|}{|z - \zeta|} \cdot \phi(\zeta)\, d\xi d\eta \tag{3.9.1}$$

3.10 Exercise. Show that f satisfies the equation

$$\Delta(f)(z) = \phi(z) \quad : z \in D$$

and use properties of Möbius functions to conclude that $f = 0$ on the unit circle. Next, set

$$M(z) = \sqrt{\iint_D \big[\log \frac{|1 - \bar{\zeta}z|}{|z - \zeta|}\big]^2\, d\xi d\eta}$$

Use the Cauchy-Schwarz inequality to conclude that

$$|f(z)|| \leq 2\pi \cdot M(z) \cdot || \cdot \sqrt{\iint_D |\phi(\zeta)|^2\, d\xi d\eta}$$

Notice that the function $M(z)$ is bounded in D and hence

$$||f||_\infty = \max_{z \in D} |f(z)| \leq C \cdot ||\phi||_2 \quad : C = \frac{||M||_\infty}{2\pi} \tag{3.10.2}$$

where $||\phi||_2$ is the norm of ϕ in $L^2(D)$.

Remark. The integral in (3.10.1) is defined for every square integrable function ϕ in D and the results above ential that Δ yieds an injective map from the Hilbert space $L^2(D)$ into the Banach space of continuous functions on the closed unit disc which are zero on T. But the range of this linear operator is not so easily found. One reason is that when f is given by (3.10.1) from some $\phi \in L^2(D)$ then it has more regularity in the open disc than mere continuity. So a more refined study of this range leads to quite involved problems which go beyond these notes. But we shall enbcounter realted results which for example deal with functions in the unit disc with a finite Dirichlet integral. The restriction to the unit disc above is not so essential, i.e. in Chapter V we construct Green's functions which enable us to extend the study of Δ on $L^2(\Omega)$ on a more extensive famiy of domains in the complex plane.

4. Some complex mappings

4.1 The inversion $z \to \frac{1}{z}$. Let w be a new complex variable and consider the map

$$z \mapsto \frac{1}{z} = w \tag{1}$$

Denote by $\mathcal{C}$ the family of all circles of the form

$$|z - z_*| = r \quad : \quad 0 < r < |z_*| \tag{*}$$

It turns out that every image of a circle in (*) gives a circle in the w-plane. To show this we express the equation for circles in $\mathcal{C}$ in another way. Namely, the equation for a circle in $\mathcal{C}$ can be given by

$$az\bar{z} + bz + \bar{b}\bar{z} + c = 0 \quad : a, c \text{ real and both } \neq 0 \tag{2}$$

The proof of (2) is left as an exercise. From (2) it is easily seen that the equation for the image with $w = \frac{1}{z}$ becomes

$$cw\bar{w} + \bar{b}w + b\bar{w} + a = 0 \tag{3}$$

Here (3) is the equation for a circle in the w-plane which again does not contain the origin. Hence the inversion map gives a *1-1 correspondence* between the class of circles whose interior discs do not contain the origin.

Exercise. Consider a circle in $\mathcal{C}$ defined by an equation

$$|z - A| = s \quad : \quad 0 < s < A$$

Thus, the center is the real point A and the radius is s. Show that the equation for the image circle becomes

$$|w - \frac{A}{A^2 - s^2}| = \frac{s}{A^2 - s^2} \tag{*}$$

Circles which contain the origin. Let $\mathcal{C}_*$ be the family of circles in the z-plane which contain $z = 0$. The equation for a circle in $\mathcal{C}_*$ is given as in (2) above where

$$a, b \neq 0 \quad : \quad c = 0$$

The image circle gets the equation

$$\bar{b}w + b\bar{w} + a = 0$$

With $b = re^{i\theta}$ and a real we get

$$\mathfrak{Re}\, e^{i\theta} \cdot w = -\frac{a}{2r} \tag{4}$$

This is the equation for a line in the w-plane and the reader should contemplate upon some specific examples to see how these lines are affected when θ varies and also verify that if C_1 and C_2 are two different circles in $\mathcal{C}_*$ then the image lines are not equal. Indeed, θ determines the direction of a line and once $\arg(b)$ is fixed, the quotients $\frac{a}{r}$ are never equal for a pair of $\mathcal{C}$-circles defined by (2) with $c = 0$. Hence there is a 1-1 correspondence between circles in the z-plane which contain the origin and lines in the w-plane which do not contain the origin. Moreover, under this correspondence a pair of circles C_1 and C_2 from $\mathcal{C}_*$ gives two *parallell lines* in the w-plane if and only if

$$\arg(b_1) = \arg(b_2) \tag{5}$$

The reader should verify that (5) holds if and only if $\mathcal{C}_1$ and $\mathcal{C}_2$ are tangent to each other at the origin and illustrate this by figures.

4.2 A specific example. Consider the family of circles

$$\mathcal{C}_r = \{(x - r)^2 + y^2 = r^2\} \quad : r > 0$$

In the complex notation the equations are:

$$z\bar{z} - rz - r\bar{z} = 0$$

So here $a = 1$ and $b = -r$. The image line has therefore the equation

$$\mathfrak{Re}\, w = \frac{1}{2r} \tag{6}$$

Take as an example $r = 1$. On $\mathcal{C}_1$ we find the point $(2, 0)$ and its image is the real w-pont $\frac{1}{2}$ is okay by (6) since $r = 1$. Moving along $\mathcal{C}$ in the positive direction we set

$$z = 1 + e^{i\theta} \quad : \, 0 < \theta < \pi$$

and get the image points $w(\theta) = \frac{1}{1+e^{i\theta}}$ where

$$\mathfrak{Im}\, w(\theta) = -\frac{\sin\theta}{2 + 2\cos\theta}$$

Notice the minus sign, i.e. $w(\theta)$ travels in the negative direction on the vertical line $\mathfrak{Re}\, w = \frac{1}{2}$ while z moves along $\mathcal{C}$ in the positive sense. This is no surprise because the inversion changes the orientation.

4.3 The map $z \mapsto z + \frac{1}{z}$

Here the geometric description is more involved. Let us give some examples of images under this map from the z-plane to the w-plane where we recall that $w = u + iv$.

4.4 Proposition 1 *Let $\mathcal{C}$ be a circle of radius $r > 1$. Then its image $\mathcal{C}_*$ in the w-plane is an ellipse defined by the equation*

$$\frac{u^2}{4r^2} + \frac{v^2}{(r - \frac{1}{r})^2} = 1$$

The verification is left to the reader. In the equation above the ellipse has focal points and again we leave as an exercise to show that they are placed at $(1, 0)$ and $(-1, 0)$ which means that

$$|w - 1| + |w + 1| = 4$$

for all points on the ellipse. Notice that the focal points do not depend on r and when $r \to 1$ the ellipse approaches the real interval $\{-2 \leq u \leq 2\}$.

4.5 Proposition *Let $\ell = \{se^{i\theta} : \, -\infty < s < \infty\}$ be a line but not the real or the imaginary axis. Then ℓ_* is the hyperbel defined by*

$$\frac{u^2}{\cos^2(\theta)} - \frac{v^2}{\sin^2(\theta)} = 1$$

4.6 Remark. We leave the proof as an exercise and the reader should verify that the focal points are the same as for the ellipses above, i.e. placed at $(1, 0)$ and $-1, 0)$. Recall a result in euclidian geometry which asserts that when an ellipse and a hyperbel have common focal points, then they intersect at right angles. Above the lines and the circles intersect with the angle $\pi/2$ in the z-plane and we shall learn that the map $z \to z + \frac{1}{z}$ is conformal which means that angles are preserved. Hence the classical result from euclidian geometry is rediscovered via Proposition 4.4-4.5 using complex calculus.

4.7 The complex exponential e^z

Consider the strip domain defined by

$$\square = \{(x,y) \quad : \ -\infty < x < \infty \quad \text{and} \quad 0 < y < 2\pi\}$$

Let $\zeta = \xi + i\eta$ be a new complex variable. We construct a map from the (x,y)-plane into the (ξ,η)-plane by

(*) $$\xi = e^x \cdot \cos(y) \quad \text{and} \quad \eta = e^x \cdot \sin(y)$$

It is clear that this gives a 1-1 map from $\square$ onto the (ξ,η)-plane where the non-negative axis $\{\xi \geq 0\} \cap \{\eta = 0\}$ has been removed.

4.8 Some geometric images. The vertical line $\ell^*(a) = \{x = a\} \cap \{0 < y < 2\pi\}$ is mapped to a circle $\xi^2 + \eta^2 = e^{2a}$ where the point $(e^{2a}, 0)$ has been excluded. The image of a horizontal line $\ell_*(b) = \{y = b\}$ becomes a half-ray defined by

$$\xi = r \cdot \cos(b) \quad \text{and} \quad \eta = r \cdot \sin(b)$$

Notice that the image circles and the half-rays intersect at a right angle. The same is true for the corresponding families of vertical, respectively horizontal line in the strip. As we shall see later on this is no accident since it reflects the conformality of the complex analytic function e^z which corresponds to the map above. In fact, $z \mapsto e^z$ is a conformal map from $\square$ onto $\mathbf{C} \setminus \mathbf{R}^*+$. Let us now describe images of a more involved nature.

4.9 Images of circles. Let $0 < r < \pi/2$ and consider the circle

$$\mathcal{C}(r) \quad : \quad x^2 + (y - \pi/2)^2 = r^2$$

In polar coordinates we write $x = r \cdot \cos(\phi)$ and $y = \pi/2 + r \cdot \sin(\phi)$ and now the image is a closed curve parametrized by ϕ where

$$\xi(\phi) = e^{r \cdot \cos(\phi)} \cdot \cos(\pi/2 + r \cdot \sin(\phi)) \quad \text{and} \quad \eta(\phi) = e^{r \cdot \cos(\phi)} \cdot \sin(\pi/2 + r \cdot \sin(\phi))$$

Using *Mathematica* the reader can plot these closed curves as $0 < r < \pi/2$ varies. When $r \to 0$ the curves become more and more circular.

4.10 The inverse map. From (*) we get

$$\xi^2 + \eta^2 = e^{2x} \implies x = \text{Log}(\sqrt{\xi^2 + \eta^2})$$

Using the complex argument we see that

$$y = \arg(\xi + i\eta)$$

where the argument is determined when it takes values in $(0, 2\pi)$. In this way the inverse map is described by the complex log-function $\log \zeta$ where a single-valued branch has been chosen when the non-negative real ξ-axis has been removed.

Remark. The discussion above shows that one can start from constructions in real analysis dealing with vector-valued functions in $\mathbf{R}^2$ and after make complexifications of these. So it is foremost a matter of notations to use z and ζ instead. But in the long run the complex notations are convenient to construct power series expansions of analytic functions. However, one should not forget the underlying real picture. As an example, from (*) we can pay attention to the function $\xi = \xi(x,y)$ and notice that the partial derivatives become

(1) $$\xi'_x = e^x \cdot \cos(y) \quad \text{and} \quad \xi'_y = -e^x \cdot \sin(y)$$

A similar computation of the partial derivatives of the η-function give the two identities

(2) $$\xi'_x = \eta'_y \quad \text{and} \quad \xi'_y = -\eta'_x$$

We shall learn in Chapter 3 that this expresses the Cauchy-Riemann equations which hold because e^z is an analytic function. Notice also that $\xi(x,y)$ satisfies the Laplace equation $\Delta(\xi) = 0$, i.e. it yields a harmonic function and similarly $\Delta(\eta) = 0$.

4.11 The harmonic angle function.

With $z = x + iy$ we refer to $\{y > 0\}$ as the upper half-plane in $\mathbf{C}$. It is denoted by U_+ and the boundary is the real x-axis. If $z \in U_+$ we use polar coordinates and write:

$$z = x + iy = r \cdot \cos\theta + i \cdot r \cdot \sin\theta$$

where $r = |z| = \sqrt{x^2 + y^2}$ and $0 < \theta < \pi$. Here θ is determined in a unique way which gives the function

$$(x, y) \mapsto \theta \tag{1}$$

Introducing the argument of complex numbers this means that:

$$\arg(z) = \theta \tag{2}$$

Now $\theta = \theta(x, y)$ is regarded as a function of x an y.

4.12 Exercise. Prove that the partial derivatives of the θ-function become:

$$\theta'_x = -\frac{y}{x^2 + y^2} \quad \text{and} \quad \theta'_y = \frac{x}{x^2 + y^2}$$

Deduce that $\Delta(\theta) = 0$, i.e. the θ-function is harmonic in U_+.

4.13 Exercise. Let $a < b$ be two real numbers. In U_+ we define the function

$$H_{a,b}(z) = \arg(z - b) - \arg(z - a)$$

Use Exercise 4.12 to show that $\Delta(H) = 0$ and show the limit formulas:

$$\lim_{y \to 0} H(x + iy) = \pi \quad \text{for all} \quad a < x < b \tag{1}$$

$$\lim_{y \to 0} H(x + iy) = 0 \quad x < a \quad \text{or} \quad x > b \tag{2}$$

4.14 Remark. The harmonic H-functions above can be used to solve the Dirichlet problem where one starts with a bounded continuous function $f(x)$ on the real x-line and seeks a harmonic function $F(x, y)$ in U_+ whose boundary values give f, i.e.

$$\lim_{y \to 0} F(x, y) = f(x) \tag{*}$$

We return to this in XXX.

5. The sterographic projection.

Introduction. A sphere Σ of diameter one is placed above the origin in the complex (x, y)-plane. Let (x, y, h) be the coordinates in $\mathbf{R}^3$. The center of the sphere is $(0, 0, 1/2)$ and Σ is defined by the equation

$$x^2 + y^2 + (h - 1/2)^2 = \frac{1}{4} \tag{0.1}$$

The point $N = (0, 0, 1)$ is called the north pole and the origin the south pole. If $(x, y, 1/2) \in \Sigma$ then (0.1) entails that

$$x^2 + y^2 = h - h^2 \tag{0.2}$$

The stereographic projection sends a point $p \in \Sigma \setminus N$ to p_* in $\mathbf{C}$ which arises when we draw the line through N and p which hits p_*. If $p = (x, y, h) \in \Sigma$ it is clear that

$$p_* = \frac{1}{1 - h} \cdot (x + iy) \tag{0.3}$$

In particular the southern hemi-sphere where $h < 1/2$ is mapped onto the unit disc D and the northern hemi-sphere $\{1/2 < h < 1\}$ is mapped onto the exterior disc $|z| > 1$. The equator circle $h = 1/2$ is mapped onto the circle of radius one centered at the origin in the z-plane.

Images of great circles.

A great circle $\mathcal{C}_\Pi$ on Σ arises when we take the intersection of Σ and a 2-dimensional plane Π in $\mathbf{R}^3$ which contains the center $(0, 0, 1/2)$. When Π contains the north-pole the image of $\mathcal{C}_\Pi$ under the map (0.2) becomes a straight line. When $\mathcal{C}_\Pi$ does not contain the northpole the image $\Pi_*$is a bounded closed curve. It turns out that it is a circle.

5.1. Theorem. *For every great circle $\mathcal{C}$ which does not contain N, the image $\mathcal{C}_*$ under (0.3) is a circle.*

Proof. Up to a rotation it suffices to consider a plane defined by the equation

$$x = A(h - 1/2) \quad : \; A > 0 \tag{1}$$

The points on $\Pi \cap \Sigma$ whose stereographic projection has maximal resp. minimal distance to the origin has y-coordinate zero, i.e. we seek points

$$p = (x, 0, h) \quad : \quad x^2 = h - h^2 \quad \text{and} \quad x = A(h - 1/2)$$

It means that h satisfies the second order algebraic equation

$$A^2(h - 1/2)^2 + (h - 1/2)^2 = 1/4$$

The two solutions become

$$h^* = \frac{1}{2}[1 + \frac{1}{\sqrt{1 + A^2}}] \quad : \quad h_* = \frac{1}{2}[1 - \frac{1}{\sqrt{1 + A^2}}]$$

The two projected points become

$$x^* = A \cdot \sqrt{\frac{h^*}{1 - h^*}} x_* = A \cdot \sqrt{\frac{h_*}{1 - h_*}}$$

At this stage the reader can finish the proof of Theorem 5.1.

Exercise. Determine the center and the radius of the image circle starting from the equation (1) above

5.2 The spherical metric.

In the complex z-plane we define a σ-metric as follows:

$$d\sigma = \frac{ds}{1 + |z|^2} \quad ds = \text{ the usual euclidian metric.}$$

This means that if γ is some parametrised C^1-curve then its σ-length becomes

$$\sigma(\gamma) = \int_0^T \frac{|\dot{z}(t)| \cdot dt}{1+|z(t)|^2}$$

where we have put $z(t) = x(t) + iy(t)$ so that $|\dot{z}(t)| = \sqrt{\dot{x}^2 + \dot{y}^2}$.

5.3 Geodesic curves. Given a pair of points p and q in the z-plane one seeks the minimum of $\sigma(\gamma)$ taken over all closed curves γ whose end-points are p and q. This leads to a variational problem. It turns out that the curve γ which minimises the σ-distance between two points p and q which both are outside the origin is a circular arc. In the case when p is the origin the minimizing curve is the straight line from p to q. To prove this one uses the stereograhic projection. First, on the sphere Σ one employs the usual metric induced by the euclidian metric in $\mathbf{R}^3$ and it is wellknown that geodesic curves on Σ are consist of arcs on great circles. For example, if p is the south pole then the shortest air-borne flight from p to another point q on the earth is to follow the great circle which passes through p and q. The crucial fact is that the stereographic projection is an *isometry* when we use the σ-metric in the complex z-plane. To prove this we first consider the case when p is the south pole and $q = (x, 0, h)$ is a point on the southern hemisphere with $y = 0$. Now $q_* = (x_*, 0)$ and the geodesic curve with respect to the σ-metric is the line from $z = 0$ to q_*. Its σ-distance becomes

$$\int_0^{x_*} \frac{dt}{1+t^2}$$

Next, the distance from the south pole to q is given by

$$\frac{\theta}{2} \quad \text{where} \quad \sin(\theta) = \frac{x}{2}$$

where θ is the angle between the vertical line from the north pole to the south pole and the line from the center $(0, 0, 1/2)$ to q. The reader may draw a picture to illustrate this and verify that

$$1 - 2h = \cos\theta \quad \text{and} \quad x = \frac{\sin\theta}{2}$$

The trigonometric formula $\cos\theta = 1 - 2 \cdot \sin^2(\theta/2)$ gives

$$h = \sin^2(\theta/2) \implies 1 - h = \cos^2(\theta/2)$$

Now $q_* = (x_*, 0)$where

$$x_* = \frac{x}{1-h} = \frac{1}{2} \cdot \frac{\sin(\theta)}{\cos^2(\theta/2)}$$

Since $\sin(\theta) = 2\sin(\theta/2) \cdot \sin(\theta/2)$ we finally obtain

$$x_* = \frac{\sin(\theta/2)}{\cos(\theta/2)} = \mathrm{tg}(\theta/2)$$

Hence the σ-distance becomes

$$\int_0^{x_*} \frac{dt}{1+t^2} = \mathrm{arctg}(x_*) = \frac{\theta}{2}$$

where the last term is the air-borne distance on the earth from the south pole to q. This proves that the stereograhic projection is an isometry in the special case when p is the south pole so that p_* is the origin, The general case is proved using a family of isometric maps with respect to the σ-metric given below.

5.4 Distance preserving maps. Let a be a complex number with $|a| < 1$. and consider the map

$$(*) \qquad M_a^*(z) = \frac{z-a}{1+\bar{a}z} \quad : \ |z| < 1$$

Set $w(z) = M_a^*(z)$. Differentiation gives

$$\frac{dw}{dz} = \frac{1}{1+\bar{a}z} - \frac{\bar{a}(z-a)}{(1+\bar{a}z)^2} = \frac{1+|a|^2}{(1+\bar{a}z)^2}$$

We have also

$$\frac{1}{1+|w|^2} = \frac{|1+\bar{a}z|^2}{|z-a|^2+|1+\bar{a}z|^2}$$

It follows that

(i)
$$\frac{|dw|}{1+|w|^2} = \frac{(1+|a|^2)\cdot|dz|}{|z-a|^2+|1+\bar{a}z|^2}$$

Next, we notice the identity

(ii)
$$(1+|a|^2)(1+|z|^2) = |z-a|^2+|1+\bar{a}z|^2 \quad : |a|<1$$

We conclude that

(*)
$$\frac{|dw|}{1+|w|^2} = \frac{|dz|}{1+|z|^2}$$

Hence the map $z \to M_a^*(z)$ preserves the σ-metric.

Exercise Suppose that $0<a<1$ is real and positive. Set

$$w^* = \frac{1+a}{1-a} \quad \text{and} \quad w_* = \frac{1-a}{1+a}$$

We get the middle point

$$w_0 = \frac{w^*+w_*}{2} = \frac{1+a^2}{1-a^2}$$

Now we get the circle centered at w_0 whose radius us

$$r = w^* - w_0 = \frac{2a}{1-a^2}$$

The reader should verify that the map $z \to M_a^*(z)$ is a bijective from the unit disc $|z|<1$ onto the disc $|w-w_0|<r$.

Remark. The transformations defined by (*) in 2.3 are not Möbius transforms since we have changed the sign for the term $\bar{a}z$.

5.5 The hyperbolic metric in D.

The hyperbolic metric in the open unit disc D is defined by

(*)
$$d\sigma_{hyp} = \frac{ds}{1-r^2} \quad : \quad |z|=r$$

5.6 Theorem. *Möbius transforms preserve the hyperbolic metric.*

Proof. Let $|a|<1$ and consider the Möbius transform

$$M_a(z) = \frac{z-a}{1-\bar{a}\cdot z} \quad : \quad z \in D$$

We must prove the differential equality

(i)
$$\lim_{\Delta z \to 0} \frac{d\sigma_{hyp}(z_0+\Delta z, z_0)}{|\Delta z|} = \lim_{\Delta z \to 0} \frac{d\sigma_{hyp}(M_a(z_0+\Delta z), M_a(z_0))}{|\Delta z|}$$

To prove (i) we take some point $z_0 \in D$ and with a small Δz we have

$$M_a(z_0+\Delta\, z)-M_a(z_0)=\frac{1-|a|^2}{(1-\bar{a}z_0)^2}\cdot\Delta\, z+O(|\Delta\, z|^2)\implies$$

(ii) $$d\sigma_{hyp}(M_a(z_0+\Delta z),M_a(z_0))=\frac{1-|a|^2}{|1-\bar{a}z_0|^2}\cdot\frac{|\Delta\, z|}{1-|M_a(z_0)|^2}+O(|\Delta z|^2)$$

Now the reader can verify the equation

(iii) $$\frac{(1-|a|^2)(1-|z_0|^2)}{|1-|M_a(z_0)|^2|^2}=\frac{1}{1-|z_0|^2}$$

so when $\Delta z\to 0$ it follows that (ii-iii) give the differential equality (i).

5.7 Geodesic curves. It is easily seen that if $z_0\neq 0$ is a point in D, then the geodesic curve in the hyperbolic metric from the origin to z_0 is the straight line whose distance becomes

(*) $$\int_0^r\frac{ds}{1-s^2}=\frac{1}{2}\cdot\log\,\frac{1+r}{1-r}\quad:\quad|z|=r$$

Thus, in the hyperbolic metric points close to the unit circle are remote from the origin.

5.8 General geodesic curves. Since Möbius transforms preserve the metric they preserve geodesic curves in the hyperbolic metric. Let a,b be a pair of points in D and take the Möbius transform

$$M(z)=\frac{z-a}{1-\bar{a}z}$$

Wirh $\beta=M(b)$ the distance between a and b is equal to that between the origin and β which gives:

$$\text{dist}(a,b)=\frac{1}{2}\,\log\frac{1+|\beta|}{1-|\beta|}\quad:\;|\beta|=|\frac{z-a}{1-\bar{a}z}|$$

For example, if $0<a<b<1$ their hyperbolic distance becomes

$$\log\,\frac{1+b-a}{1+a-b}$$

Next, with $a\neq b$ in D we already know that the geodesic line from a to b corresponds to the image of the straight line from the origin to β. It means that we regard the inverse Möbius transform

$$z\mapsto\frac{z+a}{1+\bar{a}z}$$

The geodesic line from a to b is parametrized by

$$r\mapsto\frac{re^{i\theta}+a}{1+are^{i\theta}}\quad:\;0\leq r\leq|b-a|$$

and the angle θ is the argument of β. Equivalenetly it is chosen so that

$$b=\frac{[b-a|\cdot e^{i\theta}+a}{1+a|b-a|\cdot e^{i\theta}}$$

Remark. It is instructive to make some plots to illustrate the shape of geodesic lines. Moreover, using the fact that Möbius transformations are conformal and preserve the unit circle the reader should verify the following:

5.9 Rule of Schwarz. Denote by $\mathcal{C}_*$ the family of circular arcs in D which at the two end-points intersect the unit circle at right angles. Include also the diameters, i.e. straight line segments passing the origin in this family. Show that every Möbius transform

$$M(z)=e^{i\theta}\cdot\frac{z-a}{1-\bar{a}z}$$

with $a \in D$ arbitrary yields a bijective map on the family $\mathcal{C}_*$. Moreover, for each pair of points a, b in D the geodesic curve from a to b is given by a *unique circular arc* in $\mathcal{C}_*$ which passes through a and b.

5.10 Schwarz derivatives

Let $0 < r < 1$ and $z = re^{i\phi}$ for some ϕ. If $w = e^{i\theta}$ is a point on the unit circle T we get the point $w^* \in T$ by drawing the line from w through z which hits T at w^*. Euclidian geometry shows that

$$(1) \qquad 1 - |z|^2 = |w - z| \cdot |w^* - z|$$

By this construction we get a map from T onto itself defined by

$$w \mapsto w^*$$

With $w = e^{i\theta}$ we set $w^* = e^{i\beta(\theta)}$. So now $\beta(\theta)$ is a 2π-period function.

Exercise. Show that the β-derivative is given by the formula:

$$(*) \qquad \frac{d\beta}{d\theta} = \frac{1 - |z|^2}{|w - z|^2} = \frac{1 - r^2}{1 + r^2 - 2r\cos(\theta - \phi)}$$

Remark. In the right hand side we recognize the Poisson kernel which therefore is expressed by the θ-derivative of $\beta = \beta(\theta)$ when $z = re^{i\phi}$ is kept fixed. Regarding the geometric picture for the construction of the β-function it is evident that when θ makes a full turn around the unit circle, then the β-angle also makes a full turn. Hence (**) gives:

$$(**) \qquad \int_0^{2\pi} \frac{1 - r^2}{1 + r^2 - 2r\cos(\theta - \phi)} = 2\pi$$

Remark. This geometric construction of the Poisson kernel is due to Schwarz who used it to prove that various integrals are invariant under Möbius transformations.

5.11 Die kreisgeometrischen Massbestimmung To each pair of points z_1, z_2 in D we set

$$\delta(z_1, z_2) = \left|\frac{z_1 - z_2}{1 - \bar{z}_1 \cdot z_2}\right|$$

Notice that the δ-distance is < 1 for each pair of points. When $z_2 = 0$ one has $\delta(z_1, 0) = |z_1|$, i.e. the ordinary euclidian distance from the origin to z_1. If $0 < a < b < 1$ are real and positive we see that

$$\delta(b, a) = \frac{b - a}{1 - ab} > b - a$$

So for such pair the δ distance is larger than the euclidian. Next, let $\zeta \in D$ and consider the Möbius transform

$$z \mapsto \frac{z - \zeta}{1 - \bar{\zeta} \cdot z}$$

Consider a pair z_1, z_2 and put

$$w_1 = \frac{z_1 - \zeta}{1 - \bar{\zeta} \cdot z_1} \quad : \quad w_2 = \frac{z_2 - \zeta}{1 - \bar{\zeta} \cdot z_2}$$

A computation which is left to the reader shows that

$$(*) \qquad \delta(z_1, z_2) = \delta(w_1, w_2)$$

Hence Möbius transforms preserve the δ-distance.

5.12 The triangle inequality. To prove that δ satisfies the triangle inequality we use (*) which reduces the proof to the case when $z_1 = 0$ and $z_2 = a$ is real and positive, ie. there remains to show:

$$\delta(0, z_3) \leq \delta(0, z_2) + \delta(a, z_3) \quad : \quad z_3 \in D$$

This amounts to show that

$$\text{(i)} \qquad |z_3| \leq a + \Big|\frac{z_3 - a}{1 - az_3}\Big|$$

Here (i) is obvious if $|z_3| \leq a$. Next, with b kept fixed we consider for some $a < r < 1$ all z_3 of absolute value r. Notice that

$$\min_\theta \Big|\frac{re^{i\theta} - a}{1 - re^{i\theta}}\Big|$$

is attained when $\theta = 0$, i.e. when $z_3 = r$ is real and positive. In this case we have

$$a + \frac{r - a}{1 - ar} = \frac{a - a^2 r + r - a}{1 - ar} - r = \frac{a - a^2 r + r - a - r + ar^2}{1 - ar} = \frac{ar(r - a)}{1 - ar} > 0$$

This proves the triangle inequality for the δ-function.

5.13 Exercise. Let $0 < a < 1$ and consider the set

$$E = \{w \in \mathbf{C} \quad : \quad \Big|\frac{w + a}{1 + aw}\Big| = \Big|\frac{w - a}{1 - aw}\Big|\}$$

Show that with $w = u + iv$ the equation for E becomes

$$4au(1 - a^2)(1 - u^2 - v^2) = 0$$

Hence E is the union of the unit circle and the imaginary axis. Use this result together with the invariance from (*)in § 5.11 to investigate sets of the form

$$E_{w_1,w_2} = \{w \in D \quad : \quad \Big|\frac{w - w_1}{1 - \bar{w}_1 w}\Big| = \Big|\frac{w - w_2}{1 - \bar{w}_2 w}\Big|$$

for a pair of points w_1, w_2 in D.

Chapter I:B. Series

Content of chapter

0. Introduction

1. Additive series

1:B Counting functions

2. Power series

2:B Radial limits

2:C A theorem by Landau

3. Product series

4. Blascke products

5. Estimates using the counting function

6. Convergence on the boundary

7. An example by Hardy

8. Convergence under substitution

9. The series $\sum (a_1 \cdots a_\nu)^{\frac{1}{\nu}}$

10. Thorin's convexity theorem.

11. Cesaro and Hölder limits

12. Power series and arithmetic means

13. Taylor series and quasi-analytic functions

Introduction.

Cauchy's integral formula in Chapter 3 shows that analytic functions are locally represented by a convergent power series. For this reason the study of series is important in analytic function theory. We start with additive series and after treat power series $\sum c_n z^n$ whose the radius of convergence ρ is determined by the formula

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

The study of convergence at boundary points where $|z| = \rho$ is in general involved. To begin with one studies existence of radial limits. For a fixed $0 \leq \theta < 2\pi$ one says that a power series (*) has a radial limit in the θ-direction if there exists

$$(**)\qquad \lim_{r\to 1} \sum c_n \cdot r^n \cdot e^{in\theta}$$

One can also consider less restricted limit conditions and we prove a result due to Landau of this kind at the end of § 2. *Blaschke products* play a central role in analytic function theory and are studied in § 4.

The final sections contains more advanced material which is not needed for the overall study of analytic functions. § 6 starts with Abel's theorem which gives a sufficient condition in order that an additive series $\sum a_n$ converges and has a sum s_* under the hypothesis that there exists

$$(1)\qquad \lim_{x\to 1} \sum a_n \cdot x^n = s_*$$

Abel's sufficiency condition is that

$$(2)\qquad \lim_{n\to\infty} \frac{a_n}{n} = 0$$

A weaker condition to get the convergence of $\sum a_n$ when (1) holds appears in a theorem due to Hardy and Littlewood. The result is that if $\{a_n\}$ is a sequence for which there exists a constant C such that

$$a_n \leq \frac{C}{n} \quad \text{hold for every} \quad n \geq 1 \tag{3}$$

then (1) implies that $\sum a_n$ is convergent. The proof contains several ingenious steps which demonstrates that the study of series is a rich and often quite hard subject. Thorin's convexity theorem in § 10 illustrates the usefulness of complex methods to prove inequalities which from the start are given in a real setting. In § 11 we study summations in the sense of Cesaro and Hölder and prove that they lead to equal limits whenever one of them exists. The material in sections 8,9 and 12 and 13 contains more advanced results due to Carleman which go beyond the basic study of analytic functions.

I. Additive series

1. Partial sums and convergent series. To a sequence $\{a_\nu\}$ of complex numbers indexed by non-negative integers one associates the partial sums:

$$S_N = \sum_{\nu=0}^{\nu=N} a_\nu \quad : \quad N = 1, 2, \ldots$$

If $\{S_N\}$ converges converge to a limit S_* one says that $\{a_\nu\}$ yields a convergent series and write

$$S_* = \sum_{\nu=0}^{\infty} a_\nu \tag{1}$$

2. Absolute convergence. The series is absolutely convergent if

$$\sum_{\nu=0}^{\infty} |a_\nu| < \infty$$

Absolute convergence implies that $\sum a_\nu$ converges. For if $\{S_N\}$ are the partial sums the triangle inequality gives

$$|S_M - S_N| = |\sum_{\nu=N+1}^{M} a_\nu| \leq \sum_{\nu=N+1}^{M} |a_\nu| \quad : \quad M > N \geq 0$$

The absolute convergence therefore implies that the sequence of partial sums is a *Cauchy sequence* of complex numbers and hence has a limit. The converse is false. The series defined by the sequence $\{a_\nu = \frac{(-1)^\nu}{\nu} : \ \nu \geq 1\}$ is not absolutely convergent since

$$\sum_{\nu=1}^{\infty} \frac{1}{\nu} = \infty$$

On the other hand the alternating series is convergent by the general result in § 4 below.

3. A majorant principle. Let $\{a_\nu\}$ be a bounded sequence and $\{b_\nu\}$ a sequence such that $\sum |b_\nu| < \infty$. Then it is obvious that

$$\sum |a_\nu \cdot b_\nu| < \infty$$

4. Alternating series. Let $a_0, a_1, \ldots$ be a sequence of positive real numbers which is strictly decreasing, i.e. $a_0 > a_1 > \ldots$. Assume also that $\lim a_n = 0$. Then one gets a convergent series by taking alternating signs, i.e. the series below converges:

$$\sum_{\nu=0}^{\infty} (-1)^\nu \cdot a_\nu \tag{*}$$

Proof. Even partial sums are expressed by a positive series:

$$S_{2N} = (a_0 - a_1) + \ldots + (a_{2N} - a_{2N-1}) > 0 \quad : \ N \geq 1$$

At the same time one has

$$S_{2N} = a_0 - [(a_1 - a_2) + \ldots + a_{2N-1} - a_{2N}]$$

where the last term is the negative of a positive series. Hence $\{S_{2N}\}$ is a non-increasing sequence of positive real numbers which obviously implies that there exists a limit:

$$\lim_{N \to \infty} S_{2N} = S_*$$

Finally, since $a_n \to 0$ we get the same limit using odd indices and conclude that (*) is convergent.

5. The partial sum formula. Consider two sequences $\{a_\nu\}$ and $\{b_\nu\}$. Set

$$S_N = \sum_{\nu=0}^{\nu=N} a_\nu \quad : \quad T_N = \sum_{\nu=0}^{\nu=N} a_\nu \cdot b_\nu$$

Since $a_\nu = S_\nu - S_{\nu-1}$ it follows that

$$(*) \qquad T_N = \sum_{\nu=0}^{\nu=N} (S_\nu - S_{\nu-1}) \cdot b_\nu = a_0 b_0 + \sum_{\nu=1}^{\nu=N-1} S_\nu \cdot (b_\nu - b_{\nu+1}) + S_N \cdot b_N$$

This formula which resembles partial integration of functions is quite useful.

6. Exercise Let $b_1 \geq b_2 \geq \ldots$ be a non-increasing sequence of positive real numbers. Show that for every finite n-tuple $a_1, \ldots, a_n$ of complex numbers one has the inequality

$$\big|b_1 a_1 + \ldots + b_n a_n\big| \leq b_1 \cdot M \quad \text{where} \quad M = \max_{1 \leq k \leq n} |a_1 + \ldots + a_k|$$

7. Theorem by Abel. *Assume that the partial sums $\{S_N\}$ of $\{a_\nu\}$ is a bounded sequence and that the positive series $\sum |b_\nu - b_{\nu+1}| < \infty$ where $b_\nu \to 0$ as $\nu \to \infty$. Then the series below is convergent*

$$\sum_{\nu=0}^{\infty} a_\nu \cdot b_\nu$$

Proof. By (3) the series $\sum_{\nu=0}^{\infty} S_\nu \cdot (b_\nu - b_{\nu+1})$ is absolutely convergent and by the hypothesis we also have $S_N \cdot b_N \to 0$ as $N \to \infty$. Hence (*) in (5) shows that $\{T_N\}$ has a limit T_* expressed by the absolutely convergent series

$$T_* = a_0 b_0 + \sum_{\nu=1}^{\infty} S_\nu \cdot (b_\nu - b_{\nu+1})$$

Next, let $\{b_\nu(p)\}$ be a doubly indexed sequence of non-negative numbers which satisfies

$$\nu \mapsto b_\nu(p) \quad \text{is non-increacing for each} \quad p = 0, 1, \ldots$$

$$\lim_{p \to \infty} b_\nu(p) = 1 \quad \text{for each} \quad \nu = 0, 1, \ldots$$

8. Theorem. *For each convergent series $\sum a_\nu$ it follows that*

$$\sum_{\nu=0}^{\infty} a_\nu \cdot b_\nu(p)$$

converges and one has the limit formula

$$\lim_{p \to \infty} \sum_{\nu=0}^{\infty} a_\nu \cdot b_\nu(p) = \sum_{\nu=0}^{\infty} a_\nu$$

9. Exercise. Prove this result. The hint is to employ the partial sum formula. We finish with another useful result.

10. Theorem *Let $\{a_\nu\}$ be a non-decreasing sequence of real numbers such that*

$$(*) \qquad \sum_{k=1}^{\infty} 2^{-k}(a_{k+1} - a_k) < \infty$$

Then

$$(**) \qquad \lim_{k \to \infty} 2^{-k} a_k \to 0$$

Proof. Let S_N denote partial sums of (*). If $M > N$ the partial summation formula gives:

$$(\text{i}) \qquad S_M - S_N = 2^{-M} a_{M+1} - 2^{-N} a_N + \sum_{\nu=N+1}^{M} 2^{-\nu} a_\nu$$

By the assumption $a_N \leq a_\nu$ when $\nu > N$ which gives:

$$S_M - S_N \geq 2^{-M} a_{M+1} - 2^{-N} a_N + a_N \cdot \sum_{\nu=N+1}^{\nu=M} 2^{-\nu} =$$

$$2^{-M}a_{M+1} - 2^{-N}a_N + a_N \cdot 2^{-N}(1 - 2^{-M+N}) = 2^{-M}a_{M+1} - a_N \cdot^{-M}$$

Now we can argue as follows. Since (*) holds the partial sums is a Cauchy sequence. So if $\epsilon > 0$ there exists N_* such that $S_M < S_{N_*} + \epsilon$ for every $M > N_*$. With $N = N_*$ above we therefore get

(iii) $$2^{-M}a_{M+1} \leq \epsilon + a_{N_*} \cdot 2^{-M} \quad : \quad M > N_*$$

With N_* fixed we find a large M such that (iii) entails

$$2^{-M-1}a_{M+1} \leq 2 \cdot \epsilon$$

Since ϵ is arbitrary we get the limit (**).

B. Counting functions.

A counting function $N(s)$ is an integer valued function with jumps at some strictly increasing sequence $0 < s_1 < s_2 < \ldots$. Suppose that $N(s)$ is defined for $0 < s < 1$ and assume that:

(*) $$\int_0^1 (1 - s) \cdot dN(s) < \infty$$

1.B Theorem. When (*) holds it follows that

(**) $$\lim_{s \to 1} (1 - s)N(s) = 0$$

Proof. Put $a_k = N(1-2^{-k})$. The interval $[0, 1]$ can be divided into the intervals $[1-2^{-k}, 1-2^{-k-1})$. It is easily seen that (*) implies that

$$\sum 2^{-k}(a_{k+1} - a_k) < \infty$$

Hence Theorem 7 gives

(i) $$\lim_{k \to \infty} 2^{-k}N(1 - 2^{-k}) = 0$$

Next, if $s < 1$ we choose k such that $1 - 2^{-k} \leq s < 1 - 2^{-k-1}$ and then

(ii) $$(1 - s)N(s) \leq 2^{-k} \cdot N(1 - 2^{-k-1}) = 2 \cdot 2^{-k-1} \cdot N(1 - 2^{-k-1})$$

Hence (i) implies that (**) tends to zero as required.

2.B A study of $\sum(1 - a_k)$

.

Let $0 < a_k < 1$ and suppose that the series

(1) $$\sum(1 - a_k) < \infty$$

Let $N(s)$ be the counting function with jumps at $\{a_k\}$. So (1) means that

(2) $$\int_0^1 (1 - s)dN(s) < \infty$$

With $s = 1 - \xi$ close to 1, i.e. when ξ is small the Taylor expansion of the Log-function at $s = 1$ gives

$$\log \frac{1}{s} = \log \frac{1}{1 - \xi} = \xi + \text{higher order terms in } \xi$$

Using this it is easily seen that (2) holds if and only if

(3) $$\int_0^1 \log(\frac{1}{s}) \cdot dN(s) < \infty$$

Next, for each $0 < r < 1$ we set

(4) $$S(r) = \int_0^r \log \frac{1}{s} \cdot dN(s) \quad \text{and} \quad T(r) = \int_0^r \log \frac{r}{s} \cdot dN(s)$$

Since $\log(\frac{1}{s}) - \log(\frac{r}{s}) = \log\frac{1}{r}$ it follows that

$$S(r) - T(r) = \log\frac{1}{r}\int_0^r dN(s) = \log\frac{1}{r}\cdot N(r) \tag{5}$$

Now $\log\frac{1}{r} \simeq 1-r$ as $r \to 1$ and since (2) above is assumed, it follows from Theorem 1.B that we have:

$$\lim_{r\to 1} \log\frac{1}{r}\cdot N(r) = 0$$

Hence (5) gives

$$\lim_{r\to 1} S(r) - T(r) = 0 \tag{6}$$

With the notations above we have therefore proved

3.B Theorem. *Assume that (2) holds. Then the following limit formula holds*

$$\lim_{r\to 1}\int_0^r \log\big(\frac{r}{s}\big)\cdot dN(s) = \int_0^1 \log\big(\frac{1}{s}\big)\cdot dN(s)$$

4.B Remark. By the multiplicative property of the Log-function the last term becomes

$$\sum \log\frac{1}{a_k} = \log\prod\frac{1}{a_k} \tag{i}$$

In the right hand side there appears the infinite product series defined by the a-sequence. In section III we study product series in more detail but already here we have seen an example of the interplay between additive series and product series. Notice also that via the equivalence of (1) and (2) above, Theorem 3.B gives the following:

5.B Theorem. *Let $\{a_k\}$ be a sequence with each $0 < a_k < 1$. Then the additive series $\sum(1-a_k)$ is convergent if and only if the product series*

$$\prod_{k=1}^{\infty}\frac{1}{a_k} < \infty \tag{i}$$

Moreover, when (i) holds one has the limit formula

$$\lim_{r\to 1}\prod_r \frac{r}{a_k} = \prod_{\nu=1}^{\infty}\frac{1}{a_k} < \infty \tag{ii}$$

where $\prod_r$ is extended over those k for which $a_k \leq r$.

6.B Asymptotic formulas. Let $\{\lambda_\nu\}$ be a strictly increasing sequence of positive numbers where $\lambda_n = +\infty$ as $n \to \infty$. Consider another sequence of positive numbers $\{a_\nu\}$ and assume that the series

$$\sum_{\nu=1}^{\infty}\frac{a_\nu}{\lambda_\nu} < \infty$$

For each $s > 0$ we denote by $\Lambda(s)$ be the largest integer ν such that $\lambda_\nu \leq s$. So the Λ-function is a non-decreasing integer-valued function with jumps at every λ_ν. Next, define the following pair of functions when $0 < x < \infty$:

$$\mathcal{A}(x) = \sum_{\nu<\Lambda(x)} a_\nu \quad\text{and}\quad f(x) = \sum_{\nu=1}^{\infty}\frac{a_\nu}{\lambda_\nu + x}$$

Notice that $\mathcal{A}(x)$ is non-decreasing while $f(x)$ is decreasing. With these notations the following hold:

7.B Theorem. *Assume that there exists some $0 < \alpha < 1$ and a positive constant C such that*

$$\lim_{x\to\infty} x^\alpha\cdot f(x) = C$$

Then there also exists the limit

$$\lim_{x\to\infty} x^{\alpha-1}\cdot\mathcal{A}(x)=\frac{C}{\pi}\cdot\frac{\sin\pi\alpha}{1-\alpha}$$

Remark. The result above is due to Carleman from his lectures at Institute Mittag-Leffler in 1935. The proof requires analytic methods based upon Fourier transforms and is given in § XX from Special Topics. In the special case when $\lambda_\nu=\nu$ we see that $\mathcal{A}(n)$ is equal to the partial sum S_n of $\{a_\nu\}$ for positive integers n. So in this case the theorem above asserts that if the positive series

$$\sum\frac{a_\nu}{\nu}<\infty$$

and if there exists

$$\lim_{x\to\infty} x^{\alpha}\cdot\sum\frac{a_\nu}{\nu+x}=C$$

then

$$\lim_{n\to\infty}\frac{n^\alpha\cdot S_n}{n}=\frac{C}{\pi}\cdot\frac{\sin\pi\alpha}{1-\alpha}$$

II. Power series.

Starting with a sequence $\{a_\nu\}$ and a complex number $z \neq 0$ we get the sequence $\{a_\nu \cdot z^\nu\}$. If this sequence yields a convergent additive series the sum is denoted by $S(z)$.

1. Definition *The set of all $z \in \mathbf{C}$ for which the series*

$$\sum_{\nu=0}^{\infty} a_\nu \cdot z^\nu$$

converges is denoted by $\mathfrak{conv}(\{a_\nu\})$ and called the set of convergence for the a-sequence.

Remark. It may occur that $\mathfrak{conv}(\{a_\nu\})$ just contains $z = 0$. An example is when $a_\nu = \nu\,!$. But if the absolute values $|a_\nu|$ do not increase too fast the domain of convergence is non-empty. Since the terms of a convergent sequence must be uniformly bounded there exists a constant M such that

$$(1) \qquad |a_\nu| \cdot |z_0|^\nu \leq M \quad : \quad \nu = 0, 1, \ldots \quad : \; z_0 \in \mathfrak{conv}(\{a_\nu\})$$

If $|z| < |z_0|$ it follows that the series defined by $\{a_\nu \cdot z^\nu\}$ is absolutely convergent. Indeed, we have

$$|a_\nu \cdot z^\nu| \leq M \cdot \frac{|z|^\nu}{|z_0|^\nu}$$

Here $r = \frac{|z|}{|z_0|} < 1$ and the geometric series $\sum r^\nu$ is convergent. Hence the *Majorant principle* from I.3 yields the the absolute convergence of $\{a_\nu \cdot z^\nu\}$.

2. The radius of convergence. Above we saw that if $z_0 \in \mathfrak{conv}(\{a_\nu\})$ then the domain of convergence contains the open disc of radius $|z_0|$. Put

$$(*) \qquad \mathfrak{r} = \max |z| \quad : \quad z \in \mathfrak{conv}(\{a_\nu\})$$

Assume that $\mathfrak{conv}(\{a_\nu\})$ is not reduced to $z = 0$. Then $\mathfrak{r}$ is a positive number or $+\infty$. It is called the radius of convergence for $\{a_\nu\}$). The case $\mathfrak{r} = +\infty$ means that the series

$$\sum a_\nu \cdot z^\nu$$

converges for all $z \in \mathbf{C}$.

3. Hadamard's formula for $\mathfrak{r}$. Given a sequence $\{a_\nu\}$ its radius of convergence is found by taking a limes superior. More precisely

$$(*) \qquad \frac{1}{\mathfrak{r}} = \limsup_{\nu \to \infty} \; |a_\nu|^{\frac{1}{n}}$$

The proof of this wellknown result is left to the reader.

Example. A sufficient condition in order that $\mathfrak{r} \geq 1$ for a given sequence $\{a_\nu\}$ can be checked as follows. Suppose that

$$|a_\nu| \leq e^{\rho(\nu)}$$

for some sequence $\{\rho(\nu)\}$. With $r < 1$ we can write $r = e^{-\delta}$ for some $\delta > 0$ and obtain

$$|a_n| \cdot r^n \leq \exp(\rho(n) - \delta \cdot n) \quad : \; n = 1, 2, \ldots$$

From this we conclude that $\mathfrak{r} \geq 1$ holds if

$$(**) \qquad \lim_{n \to \infty} \rho(n) - \delta \cdot n = -\infty \quad \text{for each} \quad \delta > 0$$

4. Application. Let $\sum a_n \cdot z^n$ be a power series whose radius of convergence is one. Let $\{b_n\}$ be some other sequence of complex numbers. We seek for conditions in order that the series $\sum b_n a_n \cdot z^n$ also converges when $|z| < 1$. The result below gives a sufficient condition for this to hold.

5. Theorem. *Let $\{\gamma_n\}$ be a sequence of positive numbers such that*

(i) $$\lim_{n\to\infty} \frac{\gamma_n}{n} \cdot \log(n) = 0$$

Then the $\mathfrak{r}$-number of $\{b_\nu \cdot a_\nu\}$ is ≥ 1 for every b-sequence such that

$$|b_n| \leq n^{\gamma_n} \quad : \quad n = 1, 2, \ldots$$

6. Exercise. Prove this theorem. It applies in particular when $\gamma_n = k$ for some positive integer k and hence the radius of convergence of $\{\nu^k \cdot a_\nu\}$ is at least one. Of course, this can be seen directly from the formula in (3) above since

$$\lim_{n\to\infty} n^{\frac{k}{n}} = 1$$

hold for every positive integer k.

7. Hadamard's Lemma. Let $\{c_n\}$ be a sequence if numbers such that the following two conditions hold:

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

There exists some $0 < \alpha < 1$ such that

(ii) $$\left|c_{n+1}^2 - c_{n+2} \cdot c_n\right| \leq \alpha^n \quad : \quad n = 1, 2, \ldots$$

Show that (i-ii) imply that one has an unrestricted limit:

(*) $$\lim_{n\to\infty} |c_n|^{\frac{1}{n}} = 1$$

8. Exercise. Let $a_0, a_1, \ldots$ be a sequence of positive real numbers. Suppose there exists an integer m and a constant C such that

$$a_k \leq \frac{a_{k-1} + \ldots + a_{k-m}}{k} \quad \text{for all} \quad k \geq m$$

Show that no matter how $a_0, \ldots, a_{m-1}$ are determined initially it follows that the power series

$$\sum a_\nu \cdot z^\nu$$

has an infinite radius of convergence, i.e. for every $R > 0$ the positive series $\sum a_\nu \cdot R^\nu < \infty$.

II.B Convergence at the boundary

Let $\{a_\nu\}$ be a sequence with $\mathfrak{r} = 1$. If $e^{i\theta}$ is a complex number whose absolute value is one it is not always true that the series

(1) $$\sum_{\nu=}^{\infty} a_\nu \cdot e^{i\nu\theta}$$

converges. So we have a possibly empty subset of $[0, 2\pi]$ defined by

(2) $$\mathcal{F} = \{0 \leq \theta \leq 2\pi\} \quad : \quad \text{The series (1) converges for } \theta$$

1. Example Let $\{a_\nu = \frac{1}{\nu}\}$. Here $\mathfrak{r} = 1$ and the series $\sum \frac{1}{\nu}$ is divergent. On the other hand

$$\sum \frac{e^{i\nu\theta}}{\nu}$$

converges for many θ. In fact we have

(i) $$\mathcal{F} = (0, 2\pi)$$

To see this we notice that if $b_\nu = e^{i\nu\theta}$ with $0 < \theta < 2\pi$ then the partial sums are:

$$S_N = \frac{1 - e^{i(N+1)\theta}}{e^{i\theta} - 1}$$

This sequence is bounded and since the positive series $\sum(\frac{1}{\nu}-\frac{1}{\nu+1})$ converges, the reader can deduce the inclusion (i) from Abel's theorem in A.7

2. Radial limits Let $\{a_\nu\}$ be a sequence whose radius of convergence is 1. If $0<r<1$ and $0\leq\theta\leq 2\pi$ we get the convergent series

$$S(r,\theta)=\sum_{\nu=0}^\infty a_\nu\cdot r^\nu e^{i\nu\theta}$$

Keeping θ fixed we say that one has a radial limit if there exists

$$(*)\qquad \lim_{r\to 1} S(r,\theta)=S_*(\theta)$$

Denote by Let $\mathfrak{rad}(\{a_\nu\})$ the set of θ for which the limit above exists. The question arises if $\theta\in\mathfrak{rad}(\{a_\nu\})$ implies that the series $\sum a_\nu e^{i\nu\theta}$ converges. This is not true in general. The simplest example is to take $a_\nu=(-1)^\nu$ and $\theta=1$. Here $S(r,0)=\frac{1}{1+r}$ whose limit is $\frac{1}{2}$ while $\sum a_\nu$ must diverge since the a-sequence does not tend to zero. But the converse is true, i.e. one has:

3. Theorem *Let $\{a_\nu\}$ give a convergent additive series with sum S_*. Then there exists the limit*

$$\lim_{x\to 1}\sum a_n\cdot x^n$$

Moreover, the radial limit is equal to the series sum S_* of the additive series.

Proof. We can always modify a_0 and assume that $S_*=0$. Set

$$\text{(i)}\qquad \rho_N=\max_{\nu\geq N}|S_\nu|$$

So the hypothesis is now that $\rho_N\to 0$ as $N\to\infty$. For each $0<x<1$ we set:

$$S_N(x)=\sum_{nu=0}^{\nu=N} a_\nu\cdot x^\nu$$

When $0<x<1$ is fixed the infinite power series

$$\text{(ii)}\qquad S_*(x)=\sum_{\nu=0}^\infty a_\nu\cdot x^\nu$$

converges. Next, when $0<x<1$ then the sequence $\{b_\nu=x^\nu-x^{\nu+1}\}$ is non-increasing. Hence Exercise 6 from [Additive Series] implies that that for each pair $M>N$ and every $0<x<1$ one has

$$|S_M(x)-S_N(x)|\leq\rho_N$$

Since this holds for every $M>N$ and the series (ii) converges we obtain

$$\text{(iii)}\qquad [S_*(x)-S_N(x)|\leq\rho_N$$

Next, the triangle inequality gives:

$$|S_*(x)|\leq|S_*(x)-S_N(x)|+|S_N(x)-S_N|+|S_N|\leq$$
$$2\cdot\rho_N+|S_N(x)-S_N|$$

Finally, if $\epsilon>0$ we first choose N so that $2\cdot\rho_N<\epsilon/2$ and with N fixed we have

$$\lim_{x\to 1} S_N(x)=S_N$$

This proves the requested limit formula

$$\lim_{x\to 1} S_*(x)=0$$

4. A theorem of Landau.

One can also study limits on sparse sets which converge to a boundary point. Results of this nature appear in the article *Über die Konvergenz einiger Klassen von unendlichen Reihen am Rande des Konvergenzgebietes* by Landau from 1907. Here we announce and prove one of these results. Consider a sequence of complex numbers $\{z_k\}$ in the open unit disc D which converge to 1. We say that the sequence is of Landau type if there exists a constant $\mathbf{L}$ such that

(i) $$\frac{|1-z_k|}{1-|z_k|} \leq \mathbf{L} \quad : \quad \frac{1}{\mathbf{L}} \leq k \cdot |1-z_k| \leq \mathbf{L} \quad : \ k = 0,1,2,\ldots$$

Remark. The first inequality means that z_k come close to the real axis as $|z_k| \to 1$. The second condition means that the sequence of absolute values $1-|z_k|$ decreases in a regular fashion.

Theorem. *Let $\{a_\nu\}$ be a sequence such that $\nu \cdot a_\nu \to 0$ as $\nu \to +\infty$ and suppose there exists a sequence $\{z_k\}$ of the Landau type such that there exists a limit*

$$\lim_{k\to\infty} \sum a_\nu \cdot z_k^\nu = A$$

Then the series $\sum a_\nu$ is convergent and the series sum is equal to A.

Proof. Since $\nu \cdot a_\nu \to 0$ it follows that

(i) $$\lim_{k\to\infty} \frac{1}{k} \cdot \sum_{\nu=1}^{k} a_\nu = 0$$

Next, set

(ii) $$f(k) = \sum_{\nu=1}^{\nu=k} a_\nu z_k^\nu \quad \text{and} \quad S_k = \sum_{\nu=1}^{\nu=k} a_\nu$$

The triangle inequality gives

$$|S_k - f(k)| \leq |\sum_{\nu=1}^{\nu=k} a_\nu(1-z_k^\nu) - \sum_{\nu>k} a_\nu z_k^\nu| \leq$$

(iii) $$\sum_{\nu=1}^{\nu=k} |a_\nu|(1-z_k| \cdot \nu + \sum_{\nu>k} |a_\nu| \cdot |z_k|^\nu = W(k)_* + W(k)^*$$

Put

(iv) $$\epsilon(k) = \max\{\nu \cdot |a_\nu| \colon \ \nu \geq k+1\} \implies |a_\nu| \leq \frac{\epsilon(k)}{k} \quad : \ \nu \geq k+1$$

Since we also have $|z_k|^{k+1} \leq 1$ it follows from (iv) that

(v) $$W^*(k) \leq \frac{\epsilon(k)}{k} \cdot \frac{1}{1-|z_k|} \leq \frac{\mathbf{L} \cdot \epsilon(k)}{k \cdot |1-z_k|} \leq \mathbf{L}^2 \cdot \epsilon(k)$$

At the same time we have

(vi) $$W_*(k) \leq k \cdot |1-z_k| \cdot \frac{\sum_{\nu=1}^{\nu=k} \nu \cdot |a_\nu|}{k} \leq \mathbf{L} \cdot \frac{\sum_{\nu=1}^{\nu=k} \nu \cdot |a_\nu|}{k}$$

Now we are done, i.e. $W_*(k) \to 0$ by the observation in (*) and $W^*(k) \to 0$ since the hypothesis on $\{a_\nu\}$ gives $\epsilon(k) \to 0$.

III. Product series

Consider a sequence of positive real numbers $\{q_\nu\}$. To each $N \geq 1$ we define the partial product

$$\Pi_N = \prod_{\nu=1}^{\nu=N} q_\nu$$

If $\lim_{N\to\infty} \Pi_N$ exists we say that the infinite product converges and put

$$\Pi_* = \prod_{\nu=1}^{\infty} q_\nu$$

It is obvious that if the product converges then $\lim_{\nu\to\infty} q_\nu = 1$. The function $\log r$ has the Taylor expansion close to $r = 1$ given by $\log r = (r-1) + (r-1)^2/2 + \ldots$. Using this one gets following:

1. Theorem. *Let $\{q_\nu\}$ be a sequence where $0 < q_\nu < 1$ hold for all ν. Then the following three conditions are equivalent:*

$$\sum (1 - q_\nu) < \infty \quad : \quad \sum \operatorname{Log} \frac{1}{q_\nu} < \infty \quad : \quad \prod_{\nu=1}^{\infty} q_\nu > 0$$

Exercise Prove Theorem 1.

Next, when $|z| < 1$ the complex log-function has the series expansion

$$\log(1 + z) = z - z^2/2 + z^3/3 + \ldots$$

From this one easily gets

2. Proposition *One has the inequality*

$$|\operatorname{Log}(1 + z) - z| \leq |z|^2 \quad : \; |z| \leq 1/2$$

Next, consider a complex sequence $a(\cdot)$ where $|a_\nu| \leq \frac{1}{2}$ hold for all ν and put:

$$\Pi_N = \prod_{\nu=1}^{\nu=N} (1 - a_\nu) \implies \log(\Pi_N) = \sum_{\nu=0}^{\nu=N} \log(1 - a_\nu)$$

Proposition 2 gives the inequality

$$(*) \qquad |\log(1 - a_\nu) + a_\nu| \leq |a_\nu|^2$$

This enable us to investigate the convergence of the product series with the aid of the additive series for $\{a_\nu\}$. We get for example

$$(**) \qquad |\log(\Pi_N) + \sum_{\nu=1}^{\nu=N} a_\nu| \leq \sum_{\nu=1}^{\nu=N} |a_\nu|^2$$

From (**) we can conclude:

3. Theorem. *Let $\{a_\nu\})$ be a sequence where each $|a_\nu| \leq \frac{1}{2}$ and $\sum |a_\nu|^2 < \infty$. Then $\sum a_\nu$ converges if and only if the product series $\Pi(1 - a_\nu)$ converges. Moreover, when convergence holds one has the equality*

$$\log(\Pi_*) = \sum_{\nu=1}^{\infty} \log(1 - a_\nu)$$

IV. Blaschke products.

Let $\{a_\nu\}$ be a sequene in the open unit disc D which are enumerated so that their absolute values are non-decreasing. But repetitions may occur, i.e. several a-numbers can be equal. We always assume that $|a_\nu| \to 1$ as $\nu \to +\infty$. Hence $\{a_\nu\}$ is a discrete subset of D. To each ν we set

$$(1) \qquad \beta_\nu(\theta) = \frac{e^{i\theta} - a_\nu}{1 - e^{i\theta} \cdot \bar{a}_\nu} \cdot \frac{\bar{a}_\nu}{|a_\nu|} \quad : \quad 0 \leq \theta \leq 2\pi$$

The *Blaschke product of order* N is the partial product

$$(2) \qquad B_N(\theta) = \prod_{\nu=1}^{\nu=N} \beta_\nu(\theta)$$

The question arises when the product series converges and gives a limit

$$(3) \qquad B_*(\theta) = \prod_{\nu=1}^{\infty} \beta_\nu(\theta)$$

To analyze this we use polar coordinates and put

$$a_\nu = r_\nu e^{i\theta\nu}$$

Each β-number has absolute value one and if $\theta \neq \theta_\nu$ for every ν we have

$$(4) \qquad \beta_\nu(\theta) = e^{i \cdot \gamma(r_\nu, \theta - \theta_\nu)} \quad : \quad 0 < \gamma(r_\nu, \theta - \theta_\nu) < 2\pi$$

Exercise. Show that when $-\pi/2 < \theta - \theta_\nu < \pi/2$ then the construction of the arctan-function gives

$$(4) \qquad \gamma(r, \theta - \theta_\nu) = \text{arctg}\,[\frac{(1 - r^2) \cdot \sin(\theta_\nu - \theta)}{1 + r^2 - 2r\cos(\theta_\nu - \theta)}\,]$$

4.2. Blashke's condition We impose the condition that the positive series

$$(*) \qquad \sum (1 - r_\nu) < \infty$$

Later we shall consider the analytic function defined in $|z| < 1$ by

$$(2) \qquad B(z) = \prod_{\nu=0}^{\infty} \frac{z - a_\nu}{1 - \bar{a}_\nu \cdot z} \cdot e^{-i\arg(a_\nu)}$$

A major result to be proved later on asserts that the Blaschke condition implies that the radial limit

$$\lim_{r \to 1} B(re^{i\theta}) = B_*(\theta)$$

exists almost everywhere. Moreover, the absolute value of the limit $B_*(\theta)$ is equal to one almost everywhere. But the determination of the set of all $0 \leq \theta \leq 2\pi$ for which the radial limit exists is not clear since no special assumption is imposed on the $\{\theta_\nu\}$-sequence. For example, divergence may appear when many θ_ν:s are close to θ even if $\{r_\nu\}$ tend rapidly to 1.

4.3 Exercise. Let $\{r_\nu\}$ be given where the positive series in 4.2 converges. Next, if x is a real number we set

$$\{x\} = \min_{k \in \mathbf{Z}} [x - 2\pi k|$$

Show that the Blaschke product has a radial limit at $\theta = 0$ if and only if there exists the limit

$$(5) \qquad \lim_{N \to \infty} \{ \sum_{\nu=1}^{\nu=N} \frac{(1 - r_\nu) \cdot \theta_\nu}{(1 - r_\nu)^2 + \theta_\nu^2} \}$$

Notice that θ_ν may be < 0 or > 0 and it is not necessary that all of them become close to 0. To determine all sequence of pairs (r_ν, θ_ν) where 4.1 holds and $\theta_\nu \to 0$ appears to be a very difficult problem.

V. Estimates using the counting function.

We establish some inequalities which will be used to study entire functions of exponential type in XXX. Let $\{\alpha_\nu\}$ be a complex sequence where $0 < |\alpha_1| \leq |\alpha_2| \leq \ldots\}$. This time the absolute values tend to $+\infty$. We get the counting function $N(R)$ which for every $R > 0$ is the number of α_ν with absolute value $\leq R$. Consider the situation when there exists a constant C such that

$$(*) \qquad N(R) \leq C \cdot R \quad \text{for all} \quad R > 0$$

5.1. The first estimate. To each $R > 0$ we set

$$(2) \qquad S(R) = \prod (1 + \frac{R}{|\alpha_\nu|}) \quad : \text{product taken over all } |\alpha_\nu| \leq 2R$$

Then we have

$$(*) \qquad S(R) \leq e^{KR} \quad \text{where} \quad K = 2C(1 + \operatorname{Log} \frac{3}{2})$$

To prove this we consider log $S(R)$. A partial integration gives:

$$\log\, S(R) = \int_0^{2R} \log\,(1 + \frac{R}{t}) \cdot dN(t) = \log\,(1 + \frac{1}{2}) \cdot N(2R) + \int_0^{2R} \frac{R \cdot N(t)}{t(t+R)} \cdot dt$$

Since $\frac{R}{t+R} \leq 1$ for all t, the last integral is estimated by $2R \cdot C$ and (*) follows.

5.2. The second estimate. For each $R > 0$ we consider infinite tail products:

$$(i) \qquad S^*(R) = \prod (1 + \frac{R}{\alpha_\nu}) \cdot e^{-\frac{R}{\alpha_\nu}} \quad : \text{product taken over all } |\alpha_\nu| \geq 2R$$

To estimate (i) we notice that the analytic function $(1 + \zeta)e^{-\zeta} - 1$ has a double zero at the origin. This gives a constant A such that

$$(ii) \qquad |(1 + \zeta)e^{-\zeta} - 1| \leq A \cdot |\zeta|^2 \quad : \ |\zeta| \leq \frac{1}{2}$$

Since $|\alpha_\nu| \geq 2R$ for every ν we obtain:

$$(iii) \qquad \log^+ |(1 + \frac{R}{\alpha_\nu}) \cdot e^{-\frac{R}{\alpha_\nu}}| \leq \log\,[1 + A\frac{R^2}{|\alpha_\nu|^2}) \leq A \cdot \frac{R^2}{|\alpha_\nu|^2}$$

From (6) we get

$$(iv) \qquad \log^+(S^*(R)) \leq AR^2 \int_{2R}^{\infty} \frac{dN(t)}{t^2} = A \cdot N(2R) + 2AR^2 \cdot \int_{2R}^{\infty} \frac{N(t)}{t^3}$$

The last term is estimated by

$$(8) \qquad 2AR^2 \cdot C \cdot \int_{2R}^{\infty} \frac{dt}{t^2} = AC \cdot R$$

Adding up the result we get

5.3 Theorem. *One has the inequality*

$$S^*(R) \leq \frac{5A}{4} \cdot C \cdot R$$

VI. Theorems by Abel, Tauber, Hardy and Littlewood

Introduction. Consider a power series $f(z) = \sum a_n z^n$ whose radius of convergence is one. If $r < 1$ and $0 \leq \theta \leq 2\pi$ we are sure that the series

$$f(re^{i\theta}) = \sum a_n r^n e^{in\theta}$$

is convergent. In fact, it is even absolutely convergent since the assumption implies that

$$\sum |a_n| \cdot r^n < \infty \quad \text{for all} \quad r < 1$$

Passing to $r = 1$ it is in general not true that the series $\sum a_n e^{in\theta}$ is convergent. An example arises if we consider the geometric series

$$\frac{1}{1-z} = 1 + z + z^2 + \ldots$$

So here $a_n = 1$ for all n and hence the series $\sum a_n$ is divergent. At the same time we notice that when $0 < \theta < 2\pi$ there exists the limit

$$\lim_{r \to 1} \sum r^n e^{in\theta} = \frac{1}{1 - re^{i\theta}}$$

while the series $\sum a_n e^{in\theta}$ is divergent. This leads to the following problem where we without loss of generality can take $\theta = 0$. Consider as above a convergent power series and assume that there exists the limit

$$(*) \qquad \lim_{r \to 1} \sum a_n r^n$$

When can we conclude that the series $\sum a_n$ also is convergent and that one has the equality

$$(**) \qquad \sum a_n = \lim_{r \to 1} \sum a_n r^n$$

The first result in this direction was established by Abel in a work from 1823:

A. Theorem *Let $\{a_n\}$ be a sequence such that $\frac{a_n}{n} \to 0$ as $n \to \infty$ and there exists*

$$A = \lim_{r \to 1} \sum a_n r^n$$

Then $\sum a_n$ is convergent and the sum is A.

An extension of Abel's result was established by Tauber in 1897.

B. Theorem. *Let $\{a_n\}$ be a sequence of real numbers such that there exists the limit*

$$A = \lim_{r \to 1} \sum a_n r^n$$

Set

$$\omega_n = a_1 + 2a_2 + \ldots + na_n \quad : \; n \geq 1$$

If $\lim_{n \to \infty} \omega_n = 0$ it follows that the series $\sum a_n$ is convergent and the sum is A.

C. Results by Hardy and Littlewood.

.

In their joint article *xxx* from 1913 the following extension of Abel's result was proved by Hardy and Littlewood:

C Theorem. *Let $\{a_n\}$ be a sequence of real numbers such that there exists a constant C so that $\frac{a_n}{n} \leq C$ for all $n \geq 1$. Assume also that the power series $\sum a_n z^n$ converges when $|z| < 1$. Then the same conclusion as in Abel's theorem holds.*

Remark. The proof of Theorem C requires several steps where an essential ingredient is a result about positive series from the cited article which has independent interest.

D. Theorem. *Assume that each $a_n \geq 0$ and that there exists the limit:*

(*) $$A = \lim_{r \to 1} (1-r) \cdot \sum a_n r^n$$

Then there exists the limit

(**) $$A = \lim_{N \to \infty} \frac{a_1 + \ldots + a_N}{N}$$

Notice that we do not impose any growth condition on $\{a_n\}$ above,i.e. the sole assumption is the existing limit (*).

Remark. The proofs of Abel's and Tauber's results are quite easy while C and D require more effort and we need some results from calculus in one variable. So before we enter the proofs of the theorems above insert some preliminaries.

1. Results from calculus

Below $g(x)$ is a real-valued function defined on $(0,1)$ and of class C^2 at least.

1.1 Lemma *Assume that there exists a constant $C > 0$ such that*

$$g''(x) \leq C(1-x)^{-2} \quad : \ 0 < x < 1 \quad \text{and} \quad \lim_{x \to 1} g(x) = 0$$

Then one has the limit formula:

$$\lim_{x \to 1} (1-x) \cdot g'(x) = 0$$

1.2 Lemma *Assume that the second order derivative $g''(x) > 0$. Then the following implication holds for each $\alpha > 0$:*

$$\lim_{x \to 1} (1-x)^{\alpha} \cdot g(x) = 1 \implies \lim_{x \to 1} (1-x)^{\alpha+1} \cdot g'(x) = \alpha$$

Remark. If $g(x)$ has higher order derivatives which all are > 0 on $(0,1)$ we can iterate the conclusion in Lemma 1.2 where we take α to be positive integers. More precisely, by an induction over ν the reader may verify that if

$$\lim_{x \to 1} (1-x) \cdot g(x) = 1$$

exists and if $\{g^{(\nu)}(x) > 0\}$ for all every $\nu \geq 2$ then

(*) $$\lim_{x \to 1} (1-x)^{\nu+1} \cdot g^{(\nu)}(x) = \nu\,! \quad : \ \nu \geq 2$$

Next, to each integer $\nu \geq 1$ we denote by $[\nu - \nu^{2/3}]$ the largest integer $\leq (\nu - \nu^{2/3})$. Set

$$J_*(\nu) = \sum_{n \leq [\nu - \nu^{2/3}]} n^{\nu} e^{-\nu} \quad : \quad J^*(\nu) = \sum_{n \geq [\nu + \nu^{2/3}]} n^{\nu} e^{-\nu}$$

1.3 Lemma *There exists a constant C such that*

$$\frac{J^*(\nu) + J_*(\nu)}{\nu\,!} \leq \delta(\nu) \quad : \quad \delta(\nu) = C \cdot \exp\big(-\frac{1}{2} \cdot \nu^{\frac{1}{3}}\big) \quad : \ \nu = 1, 2, \ldots$$

Proofs

We prove only Lemma 1.1 which is a bit tricky while the proofs of Lemma 1.2 and 1.3 are left as exercises to the reader. Fix $0 < \theta < 1$. Let $0 < x < 1$ and set

$$x_1 = x + (1-x)\theta$$

The mean-value theorem in calculus gives

(i) $$g(x_1) - g(x) = \theta(1-x)g'(x) + \frac{\theta^2}{2}(1-x)^2 \cdot g''(\xi) \quad \text{for some} \quad x < \xi < x_1$$

By the hypothesis

$$g''(\xi) \leq C(1-\xi)^{-2} \leq C)1 - x_1)^{-2}$$

Hence (i) gives

$$(1-x)g'(x) \geq \frac{1}{\theta}(g(x_1)-g(x)) - C\cdot\frac{\theta}{2}\frac{(1-x)^2}{1-x_1)^2} =$$

$$\frac{1}{\theta}(g(x_1)-g(x)) - \frac{C\cdot\theta}{2(1-\theta)^2}$$

Keeping θ fixed we have by assumption

$$\lim_{x\to 1} g(x) = 0$$

Notice also that $x \to 1 \implies x_1 \to 1$. It follows that

$$\liminf_{x\to 1}\ (1-x)g'(x) \geq -\frac{C\cdot\theta}{2(1-\theta)^2}$$

Above $0<\theta<1$ is arbitrary, i.e. we can choose small $\theta>0$ and hence we have proved that

(*) $$\liminf_{x\to 1}\,(1-x)g'(x) \geq 0$$

Next we prove the opposed inequality

(**) $$\limsup_{x\to 1}\,(1-x)g'(x) \leq 0$$

To get (**) we apply the mean value theorem in the form

(ii) $$g(x_1)-g(x) = \theta(1-x)g'(x_1) - \frac{\theta^2}{2}(1-x)^2\cdot g''(\eta) \quad : \ x<\eta<x_1$$

Since $(1-x_1) = \theta(1-x)(1-\theta)$ we get

(iii) $$(1-x_1)g'(x_1) = \frac{1-\theta}{\theta}\cdot(g(x_1)-g(x)) + \frac{(1-\theta)\theta}{2}\cdot(1-x)^2 g''(\eta)$$

Now $g''(\eta) \leq C(1-\eta)^{-2} \leq C(1-x_1)^{-2}$ so the right hand side in (iii) is majorized by

$$\frac{1-\theta}{\theta}\cdot(g(x_1)-g(x)) + C\cdot\frac{(1-\theta)\theta}{2}\cdot(1-x)^2(1-x_1)^2 =$$

(iv) $$\frac{1-\theta}{\theta}\cdot(g(x_1)-g(x)) + C\cdot\frac{\theta}{2(1-\theta}$$

Keeping θ fixed while $x \to 1$ we obtain:

$$\liminf_{x\to 1}\,(1-x)g'(x) \leq C\cdot\frac{\theta}{2(1-\theta}$$

Again we can choose arbitrary small θ and hence (**) holds which finishes the proof of Lemma 1.1.

2. Proof of Abel's theorem.

Without loss of generality we can assume that $a_0 = 0$. Set

$$S_N = a_1 + \ldots + a_N \quad \& \quad f(r) = \sum a_n r^n$$

where $r<1$. For every positive integer N the triangle inequality gives:

(2.1) $$\big|S_N - f(r)\big| \leq \sum_{n=1}^{n=N} |a_n|(1-r^n) + \sum_{n\geq N+1} |a_n| r^n$$

Set $\delta(N) = \max_{n\geq N} \frac{|a_n|}{n}$. Since $1 - r^n) = (1-r)(1 + \ldots + r^{n-1} \leq (1-r)n$ the right hand side in (2.1) is majorised by

$$(1-r) \cdot \sum_{n=1}^{n=N} n \cdot |a_n| + \delta(N+1) \cdot \sum_{n\geq N+1} \frac{r^n}{n}$$

Next, the obvious inequality $\sum_{n\geq N+1} \frac{r^n}{n} \leq \frac{1}{N+1} \cdot \frac{1}{1-r}$ gives the new majorisation

$$(1-r) \cdot \sum_{n=1}^{n=N} n \cdot |a_n| + \frac{\delta(N+1)}{N+1} \cdot \frac{1}{1-r} \tag{2.2}$$

This hold for all pairs N and r. To each $N \geq 2$ we take $r = 1 - \frac{1}{N}$ and then (2.2) is majorised by

$$\frac{1}{N} \cdot \sum_{n=1}^{n=N} n \cdot |a_n| + \delta(N+1) \cdot \frac{N}{N+1}$$

Here both terms tend to zero as $N \to \infty$. Indeed, Abel's condition $n \cdot a_n \to 0$ implies that both $\delta(N+1)$ and $\frac{1}{N} \cdot \sum_{n=1}^{n=N} n \cdot |a_n| \to$ tend to zero as $N \to \infty$. Hence

$$\lim_{N\to\infty} \big|s_N - f(1 - \frac{1}{N})\big| = 0 \tag{*}$$

Finally it is clear that (*) gives Abel's theorem.

3. Proof of Tauber's theorem.

We may assume that $a_0 = 0$. Notice that

$$a_n = \frac{\omega_n - \omega_{n-1}}{n} \quad : \, n \geq 1$$

It follows that

$$f(r) = \sum \frac{\omega_n - \omega_{n-1}}{n} \cdot r^n = \sum \omega_n \big(\frac{r^n}{n} - \frac{r^{n+1}}{n+1}\big)$$

Using the equality $\frac{1}{n} = \frac{1}{n+1} = \frac{1}{n(n+1)}$ we can rewrite the right hand side as follows:

$$\sum \omega_n \big(\frac{r^n - r^{n+1}}{n+1} + \frac{r^n}{n(n+1)}\big)$$

Set

$$g_1(r) = \sum \omega_n \cdot \frac{r^n - r^{n+1}}{n+1} = (1-r) \cdot \sum \frac{\omega_n}{n+1} \cdot r^n$$

By the hypothesis $\lim_{n\to\infty} \frac{\omega_n}{n+1} = 0$ and then it is clear that we get

$$\lim_{r\to 1} g_1(r) = 0$$

Since we also have $f(r) \to 0$ as $r \to 1$ we conclude that

$$\lim_{r\to 1} \sum \frac{\omega_n}{n(n+1)} \cdot r^n = 0 \tag{1}$$

Next, with $b_n = \frac{\omega_n}{n(n+1)}$ we have $nb_n = \frac{\omega_n}{n+1} \to 0$. Hence Abel's theorem applies so (1) gives convergent series

$$\sum \frac{\omega_n}{n(n+1)} = 0 \tag{2}$$

If $N \geq 1$ we have the partial sum

$$S_N = \sum_{n=1}^{n=N} \frac{\omega_n}{n(n+1)} = \sum_{n=1}^{n=N}, \omega_n \cdot \big(\frac{1}{n} - \frac{1}{n+1}\big)$$

The last term becomes

$$\sum_{n=1}^{n=N}\frac{1}{n}(\omega_n-\omega_{n-1})-\frac{\omega_N}{N+1}=\sum_{n=1}^{n=N}a_n-\frac{\omega_N}{N+1}$$

Again, since $\frac{\omega_N}{N+1}\to 0$ as $N\to\infty$ we conclude that the convergent series from (2) implies that the series $\sum a_n$ also is converges and has sum equal to zero. This finishes the proof of Tauber's result.

4. Proof of Theorem D.

Set $g(x)=\sum s_nx^n$ when $0<x<1$. Notice that

$$(1-x)g(x)=\sum a_nx^n$$

Since $s_n\geq 0$ for all n, the higher order derivatives

$$g^{(p)}(x)=\sum_{n=p}^{\infty}n(n-1)\cdots(n-p+1)s_nx^{n-p}>0$$

when $0<x<1$. The hypothesis (*) and the inductive result in the remark after Lemma 1.2 give:

(1) $$\lim_{x\to 1}(1-x)^{\nu+2}\cdot\sum s_n\cdot n^\nu x^n=(\nu+1)!\quad:\ \nu\geq 1$$

We shall use the substitution $e^{-t}=x$ where $t>0$. Since $t\simeq 1-x$ when $x\to 1$ we see that (1) gives

(2) $$\lim_{t\to 0}t^{\nu+2}\cdot\sum s_n\cdot n^\nu e^{-nt}=(\nu+1)!\quad:\ \nu\geq 1$$

Put

$$J_*(\nu,t)=\frac{t^{\nu+2}}{(\nu+1)!}\cdot\sum_{n=1}^{\infty}s_n\cdot n^\nu e^{-nt}$$

So (2) gives for each fixed ν

(3) $$\lim_{t\to 0}J_*(\nu,t)=1$$

Next, for each pair $\nu\geq 2$ and $0<t<1$ we define the integer

(*) $$N(\nu,t)=[\frac{\nu-\nu^{2/3}}{t}]$$

Since the sequence $\{s_n\}$ is non-decreasing we get

(i) $$s_{N(\nu,t)}\cdot\sum_{n\geq N(\nu,t)}n^\nu e^{-nt}\leq\sum_{n\geq N(\nu,t)}s_n\cdot n^\nu e^{-nt}\leq\frac{(\nu+1)!\cdot J_*(\nu,t)}{t^{\nu+2}}$$

Next, the construction of $N(\nu,t)$ and Lemma 1.3 give:

(ii) $$\sum_{n\geq N(\nu,t)}n^\nu e^{-nt}\geq\frac{\nu!}{t^{\nu+1}}\cdot(1-\delta(\nu))$$

where the δ function is independent of t and tends to zero as $\nu\to\infty$. Hence (i-ii) give

(iii) $$s_{N(\nu,t)}\leq\frac{(\nu+1)}{t}\cdot\frac{1}{1-\delta(\nu)}\cdot J_*(\nu,t)$$

Next, by the construction of N one has

$$N(\nu,t)+1\geq\frac{\nu-\nu^{2/3}}{t}=\frac{\nu}{t}\cdot(1-\nu^{-1/3})$$

It follows that (iii) gives

(iv) $$\frac{s_{N(\nu,t)}}{N(\nu,t)+1} \leq \frac{\nu+1}{\nu} \cdot \frac{1}{1-\nu^{-1/3}} \cdot \frac{1}{1-\delta(\nu)} \cdot J_*(\nu,t)$$

Since $\delta(\nu) \to 0$ it follows that for any $\epsilon > 0$ there exists some ν_* such that

(v) $$\frac{\nu_*+1}{\nu_*} \cdot \frac{1}{1-\nu_*^{-1/3}} \cdot \frac{1}{1-\delta(\nu_*)} < 1+\epsilon$$

Increasing ν_* if necessary, the construction of $N(\nu_*,t)$ gives

$$\big| N(\nu_*,t) - \frac{\nu_*+1}{t}\big| < \epsilon$$

With a fixed ν_* as above, we find for each positive integer N some t_N such that $N = N(\nu_*, t_N)$, and notice that

(vi) $$N \to +\infty \implies t_N \to 0$$

Next, (iv) and (v) yield:

(vii) $$\frac{s_N}{N+1} < (1+\epsilon) \cdot J_*(\nu_*, t_N)$$

Now (vi) and the limit in (3) which applies with ν_* is kept fixed while $t_N \to 0$ entail that

(viii) $$\lim_{N\to\infty} J(\nu_*, t_N) = 1$$

At the same time $\frac{N}{N+1} \to 1$ and since $\epsilon > 0$ was arbitrary we conclude from (vii) that:

(4) $$\limsup_{N\to\infty} \frac{s_N}{N} \leq 1$$

So Theorem 2 follows if we also prove that

(5) $$\liminf_{N\to\infty} \frac{s_N}{N} \geq 1$$

To get (5) we define the integers

$$N(\nu,t) = [\frac{\nu+\nu^{2/3}}{t}] \implies$$

$$S_{N(\nu,t)} \cdot \sum_{n \leq N(\nu,t)} n^\nu e^{-nt} \geq \frac{(\nu+1)! \cdot J_*(\nu,t)}{t^{\nu+2}} - \sum_{n > N(\nu,t)} s_n \cdot n^\nu e^{-nt}$$

The last term can be estimated since (4) gives a constant C such that $s_n \leq Cn$ for all n and then

(6) $$\sum_{n>N(\nu,t)} s_n \cdot n^\nu e^{-nt} \leq C \cdot \sum_{n>N(\nu,t)} n^{\nu+1} e^{-nt} \leq C \cdot \delta(\nu) \cdot \frac{(\nu+1)!}{t^{\nu+2}}$$

where Lemma 1.3 entails that $\delta(\nu) \to 0$ as ν increases. At the same time Lemma 1.3 also gives

(7) $$\sum_{n\leq N(\nu,t)} n^\nu \cdot e^{-nt} = \frac{\nu!}{t^{\nu+1}} \cdot (1-\delta_*(\nu)$$

where $\delta_*(\nu) \to 0$ as $\nu \to +\infty$. Given $\epsilon > 0$ we choose ν_* large so that $C \cdot \delta(\nu_*) < \epsilon$ and $\delta_*(\nu_*) < \epsilon$. Increasing ν_* if necessary, the construction of $N(\nu_*,t)$ gives

$$\big| N(\nu_*,t) - \frac{\nu_*+1}{t}\big| < \epsilon$$

With a fixed ν_* as above, we find for each integer $N \geq 1$ some t_N such that $N = N(\nu_*, t_N)$, and

$$S_{(N(\nu_*,t)} \cdot \frac{\nu_*!}{t^{\nu_*+1}} \cdot (1-\epsilon) \geq \frac{(\nu_*+1)!}{t^{\nu_*+2}} \cdot [J_*(\nu_*,t) - \epsilon] \implies$$

$$S_{(N(\nu_*,t)} \cdot (1-\epsilon) \geq \frac{(\nu_*+1)}{t} \cdot [J_*(\nu_*,t) - \epsilon]$$

Here ϵ csn be arbitary small, and together with (vii) the requested limit in (5) follows.

5. Proof of Theorem C

Set $f(x) = \sum a_n x^n$. Notice that it suffices to prove Theorem C when the limit value

$$\lim_{x \to 1} \sum a_n x^n = 0$$

Next, the assumption that $a_n \leq \frac{c}{n}$ for a constant c gives

$$f''(x) = \sum n(n-1)a_n x^{n-2} \leq c \sum (n-1)x^{n-2} = \frac{c}{1-x)^2}$$

The hypothesis $\lim_{x \to 1} f(x) = 0$ and Lemma xx therefore gives

(i)
$$\lim_{x \to 1} (1-x)f'(x) = 0$$

Next, notice the equality

(ii)
$$\sum_{n=1}^{\infty} \frac{na_n}{c} x^n = \frac{x}{c} \cdot f'(x)$$

At the same time $\sum_{n=1}^{\infty} x^n = \frac{x}{1-x}$ and hence (i-ii) together give:

$$\lim_{x \to 1} (1-x) \cdot \sum (1 - \frac{na_n}{c}) \cdot x^n = 1$$

Here $1 - \frac{na-n}{c} \geq 0$ so Theorem 2 gives

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{n=N} (1 - \frac{na_n}{c}) = 1$$

It follows that

$$\lim_{N \to \infty} \frac{1}{N} \cdot \sum_{n=1}^{n=N} na_n = 0$$

This means precisely that the condition in Tauber's Theorem holds and hence $\sum a_n$ converges and has series sum equal to 0 which finishes the proof of Theorem C.

VII. An example by Hardy

Consider the series expansion of

$$(1-z)^\alpha = \sum b_n z^n$$

where α in general is a complex number. Newton's binomial formula gives:

$$(*) \qquad b_n = \frac{\alpha(\alpha+1)\ldots(\alpha+n-1)}{n\,!}$$

Apply this with $\alpha = i$. If $n \geq 1$ which entails that

$$|n \cdot b_n| = \frac{|(i+1)|\ldots|i+n-1|}{n\,!} = \sqrt{(1+\frac{1}{1^2})\cdot(1+\frac{1}{2^2})\cdots(1+\frac{1}{(n-1)^2}}$$

It follows that

$$\lim_{n\to\infty} |n \cdot b_n| = \sqrt{\prod_{\nu=1}^{\infty}(1+\frac{1}{\nu^2})}$$

In particular $|b_n| \simeq \frac{1}{n}$ when n is large and therefore the series

$$\text{(i)} \qquad \sum_{n=2}^{\infty} \frac{|b_n|}{\operatorname{Log} n} = +\infty$$

In spite of the divergence above one has:

Theorem 7.1. *The series*

$$\sum_{n=2}^{\infty} \frac{b_n}{\operatorname{Log} n} \cdot e^{in\phi}$$

converges uniformly when $0 \leq \phi \leq 2\pi$.

Before we give the proof of Theorem 7.1 we need a result if independent interest.

7.2. Theorem. *Let* $\{a_n\}$ *be a sequence of complex numbers such that* $|a_n| \leq \frac{C}{n}$ *hold for all* n *and some constant* C *and the analytic function* $f(z) = \sum a_n z^n$ *is bounded in* D, *i.e.*

$$|f(z)| \leq M \quad : z \in D$$

Then, if $B_n(z) = a_0 + a_1 + \ldots + a_n z^n$ *are the partial sums one has the inequality*

$$\max_\theta |B_m(e^{i\theta})| \leq M + 2C \quad \text{for all} \quad m = 1, 2, \ldots$$

Proof. When $0 < r < 1$ and θ are given we have

$$\big|B_m(\theta) - f(re^{i\theta})\big| = \sum_{n=0}^{n=m} a_n e^{in\theta}(1-r^n) - \sum_{n=m+1}^{\infty} a_n e^{in\theta} \cdot r^n \leq$$

$$(1-r)\sum_{n=0}^{n=m} n \cdot |a_n| + \sum_{n=m+1}^{\infty} |a_n| \cdot r^n$$

With m given we apply this when $r = 1 - 1/m$. Then the last sum above is estimated above by

$$(*) \qquad \frac{1}{m} \cdot \sum_{n=1}^{n=m} n \cdot |a_n| + \frac{c}{m} \cdot \sum_{n=m+1}^{\infty} r^n \leq C + \frac{c}{m} \cdot \sum_{n=0}^{\infty} r^n = 2C$$

Finally, since the maximum norm of f is $\leq M$ the triangle inequality gives

$$|B_m(e^{i\theta})| \leq 2C + M$$

Here m and θ are arbitrary so Theorem 7.2 follows.

Proof of Theorem 7.1 To each $m \geq 2$ we consider the partial sum series

$$S_m(\phi) = \sum_{n=2}^{n=m} \frac{b_n}{\operatorname{Log} n} \cdot e^{in\phi}$$

Theorem 7.1 follows if there to every $\epsilon > 0$ exists an integer M such that

(1) $$\max_{0 \leq \phi \leq 2\pi} |S_m(\phi) - S_M(\phi)| < \epsilon \quad : \quad \forall\, m > M$$

To prove (1) we employ the partial sums

(2) $$B_n(\phi) = \sum_{\nu=1}^{\nu=n} b_\nu \cdot e^{i\nu\phi}$$

For each pair $m > M \geq 2$, the partial summation formula in gives

$$S_m(\phi) - S_M(\phi) =$$

(3) $$\sum_{n=M}^{n=m} B_n(\phi) \cdot \Big[\frac{1}{\operatorname{Log} n} - \frac{1}{\operatorname{Log}(n+1)}\Big] - \frac{B_{M-1}(\phi)}{\operatorname{Log} M} + \frac{B_m(\phi)}{\operatorname{Log}(m+1)}$$

Now we can apply Theorem 7.2 and find a constant K such that

(4) $$|B_n(\phi)| \leq K \quad >: \quad n \geq 2 \quad \text{and} \quad 0 \leq \phi \leq 2\pi$$

Notice that if (4) holds then (3) gives the inequality

$$\big|S_m(\phi) - S_M(\phi)\big| \leq K \cdot \sum_{n=M}^{n=m} \Big[\frac{1}{\operatorname{Log} n} - \frac{1}{\operatorname{Log}(n+1)}\Big] + \frac{1}{\operatorname{Log} M} + \frac{1}{\operatorname{Log}(m+1)} = \frac{2K}{\operatorname{Log} M}$$

Hence Theorem 7.1 is proved if we establish the inequality (4). To prove this we study the analytic function defined in the open unit disc D by the convergent power series:

(*) $$(1-z)^i = \sum c_n \cdot z^n$$

Since $\mathfrak{Re}(1-z) > 0$ when $|z| < 1$ there exists a single valued branch of $\log(1-z)$ and the function above can be written as

$$g(z) = e^{i \cdot \log(1-z)}$$

Now the argument of $\log(1-z)$ stays in $(-\pi/2, \pi/2$ and we conclude that

$$|g(z)| \leq e^{\pi/2} \quad : z \in D$$

Hence the g-function is bounded in D. Now

$$g(z) = \sum b_n z^n$$

and Newton's binomial formula gives:

$$|b_n| \leq \frac{C}{n} \quad : n \geq 1$$

Then it is clear that Theorem 7.2 applied to g gives (4) above.

$0 < b < 1$ we can expand f around b and obtain another series

(2) $$f(b+z) = \sum c_n \cdot z^n$$

From the convergence of $\sum a_k$ one expects that the series

(3) $$\sum c_n \cdot (1-b)^n$$

also is convergent. This is indeed true and was proved by Hardy and Littlewood in (H-L]. A more general result was established in [Carleman] and we are going to expose results from Carleman's article. In general, consider some other power series

(1) $$\phi(z) = \sum b_\nu \cdot z^\nu$$

which represents an analytic function D where $|\phi(z)| < 1$ hold when $|z| < 1$. Then there exists the composed analytic function

(*) $$f(\phi(z)) = \sum_{k=0}^{\infty} c_k \cdot z^k$$

We seek conditions on ϕ in order that the convergence of $\{a_k\}$ entails that the series

(**) $$\sum c_k \quad \text{also converges}$$

8. Convergence under substitution.

Introduction. Let $\{a_k\}$ be a sequence of complex numbers where $\sum a_k$ is convergent. This gives an analytic function $f(z)$ defined in the open disc by

(1) $$f(z) = \sum a_n \cdot z^n$$

Let us also consider another analytic function $\phi(z)$ in the open unit disc and assume that $|\phi(z)| < 1$ for every $z \in D$. Then there exists the composed analytic function in D:

(i) $$f \circ \phi(z) = \sum_{n=0}^{\infty} a_n \cdot \phi(z)^n$$

Here $f \circ \phi$ has a series expansion

(ii) $$f \circ \phi(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$$

Under the hypothesis that the additive series $\sum a_n$ converges we seek conditions on ϕ in order that $\sum c_n$ converges. To analyze this problem we consider the series of ϕ:

(iii) $$\phi(z) = \sum_{n=0}^{\infty} b_n \cdot z^n$$

With these notations our first result goes as follows:

Theorem 8.1 *Assume that each b_n is real and non-negative and that $\sum b_n = 1$. Then the additive series $\sum c_n$ converges and the sum is equal to $\sum a_n$.*

When $\{b_n\}$ no longer cinsists of non-negative real numbers we shall prove the convergence of $\sum c_n$ via condition upon ϕ. To begin with we assume that ϕ extends to a continuous function on the closed unit disc and that $\phi(1) = 1$ while $|\phi(e^{i\theta})| < 1$ when $\theta \neq 0$ and there exists some $\delta > 0$ and a constant c such that such that

$$|1 - e^{i\theta})| \leq C \cdot |\theta|^{\delta}$$

This implies that the integrals below exist for each pair of non-negative integers n and p:

(*) $$J(n,p) = \int_{-\ell}^{\ell} \frac{\phi(e^{i\theta})^n \cdot (1 - \phi(e^{i\theta})}{e^{ip\theta} \cdot (1 - e^{i\theta})} \, d\theta$$

With these notations one has

8.2. Theorem. *Let ϕ satisfy the conditions above. and suppose in addition that there exists a constant K such that*

(*) $$\sum_{n=0}^{\infty} |J(n,p)| \leq C \quad \text{for all} \quad p \geq 0$$

Then $\sum c_n$ converges and the sum is equal to $\sum a_n$.

Proof. Since $\{b_\nu\}$ are real and non-negative the Taylor series for ϕ^k also has non-negative real coefficients for every $k \geq 2$.

Put

$$\phi^k(z) = \sum B_{k\nu} \cdot z^{\nu}$$

and for each pair of integers k, p we set

$$\Omega_{k,p} = \sum_{\nu=0}^{\nu=p} B_{k\nu}$$

The assumption on $\{b_\nu\}$ entails that $\lim_{x \to 1} \phi(x)^k = 1$ for each $k \geq 1$ and it is clear that the following hold:

(i) $$\lim_{N \to \infty} \Omega_{N,p} = 0 \quad \text{for every} \quad p$$

(ii) $$k \mapsto \Omega_{k,p} \quad \text{decreases for every} \quad p$$

(iii) $$\sum_{\nu=0}^{\infty} B_{k\nu} = 1 \quad \text{hold for every} \quad k$$

The Taylor series of the composed analytic function $f(\phi(z))$ is given by

$$\sum a_k \cdot \phi^k(z) = \sum_{\nu=0}^{\infty} [\sum_{k=0}^{\infty} a_k \cdot B_{k\nu}] \cdot z^\nu$$

Next, for each positive integer n^* we set

(1) $$\sigma_p[n^*] = \sum_{\nu=0}^{\nu=p} [\sum_{k=0}^{k=n^*} a_k \cdot B_{k,\nu}]$$

(2) $$\sigma_p(n^*) = \sum_{\nu=0}^{\nu=p} [\sum_{k=n^*+1}^{\infty} a_k \cdot B_{k,\nu}] = \sum_{k=n^*+1}^{\infty} a_k \cdot \Omega_{k,p}$$

Notice that

$$\sigma_p[n^*] + \sigma_p(n^*) = \sum_{k=0}^{k=p} c_k \quad \text{hold for each} \quad p$$

Our aim is to show that the last partial sums have a limit. To obtain this we study the σ-terms separately. Introduce the partial sums

$$s_n = \sum_{k=0}^{k=n} a_k$$

By assumption there exists a limit $s_n \to S$ which entails that the sequence $\{s_k\}$ is bounded and so is the sequence $\{a_k = s_k - s_{k-1}\}$. By (i) above it follows that the last term in (2) ends to zero when n^* increases. So if $\epsilon > 0$ we find $n*$ such that

(3) $$n \geq n^* \implies |\sigma_p(n)| \leq \epsilon$$

A study of $\sigma_p[n^*]$. Keeping n^* and ϵ fixed we apply (iii) for each $0 \leq k \leq n^*$ and find an integer p^* such that

$$1 - \sum_{\nu=0}^{\nu=p} B_{k,\nu} \leq \frac{\epsilon}{n^*+1} \quad \text{for all pairs} \quad p \geq p^* : 0 \leq k \leq n^*$$

The triangle inequality gives

(6) $$|\sigma_p(n^*) - s_{n^*}| \leq \frac{\epsilon}{n^*+1} \cdot \sum_{k=0}^{k=n^*} |a_k| \quad \text{for all} \quad p \geq p^*$$

Since $\sum a_k$ converges the sequence $\{a_k\}$ is bounded, i.e. we have a constant M such that $|a_k| \leq M$ for all k. Hence (4) and (6) give

(6) $$|\sigma_p(n^*) - S| \leq \epsilon + \epsilon \cdot M \quad : \quad p \geq p^*$$

Together with (5) this entails that

$$n \geq n^* \implies |\sum_{k=0}^{n^*} |c_k - s| \leq 2\epsilon + M \cdot \epsilon$$

Since we can chose ϵ arbitrary small we conclude that $\sum c_k$ converges and the limit is equal to S which finishes the proof of Theorem 8.1.

Now we relax the condition that $\{b_\nu\}$ are real and nonnegative but impose extra conditions on ϕ. First we assume that $\phi(z)$ extends to a continuous function on the closed disc, i.e. ϕ belongs to the disc-algebra. Moreover, $\phi(1) = 1$ while $|\phi(z)| < 1$ for all $z \in \bar{D} \setminus \{1\}$ which means that $z = 1$ is a peak point for ϕ. Consider also the function $\theta \mapsto \phi(e^{i\theta})$ where θ is close to zero. The final condition on ϕ is that there exists some positive real number β and a constant C such that

$$|\phi(e^{i\theta}) - 1 - i\beta| \leq C \cdot \theta^2 \tag{1}$$

holds in some interval $-\ell \leq \theta \geq \ell$. This implies that for every integer $n \geq 2$ we get another constant C_n so that

$$|\phi^n(e^{i\theta}) - 1 - in\beta| \leq C_n \cdot \theta^2 \tag{2}$$

Hence the following integrals exist for all pairs of integers $p \geq 0$ and $n \geq 1$:

$$J(n.p) = \int_{-\ell}^{\ell} \frac{\phi(e^{i\theta})^n \cdot (1 - \phi(e^{i\theta})}{e^{ip\theta} \cdot (1 - e^{i\theta})} \cdot d\theta \tag{3}$$

With these notations one has

8.2. Theorem. *Let ϕ satisfy the conditions above. Then, if there exists a constant C such that*

$$\sum_{k=0}^{\infty} |J(k,p)| \leq C \quad \text{for all} \quad p \geq 0 \tag{*}$$

it follows that the series (**) from the introduction converges and the sum is equal to $\sum a_k$.

Proof With similar notations as in the previous proof we introduce the Ω-numbers by:

$$\Omega_{k,p} = \sum_{\nu=0}^{\nu=k} B_{k\nu}$$

Repeating the proof of Theorem 8.1 the reader may verify that the series $\sum c_k$ converges and has the limit S if the following two conditions hold:

$$\lim_{N \to \infty} \Omega_{N,p} = 0 \quad \text{holds for every} \quad p \tag{i}$$

$$\sum_{k=0}^{\infty} \left| \Omega_{k+1,p} - \Omega_{k,p} \right| \leq C \quad \text{for a constant} \quad C \tag{ii}$$

where C is is independent of p. Here (i) is clear since $\{g_N(z) = \phi^N(z)\}$ converge uniformly to zero in compact subsets of the unit disc and therefore their Taylor coefficients tend to zero with N. To get (ii) we use residue calculus which gives:

$$\Omega_{k+1,p} - \Omega_{k,p} = \frac{1}{2\pi i} \int_{|z|=1} \frac{\phi^k(z)}{z^{p+1}} \cdot \frac{1 - \phi(z)}{1 - z} \cdot dz \tag{iii}$$

Let ℓ be a small positive number and T_ℓ denotes the portion of the unit circle where $\ell \leq \theta \leq 2\pi - \ell$. Since 1 is a peak -point for ϕ there exists some $\mu < 1$ such that

$$\max_{z \in T_\ell} |\phi(z)| \leq \mu$$

This gives

$$\frac{1}{2\pi} \cdot \Big| \int_{z \in T_\ell} \frac{\phi^k(z)}{z^{p+1}} \cdot \frac{1 - \phi(z)}{1 - z} \cdot dz \Big| \leq \mu^k \cdot \frac{2}{|e^{i\ell} - 1} \Big| \tag{iv}$$

Since the geometric series $\sum \mu^k$ converges it follows from (iii) and the construction of the J_ℓ-functions in Theorem 8.2 that (ii) above holds precisely when

$$\sum_{k=0}^{\infty} |J_\ell(k,p]| \leq C$$

hold for a constant which is independent of p which finishes the the proof of Theorem 8.2.

8.3. An oscillatory integrals. The condition (*) Theorem 8.2 is implicit. A sufficient condition in order that the J-integrals satisfy (*) can be expressed by local conditions on the ϕ-function close to $z = 1$. To begin with the condition (1) above Theorem 8.2 entails that

(i) $$\phi(e^{i\theta}) = e^{i\beta\theta + \rho(\theta)}$$

holds in a neighborhood of $\theta = 0$ where the ρ-function behaves like big ordo of θ^2 when $\theta \to 0$. The next result gives the requested convergence of the composed series expressed by an additional condition on the ρ-function in (i) above.

8.4. Theorem. *Assume that $\rho(\theta)$ is a C^2-function on some interval $-\ell < \theta < \ell$ and that the second derivative $\rho''(0)$ is real and negative. Then (*) in Theorem 8.2 holds.*

Remark. We leave the proof as a (hard) exercise to the reader. If necessary, consult Carleman's article [Car] which contains a detailed proof.

IX. The series $\sum \big[a_1 \cdots a_\nu\big]^{\frac{1}{\nu}}$

Introduction. We shall prove a result from [Carleman:xx. Note V page 112-115]. Let $\{a_\nu\}$ be a sequence of positive real numbers such that $\sum a_\nu < \infty$ and e denotes Neper's constant.

9.1 Theorem. *Assume that the series $\sum a_\nu$ is convergent and let S be the sum. Then one has the strict inequality*

$$\sum_{\nu=1}^{\infty} \big[a_1 \cdots a_\nu\big]^{\frac{1}{\nu}} < e \cdot S \tag{*}$$

Remark. The result is sharp in the sense that e cannot be replaced by a smaller constant. To see this we consider a large positive integer N and take the finite series $\{a_\nu = \frac{1}{\nu} \quad : 1 \leq \nu \leq N\}$. Stirling's limit formula gives:

$$\big[a_1 \cdots a_\nu\big]^{\frac{1}{\nu}} \simeq \frac{e}{\nu} \quad : \nu >> 1$$

Since the harmonic series $\sum \frac{1}{\nu}$ is divergent we conclude that for every $\epsilon > 0$ there exists some large integer N such that $\{a_\nu = \frac{1}{\nu}\}$ gives

$$\sum_{\nu=1}^{\nu=N} \big[a_1 \cdots a_\nu\big]^{\frac{1}{\nu}} > (e-\epsilon) \cdot \sum_{\nu=1}^{\nu=N} \frac{1}{\nu}$$

There remains to prove the strict upper bound (*) when $\sum a_\nu$ is a convergent positive series. To attain this we first establish inequalities for finite series. Given a positive integer m we consider the variational problem

$$\max_{a_1,\ldots,a_m} \sum_{\nu=1}^{\nu=m} \big[a_1 \cdots a_\nu\big]^{\frac{1}{\nu}} \quad \text{when} \quad a_1 + \ldots + a_m = 1 \tag{1}$$

Let $a_1^*, \ldots, a_m^*$ give a maximum and set $a_\nu^* = e^{-x_\nu}$. The Lagrange multiplier theorem gives a number $\lambda^*(m)$ such that if

$$y_\nu = \frac{x_\nu + \ldots + x_m}{\nu}$$

then

$$\lambda^*(m) \cdot e^{-x_\nu} = \frac{1}{\nu} \cdot e^{-y_\nu} + \ldots + \frac{1}{m} \cdot e^{-y_m} \quad : \quad 1 \leq \nu \leq m \tag{2}$$

A summation over all ν gives

$$\lambda^*(m) = e^{-y_1} + \ldots + e^{-y_m} = \sum_{\nu=1}^{\nu=m} \big[a_1^* \cdots a_\nu^*\big]^{\frac{1}{\nu}}$$

Hence $\lambda^*(m)$ gives the maximum for the variational problem which is no surprise since $\lambda^*(m)$ is Lagrange's multiplier. Now we shall prove the strict inequality

$$\lambda^*(m) < e \tag{3}$$

We prove (3) by contradiction. To begin with, subtracting the successive equalities in (2) we get the following equations:

$$\lambda^*(m) \cdot [e^{-x_\nu} - e^{-x_{\nu+1}}] = \frac{1}{\nu} \cdot e^{-y_\nu} \quad : \quad 1 \leq \nu \leq m-1 \tag{4}$$

$$m \cdot \lambda^*(m) = e^{x_m - y_m} \tag{5}$$

Next, set

$$\omega_\nu = \nu(1 - \frac{a_{\nu+1}}{a_\nu}) : \quad 1 \leq \nu \leq m-1 \tag{6}$$

With these notations it is clear that (4) gives

$$\lambda^*(m) \cdot \omega_\nu = e^{x_\nu - y_\nu} \quad : \quad 1 \leq \nu \leq m-1 \tag{7}$$

It is clear that (7) gives:

$$\big(\lambda^*(m)\cdot\omega_\nu\big)^\nu = e^{\nu(x_\nu - y_\nu)} = \frac{a_1\cdots a_{\nu-1}}{a_\nu^{\nu-1}} \tag{8}$$

By an induction over ν which is left to the reader it follows the ω-sequence satisfies the recurrence equations:

$$\omega_\nu^\nu = \frac{1}{\lambda^*(m)}\cdot\big(\frac{\omega_{\nu-1}}{1-\frac{\omega_{\nu-1}}{\nu-1}}\big)^{\nu-1} \quad : \quad 1\leq\nu\leq m-1 \tag{9}$$

Notice that we also have

$$\omega_1 = \frac{1}{\lambda^*(m)} \tag{10}$$

A special construction. With λ as a parameter we define a sequence $\{\mu_\nu(\lambda)\}$ by the recursion formula:

$$(**) \qquad \mu_1(\lambda) = \frac{1}{\lambda} \quad \text{and} \quad [\mu_\nu(\lambda)]^\nu = \frac{1}{\lambda}\cdot\big[\frac{\mu_{\nu-1}(\lambda)}{1-\frac{\mu_{\nu-1}(\lambda)}{\nu-1}}\big]^{\nu-1} \quad : \quad \nu\geq 2$$

From (5) and (9) it is clear that $\lambda = \lambda^*(m)$ gives the equality

$$(***) \qquad \mu_m(\lambda^*(m)) = m$$

Now we come to the key point during the whole proof.

Lemma *If $\lambda\geq e$ then the $\mu(\lambda)$ -sequence satisfies*

$$\mu_\nu(\lambda) < \frac{\nu}{\nu+1} \quad : \quad \nu = 1,2,\ldots$$

Proof. We use an induction over ν. With $\lambda\geq e$ we have $\frac{1}{\lambda}<\frac{1}{2}$ so the case $\nu=1$ is okay. If $\nu\geq 1$ and the lemma holds for $\nu-1$, then the recursion formula (**) and the hypothesis $\lambda\geq e$ give:

$$[\mu_\nu(\lambda)]^\nu = \frac{1}{\lambda}\cdot\big[\frac{\mu_{\nu-1}(\lambda)}{1-\frac{\mu_{\nu-1}(\lambda)}{\nu-1}}\big]^{\nu-1} < \frac{1}{e}\cdot\big[\frac{\frac{\nu-1}{\nu}}{1-\frac{\nu-1}{\nu(\nu-1)}}\big]^{\nu-1}$$

Notice that the last factor is 1 and hence:

$$[\mu_\nu(\lambda)]^\nu < e < \big(1+\frac{1}{\nu}\big)^{-\nu}$$

where the last inequality follows from the wellknown limit of Neper's constant. Taking the ν:th root we get $\mu_\nu(\lambda) < \frac{\nu}{\nu+1}$ which finishes the induction.

Conclusion. If $\lambda^*(m)\geq e$ then the lemma above and the equality (***) would entail that

$$m = \mu(\lambda^*(m)) < \frac{m}{m+1}$$

This is impossible when m is a positive integer and hence we must have proved the strict inequality $\lambda^*(m) < e$.

The strict inequality for an infinite series. It remains to prove that the strict inequality holds for a convergent series with an infinite number of terms. So assume that we have an equality

$$\sum_{\nu=1}^{\infty}\big[a_1\cdots a_\nu\big]^{\frac{1}{\nu}} = e\cdot\sum_{\nu=1}^{\infty} a_\nu \tag{i}$$

Put as as above

$$\omega_n = n\big(1-\frac{a_{n+1}}{a_n}\big) \tag{ii}$$

Since we already know that the left hand side is at least equal to the right hand side one can apply Lagrange multipliers and we leave it to the reader to verify that the equality (i) gives the recursion formulas

$$\omega_n^n = \frac{1}{e}\cdot\big[\frac{\omega_{n-1}}{1-\frac{\omega_{n-1}}{n-1}}\big]^{n-1} \tag{iii}$$

Repeating the proof of the Lemma above it follows that

$$\text{(iv)} \qquad \omega_n < \frac{n}{n+1} \implies \frac{a_{n+1}}{a_n} > \frac{n}{n+1}$$

where (ii) gives the implication. So with $N \geq 2$ one has:

$$\frac{a_{N+1}}{a_1} > \frac{1 \cdots N}{1 \cdots N(N+1)} = \frac{1}{N+1}$$

Now $a_1 > 0$ and since the harmonic series $\sum \frac{1}{N}$ is divergent it would follows that $\sum a_n$ is divergent. This contradiction shows that a strict inequality must hold in Theorem 9.1.

10. Thorin's convexity theorem.

Introduction. In the article [Thorin] a convexity theorem was established which goes as follows: Let $N \geq 2$ be a positive integer and $\mathcal{A} = \{A_{\nu k}\}$ a complex $N \times N$-matrix. To each pair of real numbers a, b in the square $\square = \{0 < a, b < 1\}$ we set

$$M(a,b) = \max_{x,y} \left|\sum\sum A_{\nu k} \cdot x_k \cdot y_\nu\right| \quad : \quad \sum |x_\nu|^{1/a} = \sum |y_k|^{1/b} = 1$$

10.1 Theorem *The function $(a,b) \mapsto \log M(a,b)$ is convex in $\square$.*

The proof relies upon Hadamard's inequality for maximum norms of bounded analytic functions in strip domains. More precisely, let $f(w)$ be an entire function which is bounded in the infinite strip domain

$$\Omega = \{\sigma + is \quad : 0 \leq \sigma \leq 1 : \; -\infty < s < \infty\}$$

Set

$$M_f(\sigma) = \max_s f(\sigma + is)| \quad : \quad 0 \leq \sigma \leq 1$$

Then the following is proved in § XX:

$$\text{(*)} \qquad M_f(\sigma) \leq M_f(0)^{1-\sigma} \cdot M_f(1)^\sigma$$

Proof of Theorem 10.1. With $0 < a, b < 1$ fixed we consider N-tuples $x_\bullet$ and $y_\bullet$ in $\mathbf{C}^N$ and write

$$x_\nu = c_\nu^a \cdot e^{i\theta_\nu} \quad \text{and} \quad y_k = d_k e^{i\phi_k}$$

where the c-and the d-numbers are real and positive whenever they are $\neq 0$. It is is clear that

$$\text{(1)} \qquad M(a,b) = \max_{c,d,\theta,\phi} \left|\sum\sum A_{\nu k} \cdot c_\nu^a \cdot d_k^b \cdot e^{i\theta_\nu} e^{i\phi_k}\right|$$

where the maximum is taken over N-tuples $\{c_\bullet\}$ and $\{d_\bullet\}$ of non-negative real numbers such that

$$\text{(2)} \qquad \sum c_\nu = \sum d_k = 1$$

and $\{\theta_\nu\}$ and $\{\phi_k\}$ are arbitrary N-tuples from the periodic interval $[0, 2\pi]$. Consider a pair (a_1, b_1) and (a_2, b_2) in $\square$ and let $(\bar{a}, \bar{b})$ be the middle point. Then we find $c^*, d^*, \theta^*, \phi^*$ so that

$$\text{(3)} \qquad M(\bar{a}, \bar{b}) = \left|\sum\sum A_{\nu k} \cdot (c_\nu^*)^a \cdot (d_k^*)^b \cdot e^{i\theta_\nu^* + i\phi_k^*}\right|$$

Let $w = \sigma + is$ be a complex variable and define the analytic function f by

$$\text{(4)} \qquad f(w) = \sum\sum A_{\nu k} \cdot (c_\nu^*)^{a_1 + w(a_2 - a_1)} \cdot (d_k^*)^{b_1 + w(b_2 - b_1)} \cdot e^{i\theta_\nu^* + i\phi_k^*}$$

It is clear that $f(w)$ is an entire analytic function and $|f(1/2)| = M(\bar{a}, \bar{b})$. Next, with $w = is$ purely imaginary we have

$$\text{(5)} \qquad f(is) = \sum\sum A_{\nu k} \cdot (c_\nu^*)^{a_1} \cdot (d_k^*)^{b_1} \cdot e^{is(a_2 - a_1)\log c_\nu^* + is(b_2 - b_1)\log d_k^*} \cdot e^{i\theta_\nu^* + i\phi_k^*}$$

For each pair ν, k the exponential product

$$e^{is(a_2 - a_1)\log c_\nu^* + is(b_2 - b_1)\log d_k^*} \cdot e^{i\theta_\nu^* + i\phi_k^*} = e^{i(\theta_\nu(s) + \phi_k(s))}$$

for some pair $\theta_\nu(s), \phi_k(s)$. From (1) we see that

$$\max_s |f(is)| \leq M(a_1, b_1) \tag{6}$$

In the same way the reader can verify that

$$\max_s |f(1+is)| \leq M(a_2, b_2) \tag{7}$$

Now Hadamard's inequality (*) entails that

$$\log M(\bar{a}, \bar{b}) \leq \frac{1}{2} \cdot [\log M(a_1, b_1) + \log M(a_2, b_2)$$

This proves the required convexity.

11. Cesaro and Hölder limits

Introduction. In 1880 Cesaro introduced a certain summation procedure which which is a substitute for divergent series and leads to the notion of Cesaro summability to be defined below. Another summability was introduced by Hölder and later Knopp and Schnee proved that the conditions for Cesaro-respectively Hölder are equivalent. In Theorem 11.8 we present the elegant proof due to Schur taken from [Landau; Chapter 2]. For a given sequence of complex numbers $a_0, a_1, a_2, \ldots$ we put:

$$S_n = a_0 + \ldots + a_n$$

If $k \geq 0$ we define inductively

$$S_n^{(k+1)} = S_0^{(k)} + \ldots + S_n^{(k)} \quad \text{where} \quad S_n^{0)} = S_n$$

11.1 Definition. *For a given integer $k \geq 0$ we say that the sequence $\{a_n\}$ is Cesaro summable of order k if there exists a limit*

(*) $$s_*(k) = \lim_{n \to \infty} \frac{k!}{n^k} \cdot S_n^{(k)}$$

11.2 Exercise. Assume that $\{a_n\}$ is Cesaro summable of some order k. Show that

$$a_n = O(n^k)$$

11.3 The power series $f(x) = \sum a_n x^n$. Assume that $\{a_n\}$ is Cesaro summable of some order k. Exercise 11.2 shows that the power series $f(x)$ has a radius of convergence which is at least one and for every integer $k \geq 0$ the reader can verify the equality

(*) $$f(x) = \sum a_n x^n = (1-x)^{k+1} \cdot \sum S_n^{(k)} \cdot x^n$$

11.4 Exercise. Deduce from the above that if $\{a_n\}$ is Cesaro summable of some order k_* with limit value $s_*(k_*)$ then one has the limit formula:

$$s_*(k_*) = \lim_{x \to 1} f(x)$$

Now we prove that Cesaro summability of some order implies the summability for every higher order.

11.5 Proposition. *If (*) holds for some k_* then the Cesaro limit exists for every $k \geq k_*$ and one has the equality $s_*(k) = s_*(k_*)$.*

Proof. Cesaro summability of some order k with a limit $s_*(k)$ means that

(i) $$S_n^{(k)} = \frac{n^k}{k!} \cdot s_*(k) + o(n^k)$$

where the last term is small ordo. If (i) holds we get

$$S_n^{(k+1)} = \frac{s_*(k)}{k!} \sum_{\nu=0}^{n} n^\nu + o\Big(\sum_{\nu=0}^{n} n^\nu\Big) = \frac{s_*(k)}{k!} \cdot \Big[\frac{n^{k+1}}{k+1} - 1\Big] + o(n^{k+1})$$

From this the reader discovers the requested induction step and Proposition 11.5 follows.

11.6 Hölder's summation. To each sequence of complex numbers $a_0, a_1, a_2, \ldots$ we put

$$H_n^{(0)} = a_0 + \ldots + a_n$$

and if $k \geq 0$ we define inductively

$$H_n^{(k+1)} = \frac{H_0^{(k)} + \ldots + H_n^{(k)}}{n+1}$$

11.7 Definition. *The sequence $\{a_n\}$ is Hölder summable of order k if there exists a limit*

(**) $$\lim_{n \to \infty} H_n^{(k)}$$

11.8 Theorem *A sequence $\{a_n\}$ is Cesaro summable of of some order k if and only if it is Hölder summable of the same order and there respectively limits are the same.*

The proof of Theorem 11.8 requires several steps. First we introduce arithmetic mean value sequences attached to every sequence $\{x_0, x_1, \ldots\}$:

$$M(\{x_\nu\})[n] = \frac{x_0 + \ldots + x_n}{n+1}$$

Next, to each $k \geq 1$ we construct the sequence $T_k(\{x_\nu\})$ by

$$T_k(\{x_\nu\})[n] = \frac{k-1}{k} \cdot M(\{x_\nu\})[n] + \frac{x_n}{k}$$

So above M and $\{T_k\}$ are linear operators which send a complex sequence to another complex sequence. The reader may verify that these operators commute, i.e.

$$T_k \circ M = M \circ T_k$$

hold for every k and similarly the T-operators commute. For a given k we can also regard the passage to the Cesaro sequence $\{S_n^{[k]}\}$ as a linear operator which we denote by $\mathcal{C}^{(k)}$. Similarly we get the Hölder operators $\mathcal{H}^{(k)}$ for every $k \geq 1$.

11.9 Proposition. *The following identities hold*

(i) $$T_k \circ \mathcal{C}^{(k-1)} = M \circ \mathcal{C}^{(k)} \quad : \quad k \geq 1$$

(ii) $$\mathcal{H}^{(k)} = T_2 \circ \ldots \circ T_k \circ \mathcal{C}^{(k)} \quad : \quad k \geq 2$$

11.10 Exercise. Prove (i) and (ii) above.

As a last preparation towards the proof of Theorem 11.8 we need certain limit formulas which show that the T-operators have robust properties. First we have:

11.11 Lemma *Let $\{x_1, x_2, \ldots\}$ be a sequence of complex numbers and q a positive integer such that*

$$\lim_{n\to\infty} q \cdot \frac{x_1 + \ldots + x_n}{n} + x_n = 0$$

Then it follows that

$$\lim_{n\to\infty} x_n = 0$$

Proof. Set $y_n = q(x_1 + \ldots + x_n) + nx_n$. By an induction over n one verifies that

(1) $$\sum_{\nu=1}^{\nu=n} (\nu+1)\cdots(\nu+q-1) \cdot y_\nu = (n+1)\cdots(n+q) \cdot \sum_{\nu=1}^{\nu=n} x_\nu$$

hold for every $n \geq 1$. By the hypothesis $y_n = o(n)$ where $o(n)$ is small ordo of n. It follows that the left hand side in (1) is $o(n^{q+1})$ and since the product $(n+1)\cdots(n+q) \simeq n^q$ we conclude that

(2) $$\sum_{\nu=1}^{\nu=n} x_\nu = o(n)$$

Finally, we have

$$nx_n = y_n - q \cdot \sum_{\nu=1}^{\nu=n} x_\nu$$

and by (2) and the hypothesis the right hand side is $o(n)$ which after division with n gives $x_n = o(1)$ as required.

11.12 Proposition. *Let $\{x_\nu\}$ be a sequence and $k \geq 1$ an integer such that there exists*

$$\lim_{n\to\infty} T_k(\{x_\nu\}][n] = s$$

Then it follows that $\{x_n\}$ converges to s.

11.13 Exercise. Deduce Proposition 11.12 from Lemma 11.11.

11.14 Proof of Theorem 11.8.

The easy case $k = 1$ is left to the reader and we proceed to prove the theorem when $k \geq 2$. Assume first that $\{a_n\}$ is Cesaro summable of some order $k \geq 2$ with a limit s. Exercise 11.12 implies that $T_k \circ \mathcal{C}^{(k)}$ sends $\{a_n\}$ to a convergent sequence with limit s. If $k \geq 3$ we apply the exercise to T_{k-1} and continue until the composed operator

$$T_2 \circ \cdots \circ T_k \circ \mathcal{C}^{(k)}$$

sends the a-sequence to a convergent sequence with limit s. By (ii) in Proposition 11.8 this entails that $\{a_k\}$ is Hölder summable of order k with limit k. Conversely, assume that $\{a_n\}$ is Hölder summable of some order $k \geq 2$. The equality (ii) from Proposition 11.9 gives

$$\mathcal{H}^{(2)} = T_2 \circ \mathcal{C}^{(2)}$$

Hence Proposition 11.13 applied to T_2 shows that Hölder summability of order 2 entails Cesaro summability of the same order. Next, if $k \geq 3$ we again use (ii) in 11.9 and conclude that the sequence

$$T_3 \circ \ldots T_k \circ \mathcal{C}^{(k)}(\{a_n\})$$

is convergent. By repeated application of (ii) in 11.18 applied to $T_3, \ldots, T_k$ we conclude that the a-sequence is Cesaro summable of order k and has the same limit as the Hölder sum.

12. Power series and arithmetic means.

Consider a power series

$$f(x) = \sum a_n \cdot x^n$$

which converges when $|x| < 1$ and assume also that

$$(*) \qquad \lim_{x \to 1} \sum a_n \cdot x^n = 0$$

For each $k \geq 1$ we get the sequence $\{S_n^{(k)}\}$ from the previous section and we prove the following:

12.1 Theorem. *Assume (*) and that there exists some integer $k \geq 1$ such that*

$$\lim_{n \to \infty} S_n^{(k)} = 0$$

Then the series $\sum a_n$ converges.

Example. Consider the case $r = 1$ where

$$S_n^{(1)} = \frac{na_0 + (n-1)a_1 + \ldots + a_n}{n}$$

The sole assumption that $S_n^{(1)} \to 0$ does not imply $\sum a_n$ converges. But in addition (*) is assumed in Theorem 12.1 which will give the convergence. The proof of Theorem 12.1 is based upon the following convergence criterion where (*) above is tacitly assumed.

12.2 Proposition. *The series $\sum a_n$ converges if there to every $\epsilon > 0$ exists a pair (p_0, n_0) such that*

$$p \geq p_0 \implies J(n_0, p) = \Big| \int_0^1 \frac{\sin 2p\pi(x-1)}{x-1} \cdot \sum_{n=n_0}^{\infty} a_n x^n \cdot dx \Big| < \epsilon$$

Exercise. Prove this classic result which already was wellknown to Abel.

Proof of theorem 12.1. To profit upon Proposition 12.2 we need the two inequalities below which are valid for all pairs of positive integers p and n:

$$\text{(i)} \qquad \Big| \int_0^1 \sin(2p\pi x) \cdot x^k (1-x)^n \cdot dx \Big| \leq 2\pi (k+2)\,! \cdot \frac{p}{n^{k+2}}$$

$$\text{(ii)} \qquad \Big| \int_0^1 \sin 2p\pi x \cdot x^k (1-x)^n \cdot dx \Big| \leq \frac{C(k)}{p \cdot n^k}$$

where the constant $C(k)$ in (ii) as indicated only depends upon k. The verification of (i-ii) is left to the reader. Next, recall from (*) in § 11.4 that:

$$\text{(iii)} \qquad f(x) = \frac{(1-x)^{k+1}}{(k+1)!} \cdot \sum S_n^{(k)} n^k \cdot x^n$$

Let $\epsilon > 0$ and choose n_0 such that

$$\text{(iv)} \qquad n \geq n_0 \implies |S_n^{(k)}| < \epsilon$$

which is possible from the assumption in Theorem 12.1 Notice that (iii) gives the equality

$$\text{(iii)} \qquad \sum_{n=n_0}^{\infty} a_n x^n = \frac{(1-x)^{k+1}}{(k+1)!} \cdot \sum_{n=n_0}^{\infty} S_n^{(k)} n^k \cdot x^n$$

Hence (iv) and the triangle inequality shows that with n_0 kept fixed, the absolute value of the integral in Proposition 12.2 is majorized as follows for every p:

$$J(n_0, p) \leq \epsilon \cdot \sum_{n=n_0}^{\infty} \frac{n^k}{(k+1)!} \cdot \Big| \int_0^1 \sin(2p\pi x) \cdot x^k (x-1)^n \cdot dx \Big|$$

In (iv) we have chosen n_0 and for an arbitrary $p \geq p_0 = n_0 + 1$ we decompose the sum from n_0 up to p and after we take a sum with $n \geq p+1$ which means that $J(n_0, p)$ is majorized by ϵ times the sum of the following two expressions:

$$\sum_{n=n_0}^{n=p} \frac{n^k}{(k+1)!} \cdot \Big| \int_0^1 \sin(2p\pi x) \cdot x^k (x-1)^n \cdot dx \Big| \tag{1}$$

$$\sum_{n=p+1}^{\infty} \frac{n^k}{(k+1)!} \cdot \Big| \int_0^1 \sin(2p\pi x) \cdot x^k (x-1)^n \cdot dx \Big| \tag{2}$$

Using (i) above it follows that (1) is estimated by

$$2\pi \cdot (k+2)! \cdot \frac{C(k)}{p} \cdot (p - n_0) \leq 2\pi \cdot (k+2)! \cdot C(k) = K_1$$

Next, using (ii) it follows that (2) is estimated by

$$\pi \cdot \frac{k+2}{k+1} \cdot p \cdot \sum_{n=p+1}^{\infty} n^{-2} \leq \pi \cdot \frac{k+2}{k+1} = K_2$$

So with $K = K_1 + K_2$ we have

$$J(n_0, p) \leq 2K \cdot \epsilon$$

for every $p \geq n_0 + 1$ and since $\epsilon > 0$ was arbitrary the proof of Theorem 12.2 is finished.

13. Taylor series and quasi-analytic functions.

Introduction. Let $f(x)$ an infinitely differentiable function defined on the interval $[0,1]$. At $x=0$ we can take the derivatives and set

$$C_\nu = f^{(\nu)}(0)$$

In general the sequence $\{C_\nu\}$ does not determine $f(x)$. The standard example is the C^∞-function defined for $x>0$ by $e^{-1/x}$ and zero on $x \leq 0$. Here $\{C_\nu\}$ is the null sequence and yet the function is no identically zero. But if we impose sufficiently strong growth conditions on the derivatives of f over the whole interval $(-1,1)$ then $\{C_\nu\}$ determines f. In general, let $\mathcal{A} = \{\alpha_\nu\}$ be an increasing sequence of positive real numbers and denote by $\mathcal{C}_\mathcal{A}$ the class of C^∞-functions on $[0,1]$ where the maximum norms of the derivatives satisfy

$$(*) \qquad \max_x |f^{(\nu)}(x)| \leq k^\nu \cdot \alpha_\nu^\nu \quad : \quad \nu = 0,1,\ldots$$

for some $k>0$ which may depend upon f. One says that $\mathcal{C}_\mathcal{A}$ is a quasi-analytic class if every $f \in C_\mathcal{A}$ whose Taylor series is identically zero at $x=0$ vanishes identically on $[0,1)$. In the article [Denjoy 1921), Denjoy proved that $C_\mathcal{A}$ is quasi-analytic if the series

$$(**) \qquad \sum \frac{1}{\alpha_\nu} = +\infty$$

The conclusive result which gives a necessary and sufficient condition on the sequence $\{\alpha_\nu\}$ in order that $C_\mathcal{A}$ is quasi-analytic is proved in Carleman's book [1923]. The criterion is as follows:

Theorem. *Set $A_\nu = \alpha_\nu^\nu$ for each $\nu \geq 1$. Then $C_\mathcal{A}$ is quasi-analytic if and only if*

$$\int_1^\infty \log\Big[\sum_{\nu=1}^\infty \frac{r^{2\nu}}{A_\nu^2}\Big] \cdot \frac{dr}{r^2} = +\infty$$

For the proof of this result we refer to § XX in Special Topics.

The reconstruction theorem. Since quasi-analytic functions by definition are determined by their Tayor series at a single point there remains the question how to determine $f(x)$ in a given quasi-analytic class $C_\mathcal{A}$ when the sequence of its Taylor coefficients at $x=0$ are given. Such a reconstruction was announced by Carleman in [CR-1923] and the detailed proof appears in [Carleman-book]. Carleman considered a class of variational problems to attain the reconstruction. Let $n \geq 1$ and for a given sequence of real numbers $\{C_0,\ldots,C_{n-1}\}$ we consider the class of n-times differentiable functions f on $[0,1]$ for which

$$\text{(i)} \qquad f^{(\nu)}(0) = C_\nu \quad : \quad \nu = 0,\ldots,n-1$$

Next, let $\{\gamma_0,\gamma_1,\ldots,\gamma_n\}$ be some $n+1$-tuple of positive numbers and consider the variational problem

$$\text{(ii)} \qquad \min_f J_n(f) = \sum_{\nu=0}^{\nu=n} \gamma_\nu^{-2\nu} \cdot \int_0^1 [f^{(\nu)}(x)]^2 \cdot dx$$

where the competing family consist of n-times differentiable functions on $[0,1]$ satisfying (i) above. The strict convexity of L^2-norms entail that the variational problem has a unique minimizing function f_n which depends linearly upon $C_0,\ldots,C_n$. In other words, there exists a unique doubly indexed sequence of functions $\{\phi_{p,n}\}$ defined for pairs $0 \leq p \leq n$ such that

$$f_n(x) = \sum_{\nu=0}^{\nu=n-1} C_p \cdot \phi_{p,n-1}(x)$$

where the functions $\{\phi_{0,n-1},\ldots\phi_{n-1,n-1}\}$ only depend upon $\gamma_0,\ldots,\gamma_n$.

A specific choice of the γ-sequence. Let $\mathcal{A} = \{\alpha_\nu\}$ be a Denjoy sequence, i.e. (**) above diverges. Set $\gamma_0 = 1$ and

$$\gamma_\nu = \frac{1}{\alpha_\nu} \cdot \sum_{p=1}^{p=\nu} \alpha_p \quad : \quad \nu \geq 1$$

Given some $F(x) \in C^\infty[0,1]$ we get the sequence $\{C_\nu = F^{(\nu)}(0)\}$ and to each $n \geq 1$ we consider the variational problem above using the n-tuple $\gamma_0, \ldots, \gamma_{n-1}$ which yields the extremal function $f_n(x)$. With these notations Carelan proved the following:

13.1 Theorem. *If $F(x)$ belongs to the class $C_{\mathcal{A}}$ it follows that*

$$\lim_{n\to\infty} f_n(x) = F(x)$$

where the convergence holds uniformly on interval $[0,a]$ for every $a < 1$.

13.2 A series expansion. Using Theorem 13.1 Carleman also proved that when the series (**) diverges, then there exists a doubly indexed sequence $\{a_{\nu,n}\}$ defined for pairs $0 \leq \nu \leq n$ which only depends on the sequence $\{\alpha_\nu\}$ such that if $F(x)$ belongs to $\mathcal{C}_{\mathcal{A}}$ then it is given by a limit of series:

$$F(x) = \lim_{n\to\infty} \sum_{\nu=0}^{\nu=n} a_{\nu,n} \cdot \frac{F^{(\nu)}(0)}{\nu\,!} \cdot x^\nu \quad : \quad 0 \leq x < 1$$

Remark. Above we exposed the reconstruction for quasi-analytic classes of the Denjoy type. For a general quasi-analytic class a similar result is proved in [Carleman]. Here the proof and the result is of a more technical nature so we refrain to present the details. Concerning the doubly indexed a-sequence it is found in a rather implicit manner via solutions to the variational problems and an extra complication is that these a-numbers depend upon the given α-sequence. it appears that several open problems remain concerning effective formulas and the reader may also consult [Carleman: page xxx] for some open questions related to the reconstruction above.

.

13.4 Quasi-analytic boundary values. Another problem is concerned with boundary values of analytic functions where the set of non-zero Taylor-coefficients is sparse. In general, consider a power series $\sum a_n z^n$ whose radius of convergence equal to one. Assume that there exists an interval ℓ on the unit circle such that the analytic function $f(z)$ defined by the series extends to a continuous function in the closed sector where $\arg(z) \in \ell$. So on ℓ we get a continuous boundary value function $f^*(\theta)$ and suppose that f^* belongs to some quasi-analytic class on this interval. Let f be given by the series

$$f = \sum a_n \cdot z^n$$

Suppose that gaps occur and write the sequence of non-zero coefficients as $\{a_{n_1}, a_{n_2} \ldots\}$ where $k \mapsto n_k$ is a strictly increasing sequence. With these notations the following result is due to Hadamard:

13.5 Theorem. *Let $f(z)$ be analytic in the open unit disc and assume it has a continuous extension to some open interval on the unit circle where the boundary function $f^*(\theta)$ is real-analytic. Then there exists an integer M such that*

$$n_{k+1} - n_k \leq M$$

for all k. In other words, the sequence of non-zero coefficients cannot be too sparse.

Hadamard's result was extended to the quasi-analytic case in [Carleman] where it is proved that if f^* belongs to some quasi-analytic class determined by a sequence $\{\alpha_\nu\}$ then the gaps cannot be too sparse, i.e. after a rather involved analysis one finds that f must be identically zero if the integer function $k \mapsto n_k$ increases too fast. The rate of increase depends upon $\{\alpha_\nu\}$ and it appears that no precise descriptions of the growth of $k \mapsto n_k$ which would ensure unicity is known for a general quasi-analytic class, i.e. even in the situation considered by Denjoy. So there remains many interesting open questions concerned with quasi-analyticity.

I:C Complex vector spaces

Contents

0. Introduction

0.A The Sylvester-Franke theorem

0.B Hankel determinants

0.C The Gram-Fredhom formula

0.D Resolvents of integral operators

0.D.1 Hilbert derterminants

0.D.2 Some results by Carleman

1. Wedderburn's Theorem

2. Resolvents

3. Jordan's normal form

4. Hermitian and normal operators

5. Fundamental solutions to ODE-equations

6. Carleman's inequality for resolvents

7. Hadamad's radius formula

8. On Positive definite quadratic forms

9. The Davies-Smith inequality

10. An application to integral equations

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 and real $n \times n$-matrices and its counterpart for complex Hermitian matrices where eigenvalues are found by regarding maxima and minima of associated quadratic forms. Far-reaching studies of quadratic forms were performed by Weierstrass whose collected work contains a wealth of results related to the spectral theorem for hermitian matrices and their interplay with quadratic forms. One should also mention further investigations by Frobenius about quadratic forms and related topics. 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. Diring 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.

A. Matrices and determinants.

Let A be a matrix whose elements $\{a_{pq}\}$ are complex numbers. The 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:

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

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

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

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

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

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

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

Then (1) and (**) give the inequality

$$\text{(***)} \qquad \text{Norm}(A) \leq ||A||$$

Example. Consider a matrix A whose elements are non-negative real numbers. Then it is clear that(*) is maximixed when the z-vector is real with non-negative components. Thus,

$$\text{(**)} \qquad \text{Norm}(A) = \max_{x_1,\dots x_n} \sqrt{\sum_{p=1}^{p=n} \big(\sum_{q=1}^{q=n} a_{pq} x_q\big)^2}$$

taken over real n-vectors for which $\sum x_p^2 = 1$ and every $x_p \geq 0$. The A-norm is found via Lagrange's multiplier i.e. one employs Lagrange's criterion for extremals of quadratic forms. The result is that (**) is maximized by a real non-negative n-vector x which satsfies a linear system of equations

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

Introducing the double indexed numbers

$$\text{(2)} \qquad \beta_{jq} = \sum_{p=1}^{p=n} a_{pj} a_{pq}$$

Lagrange's equations corresponds to the system

$$\lambda \cdot x_j^* = \sum_{q=1}^{q=n} \beta_{jq} \cdot x_q^* \tag{3}$$

Notice that the β-mstrix is symmetric, i.e. $\beta_{jq} = \beta_{qj}$ hold for ech pair. So (3) 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 the largest eigenvalue was studied by Perron, with further extensions As a specific example we consider an $n \times n$-matrix of the form

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

where s is real and positive. 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 rather involved. With $n = 2$ we find the β-matrix

$$B = \begin{pmatrix} 1 & s \\ s & 1 + s^2 \end{pmatrix}$$

and here one seeks the largest root of its characteristic polynomial to find the requested norm above. For a general $n \geq 2$ and $s = 2$ is of special interest a classic result which goes back to Hankel and Frobenius is that

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

Exercise. Prove (4). If necessary, consult the literature.

The next sections discuss determinants. The reader may skip this for a while and turn to § 1 where we treat basic results about linear operators on finite dimensional vector spaces.

0.A 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

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

Exercise. Prove (1) or consult a text-book which apart from "soft abstract notions" does not ignore to discuss determinants. Personally I recommend Gerhard Kovalevski's text-book *Determinantenheorie* from 1909 where many results about determinants are proved in an elegant and detailed fashion.

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:

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

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

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 $1 \leq m \leq n-1$.

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

0.B 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 obtain the $(p+1) \times (p+1)$-matrix:

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

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

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

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.

Remark. The sequence $\{c_n\}$ gives the formal power series

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

If $n \geq 1$ we set

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

From this 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 $(a_0, \ldots, a_p)$ 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 that

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

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

B.1 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*

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

Moreover, one has the equality $p = r^*$.

B.3 A 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 expressed by a sum of simple fractions which means that

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

where $\alpha_1, \ldots, \alpha_p$ are distinct and every $d_k \neq 0$. 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 finite rank. 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].

DELETE THIS

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

$$R(z) = \frac{q(z)}{z^p - [c_1 z^{p-1} + \ldots + c_{p-1}z + c_p]}$$

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

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

Consider the $p \times p$-matrix

$$A = \begin{pmatrix} 0 & 0 & & \ldots & 0 & c_p \\ 1 & 0 & 0 & \ldots & 0 & c_{p-1} \\ 0 & 1 & 0 & \ldots & \ldots & c_{p-2} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & c_1 \end{pmatrix}$$

Prove the following for every $n \geq 1$:

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

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

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

Remark. The equality (*) 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 construct three matrices as follows. First we get three $(N-1) \times (N-1)$-matrices

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

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

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

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

B.7 Sylvester's identity. *One has the determinant formula:*

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

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

0.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 in particular the classic book [Bocher] which contains a detailed account about Fredholm determinants and their role for solutions to integral equations.

0.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

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

An eigenvalue is a complex number $\lambda\neq 0$ for which there exists a non-zero 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 $K(x,y)\neq K(y,x)$.

0.D.1 Hilbert's determinants. Let K be a kernel function for which the integral (*) is finite. A typical case is that K is singular on the diagonal subset of $\Omega \times \Omega$. 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

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

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

Some 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. below we assume that

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

D.2.1 Theorem. *The kernel of the resolvent associated to the integral operator $\mathcal{K}$ is for each complex λ outside the spectrum given by*

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

D.2.2 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 the entire function $\mathcal{D}(\lambda)$. So when λ is outside this zero set the inverse operator $(\lambda \cdot E - \mathcal{K})^{-1}$ is the integral operator defined by

$$f \mapsto \int_{\Omega} \Gamma(\xi, \eta; \lambda) \cdot f(\eta) \, d\eta$$

where ξ and η denote variable points in Ω. Using inequalities of Fredholm-Hadamard type for determinants, it is also proved in [ibid] that:

(D.2.3) $$\int_{\Omega} \Gamma(\xi, \xi; \lambda) \cdot d\xi = -\lambda \cdot \sum_{\nu=1}^{\infty} \frac{1}{\lambda_\nu(\lambda - \lambda_\nu)}$$

Another major result in [ibid] deals with the function $\mathcal{D}(\lambda)$.

D.2.4 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. Prior to this Schur had proved a representation for $\mathcal{D}(\lambda)$ as above adding a factor $e^{b\lambda^2}$. So the novelty in Carleman's work is that $b = 0$ always hold. Apart from Schur's result that (2) above is convergent, 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:

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

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

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

D.2.5 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 D.2.5 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 btoh 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 given by the 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 $\mathbf{C}^n$. it contains special matrices $e_1, \ldots, e_n$ where the elements of e_p are zero except fro a_{pp} which is equal to one. product rukes for matrices show that

$$e_p e - q = 0 \quad |colon\, p \neq q$$

At thews same time we notice that the identity element is $e_1 + \ldots + e_n$ and that $e_p = e_p^2$. it can be expressed by saying 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 $\xi_{\nu p} \neq 0$. mutlplying with a sclar we can assume that $\xi_{\nu p} = 1$ and then

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

It follows that the principal left ideal generated by ξ is equal to L_p. Thus, every non-zero element in L_p generates L_p whuich shows 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 ξ. Then $\xi \cdot e_p$ is a non-zero element in L_p so the 2-sided ideal generated by ξ contains the minmal 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 rader mat conclude that the 2-sided ideal generated by ξ is the whole ring A. This proves that A is simple.

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

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

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

$$B \cdot A = a_{11} e_2 + a_{22} e_1 \quad : \quad A \cdot B = a_{11} \cdot e_1 + a_{22} \cdot e_2$$

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

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

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

This 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 $L^\perp$. Concclude that if p is the dimension of the vector space $L^\perp$ then the dimension of L regared as a vector space is equal to $n(n-p)$. Moreover, by $L \mapsto L^\perp$ one gets a bijective map between the family of left ideals in the matrix algebra and subspaces of $\mathbf{C}^n$.

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. Identfying L with a complex vector space of some dimension k, 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

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

We shall prove that $\mathcal{M} \simeq A$ which by gives Wedderburn's theorem. To attain this we take $a \in A$ 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 b^* and 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 the map (xi is a 2-sided ideal in A. Since A is simple this kernel is zero which proves that (i) is injective.

Proof of surjectivity. First we notice that if $x \in A$ is such that $Lx \neq 0$ then this is a left ideal and since L is minimal the reader can check that we also have $Lx \in \mathcal{L}_*$. By assumption the 2-sided ideal of L is the whole ring A. Hence there exists a finite set of A-elements $\{x_\nu\}$ such that

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

Above we can choose m minimal which gives a direct sum

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

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 (i) which contraidicts the minmal chose of m. Hence one has the direct sum in (ii). Next, for every $1 \leq k \leq m$ the map from L into Lx_k defined by

$$x \to x \cdot x_k$$

is surjetive and since L was minmal the reader can check that it is also injective. It means that the vector spaces L and Lx_k are isomorphic. Counting dimensions we conclude that

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

Since the map from A into $\mathcal{M}$ was injective and B has dimension k^2 we have $m \leq k$ and there remains onoy to prove the opposite inequality

$$(*) \qquad k \leq m$$

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

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

Put $e_\nu = \xi_\nu \cdot x - \nu$. Mupltilying to themleft by some e_k in (1) we get

$$e_k = e_k e - 1 + \ldots + e_k e_m$$

The direct sum in (xx) entials that

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

This can be exrepssed by saying that $\{e_\nu\}$ 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 mulitipkication by x on the left ideal Ae_k is left A-linear, i.e. one has a map

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

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 (xx) is a divsion 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 (xx) is zero, then

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

So (xx) is injective and hence

$$e_k A e_k = \mathbf{C}$$

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

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

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 idenpotens one has

$$\xi = e_j \xi e_k$$

Conversly we notice that for every $x \in A$, one gets a ξ-element $a_k x e_j$. Hence the vector space 8xx) above can be identfiied with the subset of A given by

$$e_j A e_k$$

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

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

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 (*) and finishes the proof of Wedderburn's theorem.

we construct the principal left ideal

$$Aa = \{x \cdot a : x \in A\}$$

and since L is minimal we have $Aa = L$. Let a be an element as above. Suppose that $x \in A$ is such that $ax \neq 0$.

left ideal L in A is minimal if there does not exist any non-zero left ideal which is strictly smaller than L. Denote by $\mathcal{L}_*$ the family of all minimal left ideals. Notice that if $0 \neq x \in L$ for some minimal ideal then we must have $A \cdot x = L$, i.e. the single element x generates L. Moreover it is clear that a left principal ideal $A \cdot x$ belongs to $\mathcal{L}_*$ if and only if the left annihilator:

$$\ell(x) = \{a \in A \quad : \; ax = 0\}$$

is a maximal left ideal. So when $A \cdot x \in \mathcal{L}_*$ and $a \in A$ is such that $xa \neq 0$, then xa also generates a minimal left ideal and the maximality of $\ell(x)$ gives the equality

$$\ell(x) = \ell(xa)$$

It follows that the left A-modules Ax and Axa are isomorphic. Next, every left ideal is in particular a complex subspace. If N is the dimension of the complex vector space A then every increasing sequence of complex subspaces has at most N strict inclusions. This shows that there exist minimal left ideals. Choose some $a_0 \neq 0$ where Aa_0 is a minimal left ideal. By left multiplication every $a \in A$ gives a $\mathbf{C}$-linear operator on Aa_0 defined by

$$a^*(xa_0) = a \cdot x \cdot a_0 \quad : \quad x \in A \tag{1}$$

If a and b is a pair of elements in A the composed $\mathbf{C}$-linear operator $b^* \circ a^*$ is given by

$$b^* \circ a^*(xa) = b^*(axa_0) = baxa_0 = (ba) \cdot x \cdot a_0 = (ba)^*(xa_0) \tag{2}$$

Hence $a \mapsto a^*$ is a homomorphism from A into the algebra $\mathcal{M} = \mathrm{Hom}_{\mathbf{C}}(Aa_0, Aa_0)$

Sublemma. The map $a \mapsto a^*$ is injective.

Proof. To say that $a^* = 0$ means that

$$axa_0 = 0 \quad \text{for all} \quad x \in A$$

Hence the *two-sided* ideal generated by a is contained in $\ell(a_0)$. So if $a \neq 0$ this two-sided ideal would be the whole ring and then $1 \cdot a_0 = a_0 = 0$ which is a contradiction.

Proof continued. Let k be the dimension of the complex vector space Aa_0 which after a chosen basis identifies $\mathcal{M}$ with the algebra of $k \times k$-matrices. Wedderburn's Theorem follows from the Sublemma if prove that the map (1) is surjective. Counting dimensions this amounts to show the equality

$$\dim_{\mathbf{C}} A = k^2 \tag{3}$$

To prove (3) we consider the the two-sided ideal generated by the family $\{Aa_0x : \; x \in A\}$. Since A is simple it gives the whole ring and we find a finite set $\{x_1, \ldots, x_m\}$ such that

$$A = Aa_0x_1 + \ldots + \ldots + Aa_0x_m \tag{4}$$

Here we can choose m to be minimal which gives a direct sum in (4), i.e. now

$$A = Aa_0x_1 + \oplus \ldots \oplus \; Aa_0x_m \tag{5}$$

By previous observations the left ideal $Aa_0x_i \simeq Aa_0$ for each i. So the direct sum decomposition (5) entails that

$$\dim_{\mathbf{C}} A = m \cdot k \tag{6}$$

Hence (3) follows if we can show the equality $m = k$. To attain this we consider the unit element 1_A in the ring A which by (5) has an expression:

$$1_A = \xi_1 + \ldots + \xi_m \quad : \; \xi_i = b_i a_0 x_i \tag{7}$$

for some m-tuple $b_1, \ldots, b_m$. Since (5) is a direct sum it is easily seen that the ξ-elements satisfy:

$$\xi_i^2 = \xi_i \quad \text{and} \quad \xi_i \cdot \xi_k \; : \; i \neq k \tag{8}$$

Thus, $\{\xi_i\}$ are mutually orthogonal idempotents. Moreover, from the previous observations we have the isomorphism of left A-modules

(9) $$A\xi_i \simeq Aa_0 \quad : \quad 1 \leq i \leq m$$

Next, since ξ is an idempotent we notice that $\xi_i A \xi_i$ is a $\mathbf{C}$-algebra which is naturally identified with the Hom-space

(10) $$\mathrm{Hom}_A(A\xi_i, A\xi_i)$$

Since $A\xi_i$ is a simple left A-module this Hom-algebra is a division ring and since $\mathbf{C}$ is algebraically closed it follows that

(11) $$\xi_i A \xi_i \simeq \mathbf{C} \quad : 1 \leq i \leq m$$

Next, for each pair i, j consider the Hom-space:

(12) $$E_{ij} = \mathrm{Hom}_A(A\xi_i, A\xi_j)$$

By the isomorphisms in (9) and the equality (11) it follows that E_{ij} are one-dimensional complex vector spaces for all pairs i, j. Moreover, the reader may verify the following equality:

(13) $$E_{ij} = \{\xi_i x \xi_j \quad : x \in A\}$$

Next, consider ξ_1. From the expression of 1_A in (7) we obtain

$$A\xi_1 = \sum_{i=1}^{i=m} \xi_i \cdot A \cdot \xi_1 = \sum E_{i1}$$

Since $\{E_{i1}\}$ are 1-dimensional we get the inequality

(*) $$k = \dim_{\mathbf{C}}(A\xi_1) \leq m$$

At this stage we are done since the injectivity in the Sublemma and (6) give

$$m \cdot k \leq k^2 \implies m \leq k$$

Together with (*) it follows that $m = k$ and the requested equality (3) follows.

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

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

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:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

The unique integer m is denoted by $\text{ord}(S, v)$. The case $m = 1$ occurs if $S(v) = 0$. If $m \geq 2$ 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 $\mathcal{C}(v)$ and called a *cyclic* subspace of V generated by 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

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

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

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

For each v_α we have a postive integer

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

The construction of a quotient space means that

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

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

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

Next, put

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

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

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

The maximal choice of m^* entails that $k_\alpha^* \leq m^*$ and since the vectors $v^*, S(v^*), \ldots S^{m^*-1}(v^*)$ are linearly independent it follows that

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

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

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

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

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

Moreover, the construction of the ξ-vectors entail that

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

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 of specific matrices 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 $1 \leq k \leq n-1$. 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.

4.0.5 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*

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

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) $$x \mapsto \langle Ax, x \rangle$$

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

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

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

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

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

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

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

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

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.

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

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:

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

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.

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 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:*

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

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. $\text{tr}(B) = 0$ 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

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

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

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

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:

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

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

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

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

and the reader may also verify that the limit of the last term above is equal to $\exp\bigl[\frac{1+||B||^2}{2}\bigr]$ 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

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

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:

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

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

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

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

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

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

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:

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

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

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

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

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

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

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

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

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

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

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

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

At the same time the diagonal elements satisfy:

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

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. von Neumann's inequality.

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*

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

To prove this we first establish a general inequality which goes back to Schur. Let $g(z)$ be an analytic function in the unit disc which extends continuously to the boundary and A is some $n \times n$-matrix whose spectrum is contained in the open unit disc. If $g(z)$ has the 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 $g(A)$. There exists also the exponential matrix

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

If $g^*(z) = \sum \overline{c_k} z^k$ then

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

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*

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

Notice that (*) holds under the sole assumption that $\sigma(A) \subset D$, i.e. A need not be a contraction.

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

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

hold for every vector y. Expanding this inequality we get

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

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

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

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

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

Since the vector x was arbitrary we conclude that

$$\text{(3.1)} \qquad \text{Norm}(g_\alpha(A)) \leq 1$$

Proof of Theorem 1. By scaling we can assume that the maximum norm $|p|_D = 1$. We construct the Blaschke product taken over the zeros of p in the open unit disc and get a factorisation

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

where the zero-free analytic function $e^{g(z)}$ has maximum norm one which gives $\mathfrak{Re}(g)(z) \leq 0$ for all $z \in D$. 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 (3.1) each of these operators have norm ≤ 1 and (*) in (2) entails that the same holds for $e^{g(A)}$. So $p(A)$ is the product of operators of norm ≤ 1 and Theorem 1 follows.

Remark. The proof given by von Neumann in [1951] is carried out for operators on Hobert spaces and we remark only that the present version for matrices easilly extends to contractions on Hilbert spaces. Abve we empoyed Blaschke's factorisation which was used by Schur, while the proof by von Neumann in [1951] avoid Blasche products via certain constructions of unitary operators arising from contractions. The interested reader should consult the text-book [Davies] for this proof as well as 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

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

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

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

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 $0 \le x \le 1$ 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.

D.I On zeros of polynomials

Contents

A: Preliminary results

B: The Laguerre operators A_ζ^n

C: Properties of Laguerre forms

D: Proof of Theorem 0.1

E: Proof of Theorem 0.2 and 0.1

F: Legendre polynomials

Introduction.

The results below are foremost due to Laguerre and the subsequent material is based upon Chapter X in Volume 2 from [Polya and Szegö]. The interested reader should also consult the text-book [XX] which contains a wealth of results concerned with zeros of polynomials. For each $n \geq 1$ we denote by $\mathcal{P}_n$ the space of polynomials $p(z)$ of degree $\leq n$. Consider a pair of such polynomials

$$p(z) = a_0 + a_1 z + \ldots + a_n z^n \quad \text{and} \quad q(z) = b_0 + b_1 z + \ldots + b_n z^n$$

Their convolution is defined by:

$$f * g(z) = a_0 b_0 + a_1 b_1 z + \ldots + a_n b_n z^n \tag{*}$$

The binomial convolution. We can express polynomials in the form

$$p(z) = \sum_{\nu=0}^{\nu=n} \binom{n}{\nu} \cdot a_\nu \cdot z^\nu \quad \text{and} \quad q(z) = \sum_{\nu=0}^{\nu=n} \binom{n}{\nu} \cdot b_\nu \cdot z^\nu$$

Their binomial convolution is defined by:

$$p *_b q = \sum_{\nu=0}^{\nu=n} \binom{n}{\nu} \cdot a_\nu b_\nu \cdot z^\nu \tag{**}$$

Next, if $p = \sum a_k z^k$ is a polynomial in $\mathcal{P}_n$ we set

$$p^*(z) = \sum \binom{n}{k} \cdot a_k z^k \tag{***}$$

With the notations above we shall prove the following results.

0.1 Theorem. *Assume that the zeros of q are real and contained in $(-1, 0)$. For every open and convex set K which contains the origin and all the zeros of p, it follows that K contains the zeros of $p *_b q$.*

0.2 Theorem. *If all the roots of p are real so are the roots of p^*.*

0.3 Theorem. *Assume that the zeros of p are real and that the zeros of q are all real and strictly negative, Then the zeros of $p * q$ are real.*

The proofs are given in section D and E.

A. Preliminary results

A.1 A geometric lemma *Let D be a disc in the complex z-plane. If $\lambda \in \mathbf{C} \setminus D$ we put:*

$$\lambda^*(D) = \{\frac{1}{\lambda - z} \quad : \, z \in D\}$$

Then $\lambda^*(D)$ is an open half plane if λ belongs to the circular boundary of D and an open disc if λ is outside the closed disc $\bar{D}$. In particular $\lambda^*(D)$ is always an open convex set

A.2 Proposition. *Let D be an open disc and $\alpha_1, \ldots, \alpha_n$ some n-tuple of points in D. Then*

$$\frac{1}{n}\sum \frac{1}{\lambda - \alpha_\nu} \neq \frac{1}{\lambda - \zeta} \quad \text{hold for all pairs } \lambda, \zeta \in \mathbf{C} \setminus D$$

A.3 Exercise. Prove the two results above.

A.4 Newton's formula. Consider some $p \in \mathcal{P}_n$. Let $\alpha_1, \ldots, \alpha_n$ be the zeros where multiple zeros can occur and assume that the zeros all belong to some open disc D. Suppose that $\lambda \in \mathbf{C} \setminus D$. Newton's formula for the logarithmic derivative gives:

$$\frac{p'(\lambda)}{p(\lambda)} = \frac{1}{n}\sum \frac{1}{\lambda - \alpha_\nu}$$

The convexity of the set $\lambda^*(D)$ from A.1 gives the inclusion

$$(*) \qquad \frac{p'(\lambda)}{p(\lambda)} \in \lambda^*(D)$$

A.5 Exercise. Deduce from A.4 that if K is a convex set containing the zeros of $p(z)$ then the zeros of $p'(z)$ are also contained in K.

B. The Laguerre operators A_ζ^n

.

Let $\zeta \in \mathbf{C}$ and $n \geq 1$. Define the linear operator from $\mathcal{P}_n$ into $\mathcal{P}_{n-1}$ by

$$(*) \qquad A_\zeta^n(p)(z) = (\zeta - z)p'(z) + np(z)$$

We have for example

$$A_\zeta^n(z^k) = k\zeta z^{k-1} + (n-k)z^k \quad : 1 \leq k \leq n$$

B.1 Proposition. *Let ζ and η be two complex numbers and $n \geq 2$. Then*

$$(*) \qquad A_\eta^{n-1} \circ A_\zeta^n = A_\zeta^{n-1} \circ A_\eta^n$$

Proof. Let $p \in \mathcal{P}_n$. The left hand side becomes:

$$(\eta - z)\big[(\zeta - z)p' + np\big]' + (n-1)(\zeta - z)p' + np) =$$

$$(\eta - z)(\zeta - z)p'' + (n-1)(\eta - z)p' + (n-1)(\zeta - z)p' + (n-1)np$$

The last expression is symmetric in η and ζ and (*) follows.

Composed Laguerre operators. Let $\zeta_1, \ldots, \zeta_n$ be an arbitrary n-tuple of complex numbers. So they are not necessarily distinct. We get the composed operator from $\mathcal{P}_n$ into $\mathbf{C}$ defined by:

$$A_{\zeta_1, \ldots, \zeta_n} = A_{\zeta_1}^1 \circ A_{\zeta_2}^2 \circ \ldots \circ A_{\zeta_n}^n$$

Proposition B.1 gives

$$A_{\zeta_{k-1}}^{k-1} \circ A_{\zeta_k}^k = A_{\zeta_k}^{k-1} \circ A_{\zeta_{k-1}}^k$$

for each $k \geq 2$. Since every permutation of an n-tuple can be achieved as the composition where one makes an interchange of a pair $(k-1, k)$ for some k, it follows that the n-fold composed operator $A_{\zeta_1, \ldots, \zeta_n}$ is symmetric with respect to the n-tuple, i.e. we get the same operator after an arbitrary permutation of $\zeta_1, \ldots, \zeta_n$. Next, for every k we have the the elementary symmetric polynomial of degree k:

$$\Sigma_k(\zeta) = \sum \zeta_{\nu_1} \cdots \zeta_{\nu_k}$$

where the sum extends over all k-tuples $1 \leq \nu_1 < \ldots < \nu_k \leq n$.

B.2 Proposition. *For each $1 \leq k \leq$ one has the equality*

$$(*) \qquad \binom{n}{k} \cdot A_{\zeta_1, \ldots, \zeta_n}(x^k) = n! \cdot \Sigma_k(\zeta)$$

Proof. Notice that (*) is equivalent with

(i) $$A_{\zeta_1,\ldots,\zeta_n}(x^k) = (n-k)! \cdot k! \cdot \Sigma_k(\zeta)$$

We prove (i) by an induction over n. Denote by $\{\Sigma_k^*\}$ the elementary symmetric polynomials in $\zeta_1,\ldots,\zeta_{n-1}$. It is clear that

(ii) $$\Sigma_k(\zeta) = \zeta_n \cdot \Sigma_{k-1}^* + \Sigma_k^*$$

hold for each k. Next, we have:

(iii) $$A_{\zeta_1,\ldots,\zeta_n}(x^k) = A_{\zeta_1,\ldots,\zeta_{n-1}}(k\zeta_n x^{k-1} + (n-k)x^k)$$

By the induction over n the right hand side becomes

$$k\zeta_n \cdot (k-1)! \cdot (n-1-k)! \cdot \Sigma_{k-1}^* + (n-k)k! \cdot (n-1-k)! \cdot \Sigma_k^* =$$
$$k! \cdot (n-k)! \cdot \left[\zeta_n \cdot \Sigma_{k-1}^* + \Sigma_k^*\right] = k! \cdot (n-k)! \cdot \Sigma_k(\zeta)$$

B.3 The Laguerre form Consider a polynomial of degree n expressed in the form:

(*) $$q(x) = \binom{n}{0} \cdot b_0 + \binom{n}{1} \cdot b_1 x + \ldots + \binom{n}{n} \cdot b_n x^n$$

Let $\zeta_1,\ldots\zeta_n$ be the zeros of q which gives

$$q(x) = b_n \cdot \prod (x-\zeta_\nu) = b_n \cdot \sum_{k=0}^{k=n} (-1)^k \cdot \Sigma_k(\zeta) \cdot x^{n-k}$$

It follows that

(i) $$\binom{n}{k} b_{n-k} = b_n \cdot (-1)^k \cdot \Sigma_k(\zeta)$$

Taking a sum over k, the equalities in (i) and Proposition B.2 and the equalities $\binom{n}{k} = \binom{n}{n-k}$ give:

B.4 Proposition. *With q as above and $p(x) = \sum \binom{n}{k} a_k z^k$ we have*

$$\frac{1}{n\,!} \cdot b_n \cdot A_{\zeta_1,\ldots,\zeta_n}(p) = \sum_{k=0}^{k=n} (-1)^k \binom{n}{k} \cdot a_k \cdot b_{n-k}$$

This result suggests the following:

B.5 Definition. *For each pair of polynomials p and q in $\mathcal{P}_n$ we set*

$$\mathrm{Lag}(p,q) = \sum_{k=0}^{k=n} (-1)^k \binom{n}{k} \cdot a_k \cdot b_{n-k}$$

where p and q are expressed as in (*) above.

B.6 Remark. Notice that when p and q are interchanged we have

$$\mathrm{Lag}(q,p) = (-1)^n \cdot \mathrm{Lag}(p,g)$$

Next, with $q(x)$ as in (*) we have

$$\partial^\nu q(0) = \frac{n\,!}{(n-\nu!} \cdot b_\nu$$

A similar formula holds for p and now the reader can verify that

$$\mathrm{Lag}(p,q) = \frac{1}{n\,!} \sum_{k=0}^{k=n} (-1)^k \cdot \partial^k f(0) \cdot \partial^{n-k} g(0)$$

C. Properties of the Laguerre form

The result below is crucial in the study of Laguerre forms.

C.1 Lemma *Let $p(z) \in \mathcal{P}_n$ have all zeros in some open disc D. For every $\zeta \in \mathbf{C} \setminus D$ the zeros of the polynomial $A_\zeta(p)(z)$ belong to D.*

Proof. By definition

$$A_\zeta(p)(z) = (\zeta - z)p'(z) + np(z)$$

Suppose this polynomial has a zero α which does not belong to D. It follows that

$$\text{(i)} \qquad 0 = (\zeta - \alpha)p'(\alpha) + np(\alpha) \implies \frac{p'(\alpha)}{p(\alpha)} = \frac{n}{\alpha - \zeta}$$

Let $z_1, \ldots, z_n$ be the zeros of p where eventual multiple roots are repeated. Then (i) gives:

$$\text{(ii)} \qquad \frac{p'(\alpha)}{p(\alpha)} = \sum \frac{1}{\alpha - z_\nu}$$

It would follow that

$$\frac{1}{n} \cdot \sum_{\nu=1}^{\nu=n} \frac{1}{\alpha - z_\nu} = \frac{1}{\alpha - \zeta}$$

This contradicts Proposition A.1 and Lemma C.1 follows.

C.2 Proposition. *Let p and q be two polynomials of degree n such that $Lag(p, g) = 0$. Then, if D is an open disc which contains the zeros of p, it follows that q has at least one zero in D.*

Proof. Consider first the case $n = 1$ and let ζ_1 be the zero of q while $p(z) = \alpha z - \beta$. The hypothesis entails that

$$0 = A_{\zeta_1}(p) = (\zeta_1 - z)\alpha + \alpha z - \beta = \zeta_1 \cdot \alpha - \beta$$

It follows that ζ_1 is equal to the zero of p and hence it belongs to D. Next, let $n \geq 2$ and suppose that the zeros $\zeta_1, \ldots, \zeta_n$ of q all belong to $\mathbf{C} \setminus D$. By Lemma C.1 the zeros of $A_{\zeta_n}(p)$ belong to D and we can continue until the zero of the linear polynomial $\rho = A_{\zeta_2} \circ \ldots A_{\zeta_n}(p)$ belongs to D. Now ζ_1 is outside D and we get a contraidction from the linear case above since the hypothesis entails that $A_{\zeta_1}(\rho) = 0$.

D. Proof of Theorem 0.1

For a given complex number $\lambda \neq 0$ we set

$$\text{(1)} \qquad q^*(z) = z^n \cdot q(-\frac{\lambda}{z})$$

This gives a new polynomial of degree n. The construction of q^* and Definition B.7 give

$$\text{Lag}(p, q^*) = (-1)^n \cdot \sum \binom{n}{k} \cdot a_k \cdot b_k \cdot \lambda^k$$

The right hand side is the evaluates the polynomial $p *_b q$ at λ which gives the implication

$$\text{(1)} \qquad p *_b q(\lambda) = \text{Lag}(p, g^*) = 0$$

Let λ be a zero $p *_b q$ and consider an open disc D which contains the origin and all the zeros of p. Now (1) gives $\text{Lag}(p, q^*) = 0$ and hence Proposition C.2 gives a point $z_* \in D$ such that $q^*(z_*) = 0$. The inclusion $q^{-1}(0) \subset (-1, 0)$ implies that

$$-1 < \frac{\lambda}{z^*} < 0$$

Hence

$$\text{(2)} \qquad \lambda = az_* \quad \text{for some} \quad 0 < a < 1$$

Since the origin belongs to the convex disc D it follows that $\lambda \in D$. Since this inclusion holds for every open disc D as above elementary geometry shows that λ belongs to the convex set K which finishes the proof of Theorem 0.1.

E. Proof of Theorem 0.2

Let the real roots of p be contained in an interval $(-a,b)$ where a and b both are >0. For every polynomial q with zeros contained $(-1,0)$ Theorem 0.1 gives the inclusion

$$p *_b q^{-1}(0) \subset (-a,b) \tag{1}$$

To profit upon (1) we shall use a special q-polynomial. Put

$$L_n(z) = \partial^n((1+z)^n(z-1)^n) \tag{2}$$

Leibniz's rule gives

$$L_n(z) = \sum_{k=0}^{k=n} \binom{n}{k} \cdot \partial^k((1+z)^n) \cdot \partial^{n-k}((z-1)^n) =$$

$$n\,! \cdot \sum_{k=0}^{k=n} \binom{n}{k}^2 \cdot (1+z)^{n-k} \cdot (z-1)^k$$

As explained in XXx the zeros of L_n belong to $(-1,1)$. Now we define the polynomial

$$Q_n(z) = (z-1)^n \cdot L_n(\frac{z+1}{z-1}) \implies$$

$$Q_n(z) = 2^n \cdot n\,! \cdot \sum_{k=0}^{k=n} \binom{n}{k}^2 \cdot z^k \tag{3}$$

Let c_* be the largest zero of L_n so that $0 < c_* < 1$. Then we see that the zeros of Q_n are real and <0 where the smallest zero is given by the negative number $-c^*$ where

$$c^* = \frac{1+c_*}{1-c_*}$$

If $c > c*$ then the zeros of the polynomial $q(z) = Q_n(cz)$ are contained in $(-1,0)$. The construction of the polynomial p^* in Theorem 0.2 and (3) give

$$p^*(cz) = C \cdot p *_b q(z) \tag{4}$$

for some constant C. Theorem 0.1 implies that the zeros of $p^*(cz)$ are contained in the real interval $(-a,b)$ which implies that the zeros of $p^*(z)$ are contained in $(-ca,cb)$. Here cc^* can be arbitrarily small so actually we conclude that the zeros of p^* are contained in $-c^*a, c^*b)$.

E.1 Exercise. Deduce Theorem 0.3 from Theorem 0.1 and 0.2.

F. Legendre polynomials.

For each $n \geq 1$ an inner product is defined in $\mathcal{P}_n$ by

$$\langle q,p \rangle = \int_{-1}^{1} q(x)p(x) \cdot dx$$

Now $1,x,\ldots,x^n$ is a basis in $\mathcal{P}_n$. where $1,x,\ldots,x^{n-1}$ generate a subspace whose co-dimension is one which gives:

F.1 Proposition. *There exists a unique polynomial* $Q_n(x) = x^n + q_{n-1}x^{n-1} + \ldots + q_0$ *such that*

$$\int_{-1}^{1} x^\nu \cdot Q_n(x) \cdot dx = 0 \leq \nu \leq n-1$$

To find $Q_n(x)$ we consider the polynomial $(1-x^2)^n$ whose derivative of order n belongs to $\mathcal{P}_n$ and partial integrations give:

$$\int_{-1}^{1} x^\nu \cdot \partial^n((x^2-1)^n)) \cdot dx = 0 \leq \nu \leq n-1$$

Notice that the leading coefficient of x^n becomes

$$c_n = 2n(2n-1)\cdots(n+1)$$

Hence we have

$$(*) \qquad Q_n(x) = \frac{1}{c_n}\cdot \partial^n((x^2-1)^n)$$

F.2 Definition. *The Legendre polynomial of degree n is given by*

$$L_n(x) = k_n \cdot \partial^n((x^2-1)^n)$$

where the constant k_n is chosen so that $L_n(1) = 1$.

Since L_n is equal to Q_n up to a constant we have

$$\int_{-1}^{1} x^\nu \cdot L_n(x)\cdot dx = 0 \leq \nu \leq n-1$$

From this we conclude that

$$(**) \qquad \int_{-1}^{1} x^\nu \cdot L_n(x)\cdot L_m(x) dx = 0 \quad n \neq m$$

Thus, $\{L_n\}$ is an orthogonal family with respect to the inner product defined by the integral over $[-1,1]$.

F.3 A generating function. Let w be a new variable and set

$$\phi(x,w) = 1 - 2xw + w^2$$

Notice that $\phi \neq 0$ when $-1 \leq x \leq 1$ and $|w| < 1$. Keeping $-1 \leq x \leq 1$ fixed we have the function

$$w \mapsto \frac{1}{\sqrt{1-2xw+w^2}}$$

Recall that when $|\zeta| < 1$ one has the Newton series

$$\frac{1}{\sqrt{1-\zeta}} = \sum g_n \cdot \zeta^n \quad \text{where} \quad g_n = \frac{3\cdot 5\cdots(2n-1)}{2^n}$$

It follows that

$$\frac{1}{\sqrt{1-2xw+w^2}} = \sum g_n(2xw - w^2)^2$$

With x kept fixed the series is expanded into w-powers, i.e. set

$$\frac{1}{\sqrt{1-2xw+w^2}} = \sum \rho_n(x)\cdot w^n$$

It is easily seen that as x varies then $\rho_n(x)$ is a polynomial of degree n. Moreover, we notice that the coefficient of x^n in $\rho_n(x)$ is equal to

$$g_n \cdot 2^n$$

Next, if $x = 1$ we have

$$\frac{1}{\sqrt{1-2w+w^2}} = \frac{1}{1-w} = \sum w^n$$

From this we conclude that

$$\rho_n(1) = 1 \quad \text{for all} \quad n \geq 0$$

With these notations one has:

F.4 Theorem. *One has $\rho_n(x) = L_n(x)$ for each n, i.e.*

$$\frac{1}{\sqrt{1-2xw+w^2}} = \sum L_n(x)\cdot w^n$$

holds when $-1 \leq x \leq 1$ and $|w| < 1$.

Exercise. Prove this result.

F.5 The trigonometric polynomial $L_n(\cos\theta)$

.

With x replaced by $\cos\theta$ we notice that

$$1-2\cos\theta\cdot w+w^2=(1-e^{i\theta}w)(1-e^{-i\theta}w)\implies$$

$$\frac{1}{\sqrt{1-2\cos(\theta)w+w^2}}=\frac{1}{\sqrt{1-1-e^{i\theta}w)}}\cdot\frac{1}{\sqrt{1-e^{-i\theta}w)}}$$

The last product becomes

$$\sum\sum g_m e^{im\theta}w^m\cdot g_\nu e^{-i\nu\theta}w^\nu$$

Collecting w powers the double sum becomes

$$\sum\gamma_n(\theta)\cdot w^n \quad \text{where} \qquad \gamma_n(\theta)=\sum_{m+\nu=n} g_m g_\nu e^{i(m-\nu)\theta}$$

Since $\{g_m\}$ are real numbers we see that $\gamma_n(\theta)$ is equal to

$$g_n\cdot g_0\cos(\nu\theta)+g_{n-1}g_1\cos((n-2)\theta)+\ldots+g_1g_{n-1}\cos((1-n)\theta)+g_0g_n\cos(-n\theta)$$

By Theorem F.4 the last sum represents $L_n(\cos\theta)$. We have for example

$$L_3(\cos(\theta)=2g_3\cdot\cos(3\theta)+2g_2g_1\cdot\cos(\theta)$$

where we used that $g_0=1$.

F.6 An inequality for $|L_n(x)|$. Since the g-numbers are all ≥ 0 we obtain

$$|L_n(\cos(\theta)|\leq g_ng_0+g_{n-1}g_1+\ldots+g_1g_{n-1}+g_0g_n=P_n(1) \quad : \quad 0\leq\theta\leq 2\pi$$

Hence we have proved

F.7 Theorem. *For each n one has*

$$|L_n(x)|\leq 1 \quad : \quad -1\leq x\leq 1$$

Next, we study the values when $x>1$.

F.8 Theorem. *For each $x>1$ one has $1<L_1(x)<L_2(x)<\ldots$.*
Proof. Put

$$\psi(x,w)=1+\sum_{n=1}^{\infty}\left[L_n(x)-L_{n-1}(x)\right]\cdot w^n$$

By Theorem F.4 this is equal to

(*) $$\frac{1-w}{\sqrt{1-2xw+w^2}}$$

With $x>1$ we set $x=1+\xi$ and notice that

$$1-2xw+w^2=(1-w)^2-2\xi w$$

Hence (*) becomes

(**) $$\frac{1}{\sqrt{1-\frac{2\xi w}{1-w^2}}}=\sum g_n\cdot\frac{(2\xi w)^n}{(1-w^2)^n}=\sum g_n\cdot(2\xi)^n\cdot\frac{w^n}{(1-w^2)^n}$$

Next, for each $n\geq 1$ we notice that the power series of $\frac{w^n}{(1-w^2)^n}$ has positive coefficients. Since $g_n(2\xi)^n>0$ also hold we conclude that (**) is of the form

$$1+\sum_{n=1}^{\infty}q_n(\xi)\cdot w^n \quad \text{where} \quad q_n(\xi)>0$$

Finally, since

$$L_n(1+\xi)-L_{n-1}(1+\xi)=q_n(\xi)$$

we get Theorem F.8

F.9 An L^2-inequality.

Let $n \geq 1$ and denote by $\mathcal{P}_n[1]$ the family of real-valued polynomials $Q(x)$ of degree $\leq n$ for which

$$\int_{-1}^{1} Q^2(x) \cdot dx = 1$$

Then we seek the number

$$\rho(n) = \max_{Q \in \mathcal{P}_n[1]} |Q|_\infty$$

where $|Q|_\infty$ is the maximum norm over $[-1,1]$. To find this ρ-number we use the orthonormal basis from (xx) and write

$$Q(x) = t_0 \cdot L_0^*(x) + \ldots + t_n \cdot L_n^*(x)$$

Since $Q \in \mathcal{P}_n[1]$ we have $t_0^2 + \ldots + t_n^2 = 1$. Recall also that

$$L_\nu^*(x) = \sqrt{\frac{2\nu+1}{2}} \cdot L_\nu(x)$$

Given $-1 \leq x_0 \leq 1$ the Cauchy-Schwarz inequality gives

$$Q^2(x_0) \leq \sum_{\nu=0}^{\nu=n} \frac{2\nu+1}{2} \cdot |L_\nu(x_0)| \leq \sum_{\nu=0}^{\nu=n} \frac{2\nu+1}{2}$$

where the last inequality follows since the maximum norm of each L_ν is one. Finally, we notice that

$$\sum_{\nu=0}^{\nu=n} \frac{2\nu+1}{2} = \frac{(1+n)^2}{2}$$

We conclude that

$$(*) \qquad |Q(x_0)| \leq \frac{n+1}{\sqrt{2}}$$

F.10 Exercise. Show that (*) is sharp and find also the extremal polynomial whose L^2-norm is one while the maximum norm is $\rho(n)$.

G. Fejer's orthogonal polynomials.

Let μ be a Riesz measure in $\mathbf{C}$ with a compact support K. We assume that $K$is not reduced to a finite set, and in general the measure μ is in general complex-valued. Now there exists the Hilbert space $L^2(\mu)$ whose vectors are complex-valued μ-measurable functions f for which

$$\int |f(z)|^2 \, d\mu(z) < \infty$$

The Gram-Schmidt construction gives a unique sequence of polynomials $\{p_n(z)\}$ where $p_n(z)$ is monic of degree n for each non-negative integer n and

$$\int \bar{z}^k \cdot p_n(z) \, d\mu(z) = 0$$

hold for all pairs $0 \leq k < n$. Notice that $p_0(z) = 1$ is the identity function while $p_1(z) = z - a$ and the constant a satisfies

$$\int z \, d\mu(z) = a \cdot \int d\mu(z)$$

In order to locate the zeros of these polynomials we introduce the linear operator T on $L^2(\mu)$ defined by multiplication with z. i.e. $Tf)(z) = zf(z)$, To each $n \geq 1$ we have the orthogonal projection Π_n from $L^2(\mu)$ onto the n-dimensional subspace generarted by $1, z, \ldots, z^{n-1}$ which we denote by V. Keeping n fixed we get the linear operator on V defined by

$$A(f) = \Pi_n(zf) \quad : f \in V$$

Let e_0 be the vector in V defined by the identity function 1. By definition $A(e_0) = z$ with $z \in V$. More generally, consider an A-polynomial of degree $\leq n-1$:

$$q(A) = c_{n-1}A^{n-1} + \ldots c_1 A + c_0 E$$

where E is the identity operator on V. Then it is clear that

$$q(A)(e_0) = q(z)$$

Consider the characteristic polynomial

$$P_A(z) = \det(z \cdot E_V - A) = z^n + d_{n-1}z^{n-1} + \ldots d_1 z + d_0$$

The Cayley-Hamilton equation gives $P_A(A) = 0$ which implies that $P_A(A)(e_0) =$. From the above this gives

$$0 = \Pi_n(z^n) + d_{n-1}z^{n-1} + \ldots d_1 z + d_0 = 0$$

Next, we turn to the monic polynomial $p_n(z)$ and write

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

Then

$$p_n(A)(e_0) = \Pi_n(z^n) + c_{n-1}z^{n-1} + \ldots c_1 z + c_0$$

At the same time (x) means that

$$p_n(A)(e0) = \Pi_n(p_n(z)) = 0$$

where the last equality follos since p_n by construction is $\perp$ to V and hence belongs to the kernel of Π_n. Hence (i) and (ii) imply that

$$c_{n-1}z^{n-1} + \ldots c_1 z + c_0 = d_{n-1}z^{n-1} + \ldots d_1 z + d_0$$

which shows that $p_n(z)$ is equal to the characteristic polynomial of A. Hence the zero-set $p_n^{-1}(0)$ is equal to the spectrum of the linear operator A. At this stge we use the inclusion from § xx which gives

$$\sigma(A) \subset \text{num}(A) = \{\langle Af, f\rangle : f \in S(V)\}$$

whrre $S(V)$ is the set of vctors in V with unit norm. Next, on the Hilbert space $L^2(\mu)$ multiplication with z yields a linear operator denoted by $\widehat{A}$ and we otice that

$$A(f) = \Pi_n \circ \widehat{A} \circ \Pi_n$$

and by the remark in § xx one gets the inclusion

$$\text{num}(A) \subset \text{num}(\widehat{A})$$

The case when $\mu$is a probability measure. Assume this and consider a function $f \in L^2(\mu)$ with unit norm. Now

$$\langle \widehat{A}(f), f\rangle \int z \cdot |f(z)|^2 \, d\mu(z)$$

The right hand side can be approximated by Stieltjes sums

$$\sum \int E_k \, |f(z)|^2 \, d\mu(z) \cdot z_k$$

where $\{E_k\}$ are disjoint Borel sets and $z_k \in E_k$. With $\{c_k = \int E_k \, |f(z)|^2 \, d\mu(z)\}$ we have $\sum c_k = 1$ since f has unit norm. Hence every Siteltjes' sum in (x) belongs to the convex hull of K which shows that the numerical range of $\widehat{[A}$ is contained in this convex hull. Together with th inclusion xx we arrive at the following:

Fejer's theorem. *If μ is a probability measure the zeros of the orthogonal polynomials from (xx) are contained in the convex hul of K.*

I:E. Fourier series

Contents

A: The kernels of Dini, Fejer and Jackson

B: Legendre polynomials

C. The space $\mathcal{T}_n$

D. Tchebysheff polynomials

E. Fejer series and Gibbs phenomenon

F. Partial Fourier sums and convergence in the mean

G. Best approximation by trigonometric polynomials.

Introduction.

Fourier series were invented by Fourier around 1810. We expose results in the 1-dimensional case and remark only that one also constructs Fourier series in several variables of functions $f(x_1,\ldots,x_n)$ which are 2π-periodic with respect to each variable in $\mathbf{R}^n$. Recall the construction of Fourier's partial sums in the case $n=1$. Let $f(\theta)$ be a complex-valued and continuous function defined on the interval $\{0\leq\theta\leq 2\pi\}$ which satisfies $f(0)=f(2\pi)$. For every integer n we set

$$\widehat{f}(n)=\frac{1}{2\pi}\cdot\int_0^{2\pi} e^{-in\phi}f(\phi)\cdot d\phi$$

One refers to $\{\widehat{f}(n)\}$ as the Fourier coefficients of f and Fourier's partial sum function of degree N is defined by

$$S_N(\theta)=\sum_{n=-N}^{n=N}\hat{f}(n)\cdot e^{in\theta} \tag{0.0}$$

The question arises if

$$\lim_{N\to\infty}\max_\theta |S_N(\theta)-f(\theta)|=0 \tag{0.1}$$

So (0.1) means that Fourier's partial sums convege uniformyly to f. Examples where (0.1) fails were discovered at an early stage and lead to Gibb's phenomenon. More precisely, there exists contionuous functions f where the uniform not only fails, but for certain θ-values the sequence $\{S_N(\theta)\}$ has no limit. A relaxed condition is to ask if the pointwise limit

$$\lim_{N\to\infty} S_NN(\theta)=f(\theta) \tag{0.2}$$

exists for all θ outside a null set in the sense of Lebesgue, i.e. is it true that Fourier's partial sums converge almost everywhere to f. This question was open for more than a half century until the affirmative answer was established by Carleson in 1965. This result constitutes one of the greatest achievements ever in analysiand the proof goes beyond the level of these notes. The reader may consult Carleson's article [xxx] for the proof of almost everywhere convergence which includes a remarkable inequality which goes as follows: Let ℓ^2 be the Hilbert space of sequences of complex numbers $c_0,c_1,\ldots$ such that $\sum|c_n|^2<\infty$. To every such a sequence we introduce trigonometric polynomials

$$S_N(\theta)=\sum_{n=0}^{n=N} c_ne^{in\theta}$$

Define the maximal function by

$$S^*(\theta)=\max_{N\geq 1}|S_N(e^{i\theta})|$$

Carleson's inequality. *There exists a constant C such that the following hold when $\{c_n\} \in \ell^2$:*

$$\int_0^{2\pi} S^*(\theta)^2 \, d\theta \leq C \cdot \sum_{n=1}^{\infty} |c_n|^2$$

Fejer's inequality. Several inequalities for trigonometric polynomials were established in [Fejer] where a central issue is to construct trigonometric polynomials expressed by a sine series which are ≥ 0 on the interval $[0, \pi]$. Consider as an example is the sine-series

$$S_n(\theta) = \sum_{k=1}^{k=n} \frac{\sin k\theta}{k}$$

Here $\{S_n(\theta)\}$ are trigonometric polynomials which are odd functions of θ. Consider the following analytic function in the unit disc D:

$$f(z) = \sum_{k=1}^{\infty} \frac{z^k}{k}$$

Notice that this series represents the analytic function in D given by $\log(1-z)$. This complex log-function extends analytically across the unit circle T outside $\{z = 1\}$. If $0 < \theta < \pi$ we notice that

$$\mathfrak{Im}(f(e^{i\theta}) = -\arg(1 - e^{i\theta}) = \frac{\pi - \theta}{2}$$

At the same time $e^{ik\theta} = \sin k\theta$ and therefore

$$\sum_{k=1}^{\infty} \frac{\sin k\theta}{k} = \frac{\pi - \theta}{2} \quad : 0 < \theta < \pi$$

Moreover, there exist pointwise limits

$$\lim_{n \to \infty} S_n(\theta) = \frac{\pi - \theta}{2} \quad : 0 < \theta < \pi$$

At the same time $S_n(0) = 0$ for every n so one cannot expect that the pointwise convergence for small positive θ holds uniformly. Fejer proved the following:

0.2 Theorem. *For every $n \geq 1$ one has the inequality*

$$0 < S_n(\theta) \leq 1 + \frac{\pi}{2} \quad : \quad 0 < \theta < \pi$$

The upper bound was proved by in [Fej] and Fejer conjectured that $S_n(\theta)$ stays positive on $(0, \pi)$. This was later confirmed by Jackson in [xx] and Cronwall in [xx]. Here is an occasion to use a computer and plot graphs of the functions $\{S_n(\theta)\}$ to analyze the rate of convergence when $\theta \simeq 0$ and also confirm the inequality in Fejer's theorem numerically, i.e. plot graphs of $S - N(\theta)$ for large values of n and check the validity of the inequalities above numerically.

0.3 A result by Carleman. Let $f(\theta)$ be a 2π-periodic and continuous function. If $\epsilon > 0$ and $N \geq 1$ we denote by $\rho(N; \epsilon)$ the number of integers $0 \leq n \leq N$ for which the maximum norm

$$\max_\theta |S_n(\theta) - f(\theta)| \geq \epsilon$$

With these notations we prove in § x that

$$\text{(i)} \qquad \lim_{N \to \infty} \frac{\rho(N; \epsilon)}{N + 1} = 0$$

hold for every $\epsilon > 0$. It means that Gibb's phenomenon from a statistical point of view is exceptional, i.e. "failure of convergence" occurs only for a sparse subsequence of Fourier's partial sums. Actually (i) is a consequence of a more precise result which goes as follows: The continuous function f is uniformly continuos and put:

$$\omega_f(\delta) = \max |f(\theta_1) - f(\theta_2)| \quad : |\theta_1 - \theta_2| \leq \delta$$

Theorem. *There exists an absolute constant K such that the following hold for every 2π-periodic continuous function f whose maximum norm is ≤ 1*

$$(*) \qquad \frac{\rho(N;\epsilon)}{N} \leq \frac{K}{\epsilon^2} \cdot \Big(\frac{1}{N} + \omega_f\big(\frac{1}{N}\big)^2\Big)$$

Remark. Since $\frac{1}{N} + \omega_f(\frac{1}{N})^2$ tends to zero as $N \to +\infty$ and (*) holds for each $\epsilon > 0$ we get (i).

0.4 Fekete's inequality. The interplay between Fourier series and analytic functions in the unit disc $D = \{|z| < 1\}$ leads to many interesting results. Consider as in (0.2) the sine series

$$\phi(\theta) = \sum_{n=1}^{\infty} \frac{\sin n\theta}{n}$$

In D we have the analytic function

$$f(z) = \sum_{n=1}^{\infty} \frac{z^n}{n} = -\log(1-z)$$

Notice that

$$f(x) = -\log(1-x)$$

tends to $+\infty$ as $x \to 1$ along the real axis. The Taylor polynomials

$$S_N(z) = \sum_{n=1}^{n=N} \frac{z^n}{n}$$

attain their maximum norms on the closed unit disc when $z = 1$ and here

$$S_N(1) = \sum_{n=1}^{n=N} \frac{1}{n} \simeq \log N$$

At the same time we have seen in (0.2) that $\mathfrak{Im}(f(z))$ is a bounded function in D. Fekete proved that the example above is extremal in the following sense:

Theorem. *There exists an absolute constant C such that if $g(z) = \sum c_n z^n$ is an analytic function in D for which the maximum norm of $\mathfrak{Im}(g(z))$ is ≤ 1, then*

$$\max_\theta \Big|\mathfrak{Re} \sum_{n=0}^{n=N} c_n \cdot e^{in\theta}\Big| \leq C \cdot \log N \quad : N \geq 2$$

Fekete's result will be proved in § XX during a closer study about analytic functions in the unit disc and their associated Fourier series.

0.5 Bernstein's example. A remarkable construction was given by S. Bernstein in the article [Comptes Rendus 1914]. Let p be a prime number of the form $4\mu+1$ where μ is a positive integer. For each integer $n \geq 1$ we have the Legendre symbol $L(n;p)$ which is $+1$ is k has a quadratic remainder modolu p and otherwise $L(n;p) = -1$. Define the trigonometric polynomial

$$\mathcal{B}_p(\theta) = \frac{2}{p^{\frac{3}{2}}} \cdot \sum_{n=1}^{n=p-1} (p-n) \cdot L(n;p) \cdot \cos n\theta$$

Then Bernstein proved that

$$\text{(i)} \qquad \max_\theta |\mathcal{B}_p(\theta)| \leq 1$$

At the same time we notice that

$$\text{(ii)} \qquad \frac{2}{p^{\frac{3}{2}}} \cdot \sum_{n=1}^{p-1} |(p-n) \cdot L(n;p)| = \frac{p-1}{\sqrt{p}} \simeq \sqrt{p}$$

Bernstein's trigonometric polynomials have extremal properties. For consider an arbitrary cosine series

$$u(\theta)=\sum_{n=1}^{n=N} a_n\cdot\cos\theta$$

Now the L^2-integral

$$\frac{1}{\pi}\int_0^{2\pi} u^2(\theta)\,d\theta=\sum_{n=1}^{n=N} a_n^2$$

If the maximum norm of u is one the L^2 integral is majorized by 2 and the Cauchy-Schwarz inequality gives

$$\sum|a_n|\leq\sqrt{2\cdot N}$$

Bernstein's example shows that this inequality is essentially sharp. Another notable phenomenon in Bernstein's example is the following: We have

$$\int_0^{2\pi}\mathcal{B}_p^2(\theta)\,d\theta=\frac{4}{p^3}\cdot\pi\cdot\sum_{n=1}^{n=p-1} n^2$$

The right hand side is bounded by an absolute constant C. Hence the maximum norm and the L^2-norm of $\mathcal{B}_p$ are comparable , i.e. there is a fixed constant $0<c<1$ such that

$$\frac{c}{\pi}\leq\frac{||\mathcal{B}_p||_\infty}{||\mathcal{B}_p||_2}\leq\frac{1}{\pi c}$$

Remark. Bernstein's construction was based upon arithmetic. Later Salem proved that the Bernstein's example is generic in the sense that by random choice of signs in a given sequence $\{a_k\}$ with prescribed L^2-norm equal to one, the corresponding maximum norms of the partial sums are not so large with "high probabilities". To give an example: Let $N\geq 2$ and consider the family $\mathcal{F}_N$ of cosine series

$$f(\theta)=\frac{1}{\sqrt{N}}\cdot\sum_{n=1}^{n=N}\epsilon_n\cdot\cos n\theta$$

Here$\{\epsilon_n\}$ is random sequence where each ϵ_n is +1 or -1. Notice that the L^2-integrals

$$\int_0^{\pi} f(\theta)^2\,d\theta=\frac{\pi}{2}$$

Above one has a sample space where each choice of a siugn-sequence $\{\epsilon_n\}$ produces a function in $\mathcal{F}_N$. To be precise, we get 2^N many functions in this family. The evaluation at $\theta=0$ corerssponds to a Bernoulli trial, i.e,. tossing a coin N times and measure the difference of heads and tails, divided by $\sqrt{N}$. Here the centeral limit theorem applies, i.e by de Moivre's discovery from 1733, the random outcome of the numbers $\{f(0):F\in\mathcal{F}\}$ is expressed by a discrete random variable whose densities converge to the normal distribution as $N\to\infty$.

A more involved study arises when one regards vakues of the f-functions over the whole interval $[0,\pi]$. Fir example, one can consider the random variabke on the sampåle space above defined by

$$f\mapsto\max_{0\leq\theta\leq\pi}|f(\theta)|$$

Results about the asymptotic behaviour of the distributions of these random variables as N increases have been obtained by Salem and inspired much later work. The reader may consult [Salem] and [Kahane] for an account about Fourier series with random coefficients. It goes without saying that this leads to a quite involved theory.

Outline of contents.

Section A contains basic material about Fourier series where the kernels of Dini and Fejer are introduced. At the end of § A we construct the Jackson kernel which give approximations of a given periodic function f by trigonometric polynomials where the rate of approximation is controlled by the modulos of continuity of f. Sections B-C are devoted to results about extremal polynomials. A complex version appears in § D where Theorem D.4 relates the transfinite diameter of compact subsets of $\mathbf{C}$ with Tchebysheff polynomials. § F is devoted to a result by Carleman about convergence in the mean of partial Fourier sums. From a statistical point of view this result confirms the convergence of Fourier's partial sums where Theorem F.2 gives an absolute constant K such that for every 2π-periodic and continuous function f whose maximum norm is ≤ 1, the following inequality holds for every positive integer n and each $0 < \delta < \pi$ where $\{s_\nu\}$ are Fourier's partial sums of f:

$$\sqrt{\frac{1}{n+1} \cdot \sum_{\nu=0}^{\nu=n} ||s_\nu - f||^2} \leq \frac{1}{\sqrt{n+1}} \cdot [n^{1+1/2} \cdot \delta \cdot \omega_f(\delta) + 2K\delta^{-1/2} + K]$$

where $\{||s_\nu - f||\}$ denote maximum norms over $[0, 2\pi]$ and ω_f the modulos of continuity. If $\epsilon > 0$ we take $\delta = \frac{1}{\epsilon n}$ for large n, the left hand side is majorised by

$$\frac{\omega_f(1/\epsilon n)}{\epsilon} + 2K\sqrt{\epsilon} + \frac{1}{\sqrt{n+1}} \cdot K$$

Keeping ϵ fixed while n increases this tends to zero which entials that "with high probability" the maximum norms of $|||s_\nu - f||$ are small as ν varies over large integer intervals.

Section § H treats results due to de Vallé Poussin about best approximations by trigonometric polynomials of prescribed degree where one starts with some real-valued and continuous 2π-periodic function f. If $n \geq 1$ we denote by $\mathcal{T}_n$ the $2n + 1$)-dimensional real vector space of trigonometric polynomials of degree $\leq n$, i.e. functions of the form

$$P(x) = \frac{a_0}{2} + \sum_{\nu=1}^{\nu=n} a_\nu \cdot \cos \nu x + \sum_{\nu=1}^{\nu=n} b_\nu \cdot \sin \nu x$$

The best approximation of degree n is defined by:

(*) $$\rho_f(n) = \min_{P \in \mathcal{T}_n} ||f - P||$$

Among the results in § H we mention the following lower bound inequality expressed by the Fourier coefficients of f defined by

$$\widehat{f}(n) = \frac{1}{2\pi} \int_0^{2\pi} e^{-inx} \cdot f(x)\, dx$$

Theorem. *For each $n \geq 1$ one has the inequality*

$$\rho_f(n) \geq |\widehat{f}(n+1)| - \sum_{j=1}^{\infty} |\widehat{f}((n+1)(2j+1))|$$

We remark that this lower bound of the ρ-numbers are of special interest when the Fourier coefficients of f have many gaps.

A: The kernels of Dini, Fejer and Jackson

Denote by $C^0_{\text{per}}[0,2\pi]$ the family of complex-valued continuous functions $f(\theta)$ on $[0,2\pi]$ which satisfy $f(0)=f(2\pi)$. The Fourier coefficients of such a function are defined by:

$$\widehat{f}(n)=\frac{1}{2\pi}\cdot\int_0^{2\pi}e^{-in\phi}f(\phi)\cdot d\phi$$

where n are integers. For each non-negative integer N we define Fourier's partial sum

$$S^f_N(\theta)=\sum_{n=-N}^{n=N}\widehat{f}(n)\cdot e^{in\theta} \tag{A.0}$$

A.1.The Dini kernel. If $N\geq 0$ we set

$$D_N(\theta)=\frac{1}{2\pi}\sum_{n=-N}^{n=N}e^{in\theta}$$

A.2 Proposition. *One has the formula*

$$D_N(\theta)=\frac{1}{2\pi}\cdot\frac{\sin((N+\frac{1}{2})\theta)}{\sin\frac{\theta}{2}} \tag{A.2.1}$$

Proof. We have

$$\sum_{n=-N}^{n=N}e^{in\theta}=e^{-iN\theta}\cdot\sum_{n=0}^{n=2N}e^{in\theta}=e^{-iN\theta}\cdot\frac{e^{i(2N+1)\theta}-1}{e^{i\theta}-1}=$$

$$e^{-iN\theta-i\theta/2}\cdot\frac{e^{i(2N+1)\theta}-1}{2i\cdot\sin\theta/2}=\frac{2i\cdot\sin((N+1/2)\theta)}{2i\cdot\sin\theta/2}$$

and (A.2.1) follows after division with $2i$.

A.3 Exercise. Show that the following hold for each $N\geq 0$:

$$S^f_N(\theta)=\int_0^{2\pi}D_N(\theta-\phi)\cdot f(\phi)\cdot d\phi=\int_0^{2\pi}D_N(\phi)\cdot f(\theta+\phi)\cdot d\phi$$

A.4 The Fejer kernel. For each $N\geq 0$ we set

$$\mathcal{F}_N(\theta)=\frac{D_0(\theta)+\ldots+D_N(\theta)}{2\pi(N+1)}$$

A.5 Proposition *One has the formula*

$$\mathcal{F}_N(\theta)=\frac{1}{2\pi(N+1)}\cdot\frac{1-\cos((N+1)\theta)}{2\cdot\sin^2(\frac{\theta}{2})} \tag{A.5.1}$$

Proof. To each $\nu\geq 0$ we have $\sin((\nu+1/2)\theta)=\mathfrak{Im}\big[e^{i(\nu+1/2)\theta)}\big]$. Hence $F_N(\theta)$ is the imaginary part of

$$\frac{1}{2\pi(N+1)}\cdot\frac{e^{i\theta/2}}{\sin(\theta/2)}\cdot\sum_{\nu=0}^{\nu=N}e^{i\nu\theta}$$

Next, we have

$$e^{i\theta/2}\cdot\sum_{\nu=0}^{\nu=N}e^{i\nu\theta}=e^{i\theta/2}\cdot\frac{e^{i(N+1)\theta}-1}{e^{i\theta-1}}=\frac{e^{i(N+1)\theta}-1}{2i\cdot\sin(\theta/2)}$$

Since $i^2=-1$ we see that the imaginary part of the last term is equal to

$$\frac{1-\cos((N+1)\theta)}{2\cdot\sin(\frac{\theta}{2})}$$

and (A.5.1) follows.

A.6 Fejer sums. For each f and every $N \geq 0$ we set

$$F_N^f(\theta) = \int_0^{2\pi} \mathcal{F}_N(\phi) \cdot f(\theta + \phi) \cdot d\phi$$

A.7 An inequality. Notice that the Fejer kernels are non-negative. If $a > 0$ and $a \leq \theta \leq 2\pi - a$ we have the inequality

$$\text{(i)} \qquad \sin^2(\theta/2) \geq \sin^2(a/2)$$

Let f be given and denote by $M(f)$ the maximum norm of $|f(\theta)|$ over $[0, 2\pi]$. Then (i) gives

$$\int_a^{2\pi - a} \mathcal{F}_N(\phi) \cdot f(\theta + \phi) \cdot d\phi \leq$$

$$\text{(A.7.1)} \qquad \frac{M}{2\pi(N+1) \cdot \sin^2(a/2)} \int_a^{2\pi - a} (1 - \cos(N\phi)) \cdot d\phi \leq \frac{2M}{(N+1) \cdot \sin^2(a/2)}$$

A.8 Exercise. Given some θ_0 and $0 < a < \pi 0$ we set

$$\omega_f(a) = \max_{|\theta - \theta_0| \leq a} |f(\theta) - f(\theta_0)|$$

Use (A.7.1) to prove that

$$|F_N^f(\theta_0) - f(\theta_0)| \leq \frac{2M}{(N+1) \cdot \sin^2(a/2)} + \omega_f(a)$$

Conclude by *uniform continuity* of the function f on $[0, 2\pi]$ implies that the sequence $\{F_N^f\}$ converges uniformly to f over the interval $[0, 2\pi]$.

A.9 The case when f is real-valued. When f is real-valued the Fourier series takes the form

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cdot \cos kx + \sum_{k=1}^{\infty} b_k \cdot \sin kx$$

Here $a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) \cdot dx$ and when $k \geq 1$ one has

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cdot \cos kx \cdot dx \quad : b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cdot \sin kx \cdot dx$$

Fourier's partial sum functions become

$$\text{(A.9.1)} \qquad S_n(f) = \frac{a_0}{2} + \sum_{k=1}^{k=n} a_k \cdot \cos kx + \sum_{k=1}^{k=n} b_k \cdot \sin kx$$

The Jackson kernel

.

Above we proved that the Fejer sums converge uniformly to f. One may ask if there exists a constant C which is independent of both f and of n such that

$$(*) \qquad \max_\theta ||f(\theta) - F_n^f(\theta)| \leq C \cdot \omega_f(\frac{1}{n})$$

Examples show that (*) does not hold in general. To obtain a uniform constant C one must include an extra factor.

A.10 Exercise. Use (A.7-8) to show that there exists an absolute constant C such that

$$(**) \qquad ||f - \mathcal{F}_n(f)|| \leq C \cdot \omega_f(\frac{1}{n}) \cdot \big(1 + \log^+ \frac{1}{\omega_f(\frac{1}{n})}\big)$$

hold for all continuous 2π-periodic functions f.

To attain (*) D. Jackson introduced a new kernel in his thesis *Über die Genauigkeit der Annährerung stegiger funktionen durch ganze rationala funktionen* from Göttingen in 1911. To each 2π-periodic and continuous function $f(x)$ on the real line and every $n \geq 1$ we set

$$\mathcal{J}_n^f(x) = \frac{3}{2\pi} \cdot \int_{-\infty}^{\infty} f(x + \frac{2t}{n}) \cdot (\frac{\sin t}{t})^4 \cdot dt$$

A.11 Theorem. *The function $\mathcal{J}_n^f(x)$ is a trigonometric polynomial of degree $2n-1$ at most and one has the inequality*

$$\max_x |f(x) - \mathcal{J}_n^f(x)| \leq (1 + \frac{6}{\pi}) \cdot \omega_f(\frac{1}{n})$$

Proof. The variable substitution $t \to nt$ gives

$$(1) \qquad \mathcal{J}_n^f(x) = \frac{3}{2\pi n^3} \cdot \int_{-\infty}^{\infty} f(x + 2t) \cdot (\frac{\sin nt}{t})^4 \cdot dt$$

Since $t \mapsto f(x+2t) \cdot sin^4\, nt$ is π-periodic it follows that (1) is equal to

$$(2) \qquad \frac{3}{2\pi n^3} \cdot \int_0^{\pi} f(x + 2t) \cdot \sum_{k=-\infty}^{\infty} \frac{\sin^4(nt)}{(k\pi + t)^4} \cdot dt$$

Next, recall from § XX that

$$\frac{1}{\sin^2 z} = \sum_{k=-\infty}^{\infty} \frac{1}{(z + k\pi)^2}$$

Taking a second derivative when $z = t$ is real it follows that

$$(3) \qquad \partial_t^2(\frac{1}{\sin^2 t}) = \frac{1}{6} \cdot \sum_{k=-\infty}^{\infty} \frac{1}{(t + k\pi)^4}$$

Hence we obtain

$$(*) \qquad \mathcal{J}_n^f(x) = \frac{1}{4\pi n^3} \cdot \int_0^{\pi} f(x + 2t) \cdot \sin^4(nt) \cdot \partial_t^2(\frac{1}{\sin^2 t})\, dt$$

Next, the function

$$\sin^4(nz) \cdot \partial_z^2(\frac{1}{\sin^2 z})$$

is entire and even and the reader may verify that it is a finite sum of entire cosine-functions which implies that the Jackson kernel is expressed by a finite sum of integrals:

$$(4) \qquad \mathcal{J}_f^n(x) = \sum_{k=0}^{2n-1} c_k \int_0^{2\pi} f(u) \cdot \cos\, k(x - u))\, du$$

In particular $\mathcal{J}_f^n(x)$ is a trigonometric polynomial of degree $2n-1$ a most. Integration by parts give the equality

$$\int_{-\infty}^{\infty} (\frac{\sin nt}{t})^4\, dt = \frac{1}{6}\int_0^{\pi} \sin^4 t\cdot\partial_t^2(\frac{1}{\sin^2 t})\, dt = \frac{4}{3}\int_0^{\pi} \cos^2 t\, dt = \frac{2\pi}{3} \tag{5}$$

Next, we leave it to the reader to verify the inequality

$$\frac{3}{2\pi}\int_{-\infty}^{\infty} (1+2|t|)\cdot(\frac{\sin t}{t})^4 \cdot dt \leq 1+\frac{6}{\pi} \tag{6}$$

From the above where we use (1) and (*) it follows that

$$\mathcal{J}_n^f(x) - f(x) = \frac{3}{2\pi}\cdot\int_{-\infty}^{\infty} [f(x+\frac{2t}{n}) - f(x)]\cdot(\frac{\sin t}{t})^4\cdot dt \tag{7}$$

Now

$$|f(x+\frac{2t}{n}) - f(x)| \leq \omega_f(\frac{2t}{n}) \leq (2|t|+1)\cdot\omega_f(\frac{1}{n})$$

where the last equality follows from Lemma XX. Hence (7) gives

$$\max_x |\mathcal{J}_n^f(x) - f(x)| \leq \omega_f(\frac{1}{n})\cdot\frac{3}{2\pi}\cdot\int_{-\infty}^{\infty}(2|t|+1)\cdot(\frac{\sin t}{t})^4\cdot dt$$

Finally, by (6) the last factor is majorized by $1+\frac{6}{\pi}$ and Jackson's inequality follows.

A.12 A lower bound for polynomial approximation.

Denote by $\mathcal{T}_n$ the linear space of trigonometric polynomials of degree $\leq n$. For a 2π-periodic and continuous function f we put

$$\rho_f(n) = \min_{T\in\mathcal{T}_n} ||f-T||$$

where $||\cdot||$ denotes the maximum norm over $[0,2\pi]$. We shall establish a lower bound for the ρ-numbers when certain sign-conditions hold for Fourier coefficients. In general, let f be a periodic function and for each positive integer n we find $T\in\mathcal{T}_n$ such that $||f-T|| = \rho_f(n)$. Since Fejer kernels do not increase maximum norms one has

$$||F_k^f - F_k^T|| \leq \rho_f(n) \tag{i}$$

for every positive integer k. Apply this with $k=n$ and $k=n+p$ where p is another positive integer. If $T\in\mathcal{T}_n$ the equation from Exercise XX gives

$$T = \frac{(n+p)\cdot\mathcal{F}_{n+p}(T) - n\cdot\mathcal{F}_n(T)}{p} \tag{ii}$$

Since (i) hold for n, $n+p$ and $||f-T||\leq\rho_f(n)$, the triangle inequality gives

$$||f - \frac{(n+p)\cdot\mathcal{F}_{n+p}(f) - n\cdot\mathcal{F}_n(f)}{p}|| \leq 2\cdot\frac{n+p}{p}\cdot\rho_f(n) \tag{iii}$$

Next, by the formula (§ xx) it follows that (iii) gives

$$||f - \frac{S_n(f)+\cdots S_{n+p-1}(f)}{p}|| \leq 2\cdot\frac{n+p}{p}\cdot\rho_f(n)$$

In particular we take $p=n$ and get the inequality

$$||f - \frac{S_n(f)+\cdots S_{2n-1}(f)}{n}|| \leq \frac{4}{n}\cdot\rho_f(n) \tag{*}$$

A.12 A special case. Assume that $f(x)$ is an even function on $[-\pi,\pi]$ which gives a Fourier series:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k\cdot\cos kx$$

A.12 Proposition *Let f be even as above and assume that $a_k \leq 0$ for every $k \geq 1$. Then the following inequality holds for every $n \geq 1$:*

$$f(0) - \frac{S_n(f)(0) + \cdots S_{2n-1}(f)(0)}{n} \leq -\sum_{k=2n}^{\infty} a_k$$

The easy verification is left to the reader. Taking the maximum norm over $[-\pi, \pi]$ it follows from (*) that

$$(**) \qquad \rho_f(n) \geq \frac{n}{4} \cdot \sum_{k=2n}^{\infty} |a_k|$$

holds when the sign conditions on the Fourier coefficients above are satisfied. Notice that (**) means that one has a lower bound for polynomial approximations of f.

A.13 The function $f(x) = \sin|x|$ It is obvious that

$$\omega_f(\frac{1}{n}) = \frac{1}{n}$$

Next, the periodic function $f(x($ is even and hence we only get a cosine-series. For each positive integer m we have:

$$a_k = \frac{2}{\pi} \int_0^{\pi} \sin x \cdot \cos kx \cdot dx$$

To evaluate these integrals we use the trigonometric formula

$$sin\,(k+1)x - \sin(k-1)x = 2\sin x \cdot \cos kx$$

Now the reader can verify that $a_\nu = 0$ when ν is odd while

$$a_{2k} = -\frac{4}{\pi} \cdot \frac{1}{2k^2 - 1}$$

Hence the requested sign conditions hold and (**) entails that

$$\rho_f(n) \geq \frac{n}{\pi} \cdot \sum_{k=n}^{\infty} \frac{1}{2k^2 - 1}$$

Here the right hand side is $\geq \frac{C}{n}$ for a constant C which is independent of n. So this example shows that the inequality (*) in § A.11 is sharp up to a multiple with a fixed constant.

B. Legendre polynomials.

If $n \geq 1$ we denote by $\mathcal{P}_n$ the linear space of real-valued polynomials of degree $\leq n$. A bilinear form is defined by

$$\langle q, p \rangle = \int_{-1}^{1} q(x)p(x) \cdot dx$$

Since $1, x, \ldots, x^{n-1}$ generate a subspace of co-dimension one in $\mathcal{P}_n$ we get:

B.1 Proposition. *There exists a unique* $Q_n(x) = x^n + q_{n-1}x^{n-1} + \ldots + q_0$ *such that*

$$\int_{-1}^{1} x^\nu \cdot Q_n(x) \cdot dx = 0 \leq \nu \leq n-1$$

To find $Q_n(x)$ we consider the polynomial $(1-x^2)^n$ which vanishes up to order n at the end-points 1 and -1. Its derivative of order n gives a polynomial of degree n and partial integrations show that

$$\int_{-1}^{1} x^\nu \cdot \partial^n((x^2-1)^n)) \cdot dx = 0 \leq \nu \leq n-1$$

The leading coefficient of x^n in $\partial^n((x^2-1)^n))$ becomes

$$c_n = 2n(2n-1)\cdots(n+1)$$

Hence we have

$$Q_n(x) = \frac{1}{c_n} \cdot \partial^n((x^2-1)^n)$$

B.2 Definition. *The Legendre polynomial of degree* n *is given by*

$$P_n(x) = k_n \cdot \partial^n((x^2-1)^n)$$

where the constant k_n *is determined so that* $P_n(1) = 1$.

Since P_n is equal to Q_n up to a constant we still have

$$\int_{-1}^{1} x^\nu \cdot P_n(x) \cdot dx = 0 \leq \nu \leq n-1$$

From this we conclude that

$$\int_{-1}^{1} x^\nu \cdot P_n(x) \cdot P_m x) dx = 0 \quad n \neq m$$

Thus, $\{P_n\}$ is an orthogonal family with respect to the inner product defined by the integral over $[-1, 1]$.

B.3 A generating function. Let w be a new complex variable and set

$$\phi(x, w) = 1 - 2xw + w^2$$

The reader should check that $\phi \neq 0$ when $-1 \leq x \leq 1$ and $|w| < 1$. Keeping $-1 \leq x \leq 1$ fixed we have the function

$$w \mapsto \frac{1}{\sqrt{1-2xw+w^2}}$$

defined when $|w| < 1$. Next, as $|\zeta| < 1$ one has the Newton series

$$\frac{1}{\sqrt{1-\zeta}} = \sum g_n \cdot \zeta^n \quad \text{where} \quad g_n = \frac{3 \cdot 5 \cdots (2n-1)}{2^n}$$

It follows that

$$\frac{1}{\sqrt{1-2xw+w^2}} = \sum g_n \cdot (2xw - w^2)^2$$

With x kept fixed we get another series which is expanded into w-powers

$$\frac{1}{\sqrt{1-2xw+w^2}} = \sum \rho_n(x) \cdot w^n$$

It is easily seen that as x varies then $\rho_n(x)$ is a polynomial of degree n and the coefficient of x^n in $\rho_n(x)$ is equal to

$$g_n \cdot 2^n$$

Next, if $x = 1$ we have

$$\frac{1}{\sqrt{1-2w+w^2}} = \frac{1}{1-w} = \sum w^n$$

From this we conclude that

$$\rho_n(1) = 1 \quad \text{for all} \quad n \geq 0$$

B.4 Theorem. *One has the equality $\rho_n(x) = P_n(x)$ for each n, i.e.*

$$\frac{1}{\sqrt{1-2xw+w^2}} = \sum P_n(x) \cdot w^n \quad \text{holds when} \quad -1 \leq x \leq 1 : \ |w| < 1$$

B.5 Exercise. Prove this result. The hint is to regard the integrals

$$J_k(w) = \int_{-1}^{1} \frac{x^k}{\sqrt{w^2 - 2xw + 1}}\, dx$$

for non-negative integers and show that the power series of $J_k(w)$ at $w = 0$ is reduced to a polynomial of degree k for each $k \geq 0$. This is proeved via partial integration which for small $w \neq 0$ gives gives

$$J_k(w) = w^{-1} \cdot \sqrt{w^2 - 2xw - 1} \cdot x^k\big|_{-1}^{1} + + \frac{k}{w} \cdot \int_{-1}^{1} (w^2 - 2xw + 1) \cdot \frac{x^{k-1}}{\sqrt{w^2 - 2xw + 1}}, \, dx$$

It follows that

$$(1+2k)J_k(w) = w^{-1} \cdot \sqrt{w^2 - 2xw - 1} \cdot x^k\big|_{-1}^{1} + k(w + w^{-1}) \cdot J_{k-1}(w)$$

B.6 The series for $P_n(\cos\theta)$. With x replaed by $\cos\,\theta$ we notice that

$$1 - 2\cos\,\theta \cdot w + w^2 = (1 - e^{i\theta}w)(1 - e^{-i\theta}w)$$

It follows that

$$\frac{1}{\sqrt{1 - 2\cos(\theta)w + w^2}} = \frac{1}{\sqrt{1 - 1 - e^{i\theta}w)}} \cdot \frac{1}{\sqrt{1 - e^{-i\theta}w)}}$$

The last product becomes

$$\sum\sum g_m e^{im\theta} w^m \cdot g_\nu e^{-i\nu\theta} w^\nu$$

Collecting w powers the double sum becomes

$$\sum \gamma_n(\theta) \cdot w^n \quad \gamma_n(\theta) = \sum_{m+\nu=n} g_m g_\nu e^{i(m-\nu)\theta}$$

By Theorem B.4 the last sum represents $P_n(\cos(\theta))$. One has for example

$$P_3(\cos(\theta) = 2g_3 \cdot \cos(3\theta) + 2g_2 g_1 \cdot \cos(\theta)$$

where we used that $g_0 = 1$.

B.7 An inequality for $|P(x)|$. Since the g-numbers are ≥ 0 we obtain

$$|P_n(\cos(\theta)| \leq g_n g_0 + g_{n-1}g_1 + \ldots + g_1 g_{n-1} + g_0 g_n = P_n(1) \quad : \ 0 \leq \theta \leq 2\pi$$

Hence we have proved

B.8 Theorem. *For each n one has*

$$|P_n(x)| \leq 1 \quad : \quad -1 \leq x \leq 1$$

Next, we study the values when $x > 1$. Here one has

B.9 Theorem. *For each $x > 1$ one has*

$$1 < P_1(x) < P_2(x) < \ldots$$

Proof. Let us put

$$\psi(x.w) = 1 + \sum_{n=1}^{\infty} [P_n(x) - P_{n-1}(x)] \cdot w^n$$

By Theorem B.4 this is equal to

$$(*) \qquad \frac{1-w}{\sqrt{1-2xw+w^2}}$$

With $x > 1$ we set $x = 1 + \xi$ and notice that

$$1 - 2xw + w^2 = (1-w)^2 - 2\xi w$$

Hence (*) becomes

$$(**) \qquad \frac{1}{\sqrt{1 - \frac{2\xi w}{1-w^2}}} = \sum g_n \cdot \frac{(2\xi w)^n}{(1-w^2)^n} = \sum g_n \cdot (2\xi)^n \cdot \frac{w^n}{(1-w^2)^n}$$

Next, for each $n \geq 1$ we notice that the power series of $\frac{w^n}{(1-w^2)^n}$ has positive coefficients. Since $g_n(2\xi)^n > 0$ also hold we conclude that (**) is of the form

$$1 + \sum_{n=1}^{\infty} q_n(\xi) \cdot w^n \quad \text{where} \quad q_n(\xi) > 0$$

Finally, Theorem B.9 follows since

$$P_n(1+\xi) - P_{n-1}(1+\xi) = q_n(\xi)$$

B.10 An L^2-inequality.

Let $n \geq 1$ and denote by $\mathcal{P}_n[1]$ the space of real-valued polynomials $Q(x)$ of degree $\leq n$ for which $\int_{-1}^{1} Q(x)^2 \cdot dx = 1$ and set

$$\rho(n) = \max_{Q \in \mathcal{P}-n[1]} |Q|_\infty$$

where $|Q|_\infty$ is the maximum norm over $[-1,1]$. To find $\rho(n)$ we use the orthonormal basis $\{P_k^*\}$ and write

$$Q(x) = t_0 \cdot P_0^*(x) + \ldots + t_n \cdot P_n^*(x)$$

Since $Q \in \mathcal{P}_n[1]$ we have $t_0^2 + \ldots + t_n^2 = 1$. Recall also that

$$P_\nu^*(x) = \sqrt{\frac{2\nu+1}{2}} \cdot P_\nu(x)$$

Given $-1 \leq x_0 \leq 1$ the Cauchy-Schwarz inequality gives

$$Q(x_0)^2 \leq \sum_{\nu=0}^{\nu=n} \frac{2\nu+1}{2} \cdot |P_\nu(x_0)| \leq \sum_{\nu=0}^{\nu=n} \frac{2\nu+1}{2}$$

where the last inequality follows since the maximum norm of each P_ν is ≤ 1. Finally, we notice that

$$\sum_{\nu=0}^{\nu=n} \frac{2\nu+1}{2} = \frac{(1-n)^2}{2}$$

We conclude that

$$|Q(x_0)| \leq \frac{n+1}{\sqrt{2}}$$

B.11 The case of equality. To have equality above we take $x_0 = 1$ and

$$t_\nu = \alpha \cdot P_\nu^*(1) \quad : \quad \nu \geq 0$$

C. The space $\mathcal{T}_n$

Let $n \geq 1$ be a positive integer. A real-valued trigonometric polynomial of degree $\leq n$ is given by

$$g(\theta) = a_0 + a_1 \cos\theta + \ldots + a_n \cos n\theta + b_1 \sin\theta + \ldots b_n \sin n\theta$$

Here $a_0, \ldots, a_n, b_1, \ldots, b_n$ are real numbers. The space of such functions is denoted by $\mathcal{T}_n$ which is a vector space over $\mathbf{R}$ of dimension $2n+1$. We can write

$$\cos kx = \frac{1}{2}[e^{ikx} + e^{-ikx}] \quad \text{and} \quad \sin kx = \frac{1}{2i}[e^{ikx} - e^{-ikx}] \quad : \ k \geq 1$$

It follows that there exist complex numbers $c_0, \ldots, c_{2n}$ such that

$$g(\theta) = e^{-in\theta} \cdot [c_0 + c_1 e^{i\theta} + \ldots + c_{2n} e^{i2n\theta}]$$

Exercise. Show that

$$c_\nu + c_{2n-\nu} = 2a_\nu \quad \text{and} \quad c_\nu - c_{2n-\nu} = 2b_\nu \implies$$
$$c_{2n-\nu} = \bar{c}_\nu \quad 0 \leq \nu \leq n$$

C.1 The polynomial $G(z)$. Given $g(\theta)$ as above we set

$$G(z) = c_0 + c_1 z + \ldots + c_{2n} z^{2n} \implies e^{-in\theta} \cdot G(e^{i\theta}) = g(\theta)$$

C.2 Exercise. Set

$$\bar{G}(z) = \bar{c}_0 + c_1 z + \ldots + \bar{c}_{2n} z^{2n}$$

and show that

$$(*) \qquad z^{2n} G(1/z) = \bar{G}(z)$$

Use this to show that if $0 \neq z_0$ is a zero of $G(z)$ then $\frac{1}{\bar{z}_0}$ is also a zero of $G(z)$.

C.3 The case when $g \geq 0$. Assume that the g-function is non-negative. Let

$$0 \leq \theta_1 < \ldots < \theta_\mu < 2\pi$$

be the zeros on the half-open interval $[0, 2\pi)$. Since $g \geq 0$ every such zero has a multiplicity given by an *even* integer. Consider also the polynomial $G(z)$. Exercise C.2 shows that $\{e^{i\theta_\nu}\}$ are complex zeros of $G(z)$ whose multiplicities are even integers. Next, if ζ is a zero where $\zeta \neq 0$ and $|\zeta| \neq 1$, then (*) in C.2 implies that $\frac{1}{\bar{\zeta}}$ also is a zero and hence $G(z)$ has a factorisation

$$G(z) = c_{2n} \cdot \prod_{\nu=1}^{\nu=\mu} (z - e^{i\theta_\nu})^{2k_\nu} \cdot \prod_{j=1}^{j=m} (z - \zeta_j)(z - \frac{1}{\bar{\zeta}_j}) \cdot z^{2r} \quad \text{where} \quad 2\mu + 2m + 2r = 2n$$

Here $0 < |\zeta_j| < 1$ hold for each j and it may occur that multiple zeros appear, i.e. the ζ-roots need not be distinct and the integer r may be zero or positive.

C.4 The h-polynomial. Let $\delta = \sqrt{|\zeta_1| \cdots |\zeta_m|}$ and put

$$h(z) = c_{2n} \dot{\delta} \cdot \prod_{\nu=1}^{\nu=\mu} (z - e^{i\theta_\nu})^{k_\nu} \cdot \prod_{j=1}^{j=m} (z - \zeta_j) \cdot z^r$$

C.5 Proposition. *One has the equality*

$$|h(e^{i\theta})|^2 = g(\theta)$$

Proof. With $z = e^{i\theta}$ and $0 < |\zeta| < 1$ one has

$$(e^{i\theta} - \zeta)(e^{i\theta} - \frac{1}{\bar{\zeta}}) = (e^{i\theta} - \zeta) \cdot (\bar{\zeta} - e^{-i\theta}) \cdot e^{i\theta} \cdot \frac{1}{\bar{\zeta}}$$

Passing to absolute values it follows that

$$\left|(e^{i\theta} - \zeta)(e^{i\theta} - \frac{1}{\bar{\zeta}})\right| = \left|e^{i\theta} - \zeta\right|^2 \cdot \frac{1}{|\zeta|}$$

Apply this to every root ζ_ν and take the product which gives Proposition C.5.

C.6 Application. Let $g \geq 0$ be as above and assume that the constant coefficient $a_0 = 1$. This means that

$$1 = \frac{1}{2\pi} \cdot \int_0^{2\pi} g(\theta) \cdot d\theta$$

With $h(z) = d_0 + d_1 z + \ldots + d_n z^n$ we get

$$1 = \frac{1}{2\pi} \cdot \int_0^{2\pi} h(e^{i\theta})|^2 \cdot d\theta = |d_0|^2 + \ldots + |d_n|^2$$

Notice that

(i) $$|d_n|^2 = |c_{2n}| \cdot \delta \quad \text{and} \quad |d_0|^2 = |c_{2n} \cdot \delta| \cdot \prod |\zeta_j|^2 = |c_{2n}| \cdot \frac{1}{\delta}$$

From this we see that

(iii) $$|c_{2n}| \cdot (\delta + \frac{1}{\delta}) = |d_0|^2 + d_n|^2 \leq 1$$

Here $0 < \delta < 1$ and therefore $\delta + \frac{1}{\delta} \geq 2$ which together with (iii) gives

$$|c_{2n}| \leq \frac{1}{2}$$

At the same time we recall that

(*) $$c_{2n} = \frac{a_n + ib_n}{2} \implies |a_n + ib_n| \leq 1$$

Summing up we have proved the following:

C.7 Theorem. *Let $g(\theta)$ be non-negative in $\mathcal{T}_n$ with constant term $a_0 = 1$. Then*

$$|a_n + ib_n| \leq 1$$

C.8 An application. Let $n \geq 1$ and consider the space of all monic polynomials

$$P(x) = x^n + c_{n-1}x^{n-1} + \ldots + c_0$$

where $\{c_\nu\}$ are real- To each such polynomial we can consider the maximum norm over the interval $[-1, 1]$. Then one has

C.9 Theorem. *For each $P \in \mathcal{P}_n^*$ one has the inequality*

$$\max_{-1 \leq x \leq 1} |P(x)| \geq 2^{-n+1}$$

Proof. Consider some $P \in \mathcal{P}_n^*$ and define the trigonometric polynomial

$$g(\theta) = (\cos\theta)^n + c_{n-1}(\cos\theta)^{n-1} + \ldots + c_0$$

So here $P(\cos\theta) = g(\theta)$ and Theorem C.9 follows if we have proved that

(1) $$2^{-n+1} \geq \max_{0 \leq \theta \leq 2\pi} |g(\theta))|$$

To prove this we set $M = \max_{0 \leq \theta \leq 2\pi} |g(\theta))|$. Next, we can write

$$g(\theta) = a_0 + a_1 \cos\theta \ldots + a_n \cos n\theta$$

Moreover, since

$$(\cos\theta)^n = 2^{-n} \cdot [e^{i\theta} + e^{-\theta}]^n$$

we get

$$a_n = 2^n$$

Now we shall apply Theorem C.8. For this purpose we construct non-negative trigonometric polynomials. First we define

$$g^*(\theta) = \frac{M - g(\theta)}{M - a_0}$$

Then $g^* \geq 0$ and its constant term is 1. We have also

$$g^*(\theta) = 1 - \frac{1}{M - a_0} \cdot \sum_{\nu=1}^{\nu=n} a_\nu \cos\nu\theta$$

Hence Theorem C.7 gives

(1) $$\frac{1}{|M-a_0|}\cdot|a_n|\leq 1 \implies |M-a_0|\geq 2^{-n+1}$$

Next, we have also the function

$$g_*(\theta)=\frac{M+g(\theta)}{M+a_0}$$

In the same way as above we obtain:

(2) $$|M+a_0|\geq 2^{-n+1}$$

Finally, (1) and (2) give

$$M\geq 2^{-n+1}$$

which proves Theorem C.9

D. Tchebysheff polynomials.

The inequality in Theorem C.9 is sharp. To see this we shall construct a special polynomial $T_n(x)$ of degree n. Namely, with $n \geq 1$ we can write

$$\cos n\theta = 2^{n-1} \cdot (\cos\theta)^n + c_{n-1} \cdot (\cos\theta)^{n-1} + \ldots + c_0$$

Set

$$T_n(x) = 2^{n-1} x^n + c_{n-1} \cdot x^{n-1} + \ldots + c_0$$

Hence

$$T_n(\cos\theta) = \cos n\theta$$

We conclude that the polynomial

$$p_n(x) = 2^{-n+1} \cdot T_n(x)$$

belongs to $\mathcal{P}_n^*$ and its maximum norm is 2^{-n+1} which proves that the inequality in Theorem C.9 is sharp.

D.1 Zeros of T_n. Set

$$\theta_\nu = \frac{\nu\pi}{n} + \frac{\pi}{2n}$$

It is clear that $\theta_1, \ldots, \theta_n$ are zeros of $\cos n\theta$. Hence the zeros of $T_n(x)$ are:

$$x_\nu = \cos\theta_\nu$$

Notice that

$$-1 < x_n < \ldots < x_1 < 1$$

Since $T_n(x)$ is a polynomial of degree n it follows that $\{x_\nu\}$ give all zeros and we have

$$T_n(x) = 2^{n-1} \cdot \prod (x - x_\nu)$$

D.2 Exercise. Show that

$$T_n'(x_\nu) \cdot \sqrt{1 - x_\nu^2} = n$$

hold for every zero of $T_n(x)$.

D.3 An interpolation formula. Since $x_1, \ldots, x_n$ are distinct it follows that if $p(x) \in \mathcal{P}_{n-1}$ is a polynomial of degree $\leq n-1$ then

$$p(x) = \sum_{\nu 01}^{\nu=n} p(x_\nu) \cdot \frac{1}{T'(x_\nu)} \cdot \frac{T(x)}{x - x_\nu}$$

By the exercise above we get

$$p(x) = \frac{1}{n} \cdot \sum_{\nu=1}^{\nu=n} (-1)^{\nu-1} p(x_\nu) \cdot \sqrt{1 - x_\nu^2} \cdot \frac{T(x)}{x - x_\nu}$$

We shall use the interpolation formula above to prove

D.4 Theorem *Let $p(x) \in \mathcal{P}_{n-1}$ satisfy*

(1) $$\max_{-1 \leq x \leq 1} \sqrt{1 - x^2} \cdot |p(x)| \leq 1$$

Then it follows that

(2) $$\max_{-1 \leq x \leq 1} |p(x)| \leq n$$

Proof. First, consider the case when

(*) $$-\cos\frac{\pi}{2n} \leq x \leq \cos\frac{\pi}{2n}$$

Then we have

$$\sqrt{1 - x^2} \geq \sqrt{1 - \cos^2\frac{\pi}{2n}} = \sin\frac{\pi}{2n}$$

Next, recall the inequality $\sin x \geq \frac{2}{\pi} \cdot x$. It follows that

$$\sqrt{1-x^2} \geq \frac{1}{n}$$

So when (1) holds in the theorem we have

$$|p(x)| = \frac{1}{\sqrt{1-x^2}} \cdot \sqrt{1-x^2} \cdot |p(x)| \leq \frac{1}{\sqrt{1-x^2}} \leq n$$

Hence the required inequality in Theorem D.4 holds when x satisfies (*) above. Next, suppose that

(**) $$x_1 \leq x \leq 1$$

On this interval $T_n(x) \geq 0$ and from the interpolation formula xx and the triangle inequality we have

$$|p(x) \leq \frac{1}{n} \sum_{\nu=1}^{\nu=n} \sqrt{1-x_\nu^2} \cdot |p(x_\nu)| \cdot \frac{T(x)}{x-x_\nu} \leq \frac{1}{n} \sum_{\nu=1}^{nu=n} \frac{T(x)}{x-x_\nu}$$

Next, the sum

$$\frac{T(x)}{x-x_\nu} = T_n'(x) = n \cdot U_{n-1}(x)$$

So when (**) holds we have

(***) $$|p(x)| \leq |U_{n-1}(x)|$$

By xx the maximum norm of U_{n-1} over $[-1,1]$ is n and hence (***) gives

$$|p(x)| \leq n$$

In the same way one proves htat

$$-1 \leq x \leq x_n \implies |p(x)| \leq n$$

Together with the upper bound in the case (xx) we get Theorem D.4.

D.5 Bernstein's inequality.

Let $g(\theta) \in \mathcal{T}_n$. The derivative $g'(\theta)$ is another trigonometric polynomial and we have

Theorem. *For each $g \in \mathcal{T}_n$ one has*

$$\max_{0 \leq \theta \leq 2\pi} |g'(\theta)| \leq n \cdot \max_{0 \leq \theta \leq 2\pi} |g(\theta)|$$

Before we prove this result we establish an inequality for certain trigonometric polynomials. Namely, consider a real-valued sine-polynomial

$$S(\theta) = c_1 \text{sin}(\theta) + \ldots + c_n \text{sin}(n\theta)$$

Now $\theta \mapsto \frac{S(\theta)}{\sin\theta}$ is an even function of θ and therefore one has

$$\frac{S(\theta)}{\sin\theta} = a_0 + a_1 \cos\theta + \ldots + a_{n-1} (\cos\theta)^{-n-1}$$

Consider the polynomial

$$p(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1}$$

Then e see that:

$$|p(\cos\theta)| = \frac{|S(\theta)|}{\sqrt{1-\cos^2\theta}}$$

Using this we apply Theorem D.4 to the polynomial $p(x)$ and conclude

D.6 Theorem. *Let $S(\theta) = c_1 sin(\theta) + c_n sin(n\theta)$ be a sine-polynomial as above. Then*

$$\max_{0 \leq \theta \leq 2\pi} \frac{|S(\theta)|}{\sin\theta} \leq n \cdot \max_{0 \leq \theta \leq 2\pi} |S(\theta)|$$

D.7 Proof of Bernstein's theorem. Fix an arbitrary $0 \leq \theta - 0 < 2\pi$. Set

$$S(\theta) = g(\theta_0 + \theta) - g(\theta_0 - \theta)$$

We notice that $S(\theta$is an odd polynomial of θ and $S(0) = 0$, It follows that $S(\theta)$ is a sine-polynomial as above of degree $\leq n$. Notice also that

$$\max_{0 \leq \theta \leq 2\pi} |S(\theta)| \leq 2 \cdot \max_{0 \leq \theta \leq 2\pi} |g(\theta)| \max_{0 \leq \theta \leq 2\pi} |g(\theta)|$$

Theorem D.6 applied to $S(\theta)$ gives

$$\text{(i)} \qquad |\frac{g(\theta_0 + \theta) - g(\theta_0 - \theta)}{\sin\theta}| \leq 2n \cdot \max_{0 \leq \theta \leq 2\pi} |g(\theta)|$$

Next, in the left hand side we can take the limit as $\theta| \to 0$ and notice that

$$2 \cdot g'(\theta_0) = \lim_{\theta \to 0} \frac{g(\theta_0 + \theta) - g(\theta_0 - \theta)}{\sin\theta}$$

Hence (i) gives

$$|g'(\theta_0)| \leq n \cdot \max_{0 \leq \theta \leq 2\pi} |g(\theta)|$$

Finally, since θ_0 was arbitrary we get Bernstein's theorem.

E. Fejers sine series and Gibb's phenomenon.

Several remarkable inequalities for trigonometric polynomials were established by Fejer in [Fejer] where a central issue is to find trigonometric polynomials expressed by a sine series which are ≥ 0 on the interval $[0, \pi]$. Consider as an example is the sine-series

$$S_n(\theta) = \sum_{k=1}^{k=n} \frac{\sin k\theta}{k}$$

E.1 Theorem. *For every $n \geq 1$ one has the inequality*

$$0 < S_n(\theta) \leq 1 + \frac{\pi}{2} \quad : \quad 0 < \theta < \pi$$

The upper bound was proved by in [Fej] and Fejer conjectured that $S_n(\theta)$ stays positive on $(0, \pi)$. This was confirmed in articles by Jackson in [xx] and Cronwall in [xx]. The series (*) has a connection with Gibb's phenomenon and Theorem E.1 can be illustrated by drawing graphs of the S-polynomials where the situation when $\theta = \pi - \delta$ for small positive δ has special interest. Since cos $\pi = -1$ the positivity entails that

$$(*) \qquad \sum_{k=2}^{n} (-1)^k \cdot \frac{\sin k\delta}{k} \geq \sin\delta \quad \text{hold for every} \quad n \geq 2 \quad \text{and small} \quad \delta > 0$$

Exercise. Prove Theorem E.1 or consult the literature. It is also instructive to confirm (*) by numerical experiments with a computer.

E.2 Mehler's integral formula. In XX we introduced the Legendre polynomials. It turns out that

$$(*) \qquad \mathcal{P}_n(x) = \sum_{\nu=1}^{\nu=n} P_\nu(x) > 0 \quad : \quad -1 < x < 1$$

is strictly positive for each $-1 < x < 1$.

Exercise. Prove (*) using Theorem E.1 and Mehler's integral formula

$$(*) \qquad \mathcal{P}_n(\cos\theta) = \frac{2}{\pi} \cdot \int_0^\pi \frac{\sin(n + \frac{1}{2})\phi \cdot d\phi}{\sqrt{2\cos\theta - 2\cos\phi}}$$

F. Convergence of arithmetical means

Let $f(x)$ be a real-valued and square integrable function on $(-\pi, \pi)$, i.e.

$$\int_{-\pi}^{\pi} |f(x)|^2 \, dx < \infty$$

We say that f has a determined value $A = f(0)$ at $x = 0$ if the following two conditions hold:

$$\text{(i)} \qquad \lim_{\delta \to 0} \frac{1}{\delta} \cdot \int_0^{\delta} |f(x) + f(-x) - 2A| \, dx = 0$$

$$\text{(ii)} \qquad \int_0^{\delta} |f(x) + f(-x) - 2A|^2 \, dx \leq C \cdot \delta \quad \text{holds for some constant} \quad C$$

Remark. In the same way we can impose this condition at every point $-\pi < x_0 < \pi$. To simplify the subsequent notations we take $x = 0$. If $x = 0$ is a Lebesgue point for f and A the Lebesgue value we have (i). Hence Lebesgue's Theorem entails that (i) holds almost everywhere when $x = 0$ is replaced by other points x_0. We leave it to the reader to show that the second condition also is valid almost everywhere when f is square integrable but in general there appears a null set $\mathcal{N}$ where (ii) fails to hold while $\mathcal{N}$ contains some Lebegue points. Next, expand f in a Fourier series

$$f(x) = \frac{a_0}{2} + \sum a_k \cdot \cos kx + \sum b_k \cdot \sin kx$$

and with $x = 0$ we consider the partial sums

$$s_n(0) = \frac{a_0}{2} + a_1 + \ldots + a_n + b_1 + \ldots + b_n$$

The result below is proved in [Carleman] and shows that $\{s_n\}$ are close to the determined value for many n-values.

F.1 Theorem. *Assume that f has a determined value A at $x = 0$. Then the following hold for every positive integer k*

$$(*) \qquad \lim_{n \to \infty} \frac{1}{n+1} \cdot \sum_{\nu=0}^{\nu=n} |s_\nu - A|^k = 0$$

Remark. Carleson's theorem asserts that $\{s_n(x)\}$ converge to $f(x)$ almost everywhere when $f \in L^2$. When a pointwise convergence holds the limit formula (*) is obvious. However, it is in general not true that the pointwise convergence exists at *every point* where f has a determined value. So "ugly points" may appear in a null-set where pointwise convergence fails and here Carleman's result offers a substitute.

The case when $f \in$ BMO(T). If f has bounded mean oscillation the results from § XX in Special Topics show that the conditions (i-ii) hold at every Lebesgue point of f. So here one has a control for averaged Fourier series of f expressed via its set of Lebesgue points.

The case when f is continuous. Here (i-ii) hold everywhere so the averaged limit formulas hold at every point. We can say more since f is uniformly continuous. Let $\omega_f(\delta)$ be the modulos of continuity function and for each $n \geq 1$, $||s_n - f||$ is the maximum norm of $s_n - f$ over $[0, 2\pi]$. Set

$$\mathcal{D}_n(f) = \sqrt{\frac{1}{n+1} \cdot \sum_{\nu=0}^{\nu=n} ||s_\nu - f||^2}$$

F.2 Theorem. *There exists an absolute constant K such that the following hold for every continuous function f with maximum norm ≤ 1:*

$$\mathcal{D}_n(f) \leq K \cdot \left[\frac{1}{\sqrt{n}} + \omega_f\left(\frac{1}{n}\right)\right]$$

Proof of Theorem F.1

Set $A = f(0)$ and $s_n = s_n(0)$. Introduce the function:

$$\phi(x) = f(x) + f(-x) - 2A$$

Applying Dini's kernel we have

$$s_n - A = \int_0^\pi \frac{\sin(n+1/2)x}{\sin x/2} \cdot \phi(x) \cdot dx$$

By trigonometric formulas the integral is expressed by three terms for each $0 < \delta < \pi$:

$$\alpha_n = \frac{1}{\pi} \cdot \int_0^\delta \sin nx \cdot \cot x/2 \cdot \phi(x) \cdot dx$$

$$\beta_n = \frac{1}{\pi} \cdot \int_\delta^\pi \sin nx \cdot \cot x/2 \cdot \phi(x) \cdot dx$$

$$\gamma_n = \frac{1}{\pi} \cdot \int_0^\pi \cos nx \cdot \phi(x) \cdot dx$$

By Hölder's inequality it suffices to show Theorem F.1 if $k = 2p$ is an even integer. Minkowski's inequality gives

$$(1) \qquad \Big[\sum_{\nu=0}^{\nu=n} |s_\nu - A|^{2p}\Big]^{1/2p} \leq \Big[\sum_{\nu=0}^{\nu=n} |\alpha_\nu|^{2p}\Big]^{1/2p} + \Big[\sum_{\nu=0}^{\nu=n} |\beta_\nu|^{2p}\Big]^{1/2p} + \Big[\sum_{\nu=0}^{\nu=n} |\gamma_\nu|^{2p}\Big]^{1/2p}$$

Denote by $o(\delta)$ small ordo and $O(\delta)$ is big ordo. When $\delta \to 0$ we shall establish the following:

$$\text{(i)} \qquad \Big[\sum_{\nu=0}^{\nu=n} |\alpha_\nu|^{2p}\Big]^{1/2p} = n^{1+1/2p)} \cdot o(\delta)$$

$$\text{(ii)} \qquad \Big[\sum_{\nu=0}^{\nu=n} |\beta_\nu|^{2p}\Big]^{1/2p} \leq K \cdot p \cdot \delta^{-1/2p}$$

$$\text{(iii)} \qquad \Big[\sum_{\nu=0}^{\nu=n} |\gamma_\nu|^{2p}\Big]^{1/2p} \leq K$$

In (ii-iii) K is an absolute constant which is independent of p, n and δ. Let us first see why (i-iii) give Theorem F.1. Write $o(\delta) = \epsilon(\delta) \cdot \delta$ where $\epsilon(\delta) \to 0$. With these notations (1) gives:

$$(*) \qquad \Big[\sum_{\nu=0}^{\nu=n} |s_\nu - A|^{2p}\Big]^{1/2p} \leq n^{1+1/2p} \cdot \delta \cdot \epsilon(\delta) + Kp \cdot \delta^{-1/2p} + K$$

Next, let $\rho > 0$ and choose b so large that

$$pKb^{-1/2p} < \rho/3$$

Take $\delta = b/n$ and with n large it follows that $\epsilon(\delta)$ is so small that

$$b \cdot \epsilon(b/n) < \rho/3$$

Then right hand side in (*) is majorized by

$$\frac{2\rho}{3} \cdot n^{1/2p} + K$$

When n is large we also have

$$K \leq \frac{\rho}{3} \cdot n^{1/2p}$$

Hence the left hand side in (*) is majorized by $\rho \cdot n^{1/2p}$ for all sufficiently large n. Since $\rho > 0$ was arbitrary we get Theorem F.1 when the power is raised by $2p$.

Proof of (i-iii)

To obtain (i) we use the triangle inequality which gives the following for every integer $\nu \geq 1$:

$$(1) \qquad |a_\nu| \leq \frac{2}{\pi} \cdot \max_{0 \leq x \leq \delta} |\sin \nu x \cdot \cot x/2| \cdot \int_0^\delta |\phi(x)|\, dx = \nu \cdot o(\delta)$$

where the small ordo δ-term comes from the hypothesis expressed by (*) in the introduction. Hence the left hand side in (i) is majorized by

$$[\sum_{\nu=1}^{\nu=n} \nu^{2p}\,]^{\frac{1}{2p}} \cdot o(\delta) = n^{1+1/2p} \cdot o(\delta)$$

which was requested to get (i). To prove (iii) we notice that

$$\gamma_0^2 + 2 \cdot \sum_{\nu=1}^\infty \gamma_\nu^2 = \frac{1}{\pi} \int_0^\pi |\phi(x)|^2\, dx$$

Next, we have

$$\sum_{\nu=1}^\infty |\gamma_\nu|^{2p} \leq [\sum_{\nu=1}^\infty |\gamma_\nu|^2\,]^{1/2p} \leq K$$

where K exists since ϕ is square-intergable on $[0, \pi]$.

Proof of (ii). Here several steps are required. For each $0 < s < \pi$ we define the function $\phi_s(x)$ by

$$\phi_s(x) = \phi(x) \quad : \quad 0 < x < s$$

and extend it to an odd function, i.e. $\phi_s(-x) = -\phi_s(x)$ while $\phi_s(x) = 0$ when $|x| > s$. This odd function has a sine series

$$(1) \qquad \phi_s(x) = \sum_{\nu=1}^\infty c_\nu(s) \cdot \sin x$$

Let us also introduce the functions

$$(2) \qquad \rho(s) = \int_0^s |\phi(x)| \cdot dx \quad \text{and} \quad \Theta(s) = \int_0^s |\phi(x)|^2 \cdot dx$$

The crucial step towards the proof of (ii) is the following:

Lemma. *One has the inequality*

$$\sum_{|nu=1}^\infty |c_\nu(s)|^{2p} \leq (\frac{2}{\pi})^{2p-1} \cdot \Theta(s) \cdot \rho(s)^{2p-2}$$

Proof. We employ convolutions and define inductively a sequence of functions $\{\phi_{n,s}(x)\}$ where $\phi_{1,s}(x) = \phi_s(x)$ and

$$\phi_{n+1,s}(x) = \frac{1}{\pi} \int_{-\pi}^\pi \phi_{n,s}(y)\phi_s(x+y) \cdot dy$$

Since convolution yield products of the Fourier coefficients and $2p$ is an even integer we have the standard formula:

$$(1) \qquad \sum_{\nu=1}^\infty c_n(s)^{2p} = \phi_{2p,s}(0)$$

Next, using the Cauchy-Schwarz inequality the reader may verify that

$$|\phi_{2,s}(x)| \leq \frac{2}{\pi} \cdot \Theta(x)$$

This entails that

$$\phi_{3,s}(x) \leq \frac{1}{\pi} \int_{-\pi}^{\pi} |\phi_{2,s}(y)| \cdot |\phi_s(x+y)| \cdot dy \leq \frac{2}{\pi^2} \cdot \Theta(s) \cdot \int_{-\pi}^{\pi} |\phi_s(x+y)| \cdot dy = (\frac{2}{\pi})^2 \cdot \Theta(s) \cdot \rho(s)$$

Proceeding in this way it follows by an induction that

$$\phi_{2p,s}(x) \leq (\frac{2}{\pi})^{2p-1} \cdot \Theta(s) \cdot (\rho(s))^{2p-2}$$

This holds in particular when $x = 0$ and then (1) above gives Lemma 1.

A formula for the β-numbers. We have by definition

$$\beta_\nu = \frac{2}{\pi} \int_\delta^\pi \sin \nu x \cdot \frac{1}{2} \cot(\frac{x}{2}) \cdot \phi(x) \cdot dx$$

An integration by parts and the construction of the Fourier coefficients $\{c_\nu(s)\}$ which applies with $s = \delta$ give:

$$(*) \qquad \beta_\nu = -\frac{1}{2} \cdot \cot \delta/2 \cdot c_\nu(\delta) + +\frac{1}{4} \int_\delta^\pi c_\nu(x) \cdot \mathrm{cosec}^2(\frac{x}{2}) \cdot dx$$

Now we profit upon Minkowski's inequality. Let q be the conjugate of $2p$, i.e $\frac{1}{q} + \frac{1}{2p} = 1$ and choose $\{\xi_\nu\}$ to be the sequence in ℓ^q of unit norm such that

$$|\sum \xi_\nu \cdot \beta_\nu| = ||\beta_\bullet||_{2p}$$

where the last term is the left hand side in (ii). At the same time (*) above and the triangle inequality give

$$||\beta_\bullet||_{2p} \leq -\frac{1}{2} \cdot \cot(\delta/2) \cdot \sum |c_\nu(\delta)| \cdot |\xi_\nu| + \frac{1}{4} \int_\delta^\pi \mathrm{cosec}^2(\frac{x}{2}) \cdot \sum |c_\nu(x) \cdot \xi_\nu| \cdot dx \leq$$

$$(**) \qquad \frac{1}{2} \cdot \cot(\delta/2) \cdot ||c_\bullet(\delta)||_{2p} + \frac{1}{4} \int_\delta^\pi \mathrm{cosec}^2(\frac{x}{2}) \cdot ||c_\bullet(x)||_{2p} \cdot dx$$

At this stage we apply Lemma 1 and the assumption which give a constant K such that

$$\Theta(s) \leq K \quad \text{and} \quad \rho(s) \leq K \cdot s$$

The last estimate actually is weaker than the hypothesis but it will be sufficient to get the requested estimate of the ℓ^{2p}-norm in (ii). Lemma 1 gives a constant K_1 such that

$$||c_\bullet(\delta)||_{2p} \leq K_1 \cdot \delta^{1-1/p}$$

At the same time we have a constant K_2 such that

$$\cot(\delta/2) \leq \frac{K_2}{\delta}$$

The product in the first term from (**) is therefore majorized by $K_1 K_2 \cdot \delta^{-1/2p}$ as requested in (ii). For the second term we use Lemma 1 which first gives

$$||c_\bullet(x)||_p \leq K \cdot x^{-1/2p}$$

At this stage we leave it to the reader to verify that we get a constant K so that

$$\int_\delta^\pi x^{-1/2p} \cdot \mathrm{cosec}^2(\frac{x}{2}) \cdot dx \leq K \cdot \delta^{-1/2p}$$

which finishes the proof of (ii).

The case when f is continuous.

Under the normalisation that the L^2-integral of f is ≤ 1 the inequalities (ii-iii) hold for an absolute constant K. In (i) we notice that the construction of ϕ and the definition of ω_f give the estimates

$$|a_\nu| \leq \nu \cdot \delta \cdot \omega_f(\delta)$$

With $p = 2$ this entails that (i) from the proof of Theorem F.1 is majorised by

$$n^{1+1/2} \cdot \delta \cdot \omega_f(\delta)$$

This holds for every $0 \leq x \leq 2\pi$ and from the previous proof we conclude that the following hold for each $n \geq 2$ and every $0 < \delta < \pi$:

(i) $$\mathcal{D}_n(f) \leq \frac{1}{\sqrt{n+1}} \cdot [n^{1+1/2} \cdot \delta \cdot \omega_f(\delta) + 2K\delta^{-1/2} + K]$$

With $n \geq 2$ we take $\delta = n^{-1}$ and see that (i) gives a requested constnt in Theorem F.2.

G. Best approximation by trigonometric polynomials.

The results below are due to de Vallé Poussin and we follow Chapter VIII in his text-book [V-P]. Consider the $2n+2$-tuple

$$x_j = \frac{2\pi j}{(2n+2)} \quad : \quad 1 \leq j \leq 2n+2$$

Let $P(x)$ be a trigonometric polynomial in $\mathcal{T}_n$:

$$P(x) = \sum_{\nu=-n}^{\nu=n} a_\nu \cdot e^{i\nu x}$$

Let f be a 2π-periodic and continuous function and put

$$\rho_P(f) = \max_j |P(x_j) - f(x_j)|$$

Assume that $\rho_P(f) > 0$ which gives a unique $(2n+2)$-tuple of complex numbers $\{u_j\}$ where every u_j has absolute value ≤ 1 and

$$(1) \qquad f(x_j) = \rho_P(f) \cdot u_j + \sum_{\nu=-n}^{\nu=n} a_\nu \cdot e^{i\nu x_j} \quad : \quad 1 \leq j \leq 2n+2$$

G.1 Proposition. *One has the equality*

$$(*) \qquad \rho_P(f) = |\frac{f(x_1) - f(x_2) + \ldots + f(x_{2n+1} - f(x_{2n})}{u_1 + u_2 + \ldots + u_{2n+1} + u_{2n+2}}|$$

Proof. Consider the $(2n+2) \times (2n+1)$-matrix

$$(i) \qquad \begin{pmatrix} e^{-inx_1} & \ldots & e^{inx_1} \\ e^{-inx_2} & \ldots & e^{inx_2} \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ e^{-inx_{2n+2}} & \ldots & e^{inx_{2n+2}} \end{pmatrix}$$

To each $1 \leq k \leq 2n+2$ we denote by $\mathcal{A}_k$ the $(2n+1) \times (2n+1)$-matrix which arises when the k:th row is deleted. Using van der Monde formulas the reader can verify that

$$(ii) \qquad A_k = \det(\mathcal{A}_k) = \prod_{1 \leq i \leq j \leq 2n+2}^{(k)} \sin \frac{x_j - x_i}{2}$$

where (k) above the product sign indicates that i and j both are $\neq 0$ in the product. We leave it to the reader to show that there exists a positive constant A_* such that

$$(iii) \qquad \det A_k = A_* \quad : \quad 1 \leq k \leq 2n+2$$

Now (1) is a system of linear equations where $\rho_*(P), a_{-n}, \ldots, a_n$ are the indeterminate variables. By Cramér's rule we can solve out $\rho_*(P)$ via the $2n+2$-matrix and (iii) gives

$$(iv) \qquad \rho_P(f) = \frac{A_1 f(x_1) - A_2 f(x_2) + \ldots + A_{2n+1} f(x_{2n+1} - A_{2n} f(x_{2n})}{A_1 u_1 + A_2 u_2 + \ldots + A_{2n+1} u_{2n+1} + A_{2n} u_{2n}}$$

Together with (iii) the requested equation (*) in Proposition G.1 follows.

G.2 Conclusion. To find

$$\rho_n(f) = \min_{P \in \mathcal{T}_n} \rho_P(f)$$

we should choose P so that all the u-numbers are +1 or -1. This determines the a-numbers in the system (1), i.e. we find a unique polynomial $P_* \in \mathcal{T}_n$ for which

$$\rho_n(f) = \max_{1 \leq j \leq 2n+2} |P_*(x_j) - f(x_j]|$$

Moreover, the deviation numbers $|P_*(x_j) - f(x_j)|$ are all equal to

(**) $$\frac{\big|f(x_1) - f(x_2) + \ldots + f(x_{2n+1} - f(x_{2n})\big|}{2n+2}$$

G.3 Exercise. Let f be given with a Fourier series expansion:

$$f(x) = \frac{1}{2} \cdot a_0 + \cdot \sum_{k=1}^{\infty} (\alpha_k \cdot \cos\, kx + \beta_k \cdot \sin\, kx)$$

Use the formula (ii) from the proof above to show that

(***) $$\frac{1}{2n+2} \cdot \sum_{\nu=1}^{\nu=2n+2} (-1)^{\nu} \cdot f(x_{\nu}) = \sum_{j=0}^{\infty} a_{(2j+1)(n+1)}$$

Together (**) and (***) give:

G.4 Theorem. *For each integer n and every 2π-periodic function $f(\theta)$ one has the equality*

$$\rho_n(f) = \big|a_{n+1} + a_{3(n+1)} + a_{5(n+1)} + \ldots\big|$$

G.5 Remark. Since the maximum norm taken over the whole interval $[0, 2\pi]$ majorizes the maximum norm over the $2n+2$-tuple above, we get the inequality:

$$\min_{P \in \mathcal{T}_n} ||f - P|| \geq \big|a_{n+1} + a_{3(n+1)} + a_{5(n+1)} + \ldots\big|$$

where $||f - P||$ is the maximum norm over $[0, 2\pi]$. This gives the result announced in § 0.X from the introduction.

SKIP THIS

Zeros of polynomials. Considerable attention is given to the location of zeros of polynomials of a complex variable z. Consider as an exampe a monic polynomial of even degree $2m$:

$$P(z) = z^{2m} + c_{2m-1}z^{2m-1} + \ldots + c_1 z + c_0$$

Separating real and imaginary parts we have $c_k = a_k + ib_k$ and get the polynomial

$$R(z) = z^{2m} + a_{2m-1}z^{2m-1} + \ldots + a_1 z + a_0$$

Suppose that $R(z)$ has some real zeros with odd multiplicity, i.e. if a is such a real zero then the signs of $P(a-\epsilon)$ and $P(a+\epsilon)$ differ for small ϵ. Let $\alpha_1 < \ldots < \alpha_k$ be this set of real zeros. Eventual real zeros of where P vanishes with an even order are not included. We have also the imaginary part of P expressed by the polynomial

$$S(z) = b_{2m-1}x^{2m-1} + \ldots + b_1 z + b_0$$

Under the hypothesis that $S(\alpha_\nu) \neq 0$ for each $1 \leq \nu \leq k$ the calculus with complex numbers gives a formula for the number of zeros of $P(z)$ counted with multiplicities in the upper half-plane $\mathfrak{Im} z > 0$. The result is that this number is equal to

$$m + \frac{1}{2} \cdot \sum_{\nu=1}^{\nu=k} (-1)^{\nu-1} \cdot \operatorname{sign}(S(\alpha_\nu)) \tag{*}$$

After the reader has become familiar with the argument principle and Cauchy's integral formula the proof of (*) becomes an easy exercise. See also § xx for details of the proof. Various special cases of the result above were known at an early stage by Gauss and Cauchy while the general result is attributed to Laguerre. A special case in questions related to stability for solutions to ordinary differential equations is the condition that all zeros of $P(z)$ belong to the open upper half-plane. From (*) a necessary and sufficient condition for this to be true is that $R(x)$ has $2n$ many simple real zeros $\alpha_1 < \ldots < \alpha_{2m}$ and at the same time $S(z)$ has degree $2n-1$ and interlacing real zeros $\beta_1 < \ldots < \beta_{2n-1}$, i.e. it holds that

$$\alpha_1 < \beta_1 < \alpha_2 < \ldots < \alpha_{2m-1} < \beta_{2m-1} < \alpha_{2m}$$

In addition the leading coefficient b_{2m-1} of the S-polynomial must be strictly negative.

A: The kernels of Dini, Fejer and Jackson

Denote by $C^0_{\text{per}}[0,2\pi]$ the family of complex-valued continuous functions $f(\theta)$ on $[0,2\pi]$ which satisfy $f(0)=f(2\pi)$. The Fourier coefficients of such a function f are defined by:

$$\widehat{f}(n)=\frac{1}{2\pi}\cdot\int_0^{2\pi} e^{-in\phi}f(\phi)\cdot d\phi$$

where n are integers. Fourier's partial sum of degree N is defined by

$$S_N^f(\theta)=\sum_{n=-N}^{n=N}\hat{f}(n)\cdot e^{in\theta} \tag{A.0}$$

A.1.The Dini kernel. If $N\geq 0$ we set

$$D_N(\theta)=\frac{1}{2\pi}\sum_{n=-N}^{n=N} e^{in\theta}$$

A.2 Exercise. Show that the folloeing hold for each $N\geq 0$:

$$S_N^f(\theta)=\int_0^{2\pi} D_N(\theta-\phi)\cdot f(\phi)\cdot d\phi=\int_0^{2\pi} D_N(\phi)\cdot f(\theta+\phi)\cdot d\phi$$

A.3 Proposition. *One has the formula*

$$D_N(\theta)=\frac{1}{2\pi}\cdot\frac{\sin((N+\frac{1}{2})\theta)}{\sin\frac{\theta}{2}} \tag{A.3.1}$$

Proof. We have

$$\sum_{n=-N}^{n=N} e^{in\theta}=e^{-iN\theta}\cdot\sum_{n=0}^{n=2N} e^{in\theta}=e^{-iN\theta}\cdot\frac{e^{i(2N+1)\theta}-1}{e^{i\theta}-1}=$$

$$e^{-iN\theta-i\theta/2}\cdot\frac{e^{i(2N+1)\theta}-1}{2i\cdot\sin\theta/2}=\frac{2i\cdot\sin((N+1/2)\theta)}{2i\cdot\sin\theta/2}$$

and (A.3.1) follows after division with $2i$.

A.4 The Fejer kernel. For each $N\geq 0$ we set

$$\mathcal{F}_N(\theta)=\frac{D_0(\theta)+\ldots+D_N(\theta)}{N+1}$$

A.5 Proposition *One has the formula*

$$\mathcal{F}_N(\theta)=\frac{1}{2\pi(N+1)}\cdot\frac{1-\cos((N+1)\theta)}{2\cdot\sin^2(\frac{\theta}{2})} \tag{A.5.1}$$

Proof. To each $\nu\geq 0$ we have $\sin((\nu+1/2)\theta)=\mathfrak{Im}\big[e^{i(\nu+1/2)\theta)}\big]$. Hence $F_N(\theta)$ is the imaginary part of

$$\frac{1}{2\pi(N+1)}\cdot\frac{e^{i\theta/2}}{\sin(\theta/2)}\cdot\sum_{\nu=0}^{\nu=N} e^{i\nu\theta}$$

Next, we have

$$e^{i\theta/2}\cdot\sum_{\nu=0}^{\nu=N} e^{i\nu\theta}=e^{i\theta/2}\cdot\frac{e^{i(N+1)\theta}-1}{e^{i\theta-1}}=\frac{e^{i(N+1)\theta}-1}{2i\cdot\sin(\theta/2)}$$

Since $i^2=-1$ we see that the imaginary part of the last term is equal to

$$\frac{1-\cos((N+1)\theta)}{2\cdot\sin(\frac{\theta}{2})}$$

and (A.5.1) follows.

A.6 Fejer sums. For each f and every $N \geq 0$ we set

$$F_N^f(\theta) = \int_0^{2\pi} \mathcal{F}_N(\phi) \cdot f(\theta + \phi) \cdot d\phi$$

A.7 An inequality. If $0 < a > 2\pi$ and $a \leq \theta \leq 2\pi - a$ we have the inequality

(i) $$\sin^2(\theta/2) \geq \sin^2(a/2)$$

Let f be given and denote by $M(f)$ the maximum norm of $|f(\theta)|$ over $[0, 2\pi]$. Then (i) gives

$$\int_a^{2\pi - a} F_N(\phi) \cdot f(\theta + \phi) \cdot d\phi \leq$$

(A.7.1) $$\frac{M}{2\pi(N+1) \cdot \sin^2(a/2)} \int_a^{2\pi - a} (1 - \cos(N\phi)) \cdot d\phi \leq \frac{2M}{(N+1) \cdot \sin^2(a/2)}$$

A.8 Exercise. Given some θ_0 and $0 < a < \pi 0$ we set

$$\omega_f(a) = \max_{|\theta - \theta_0| \leq a} |f(\theta) - f(\theta_0)|$$

Use (A.7.1) to prove that

$$|\mathcal{F}_N(\theta_0) - f(\theta_0)| \leq \frac{2M}{(N+1) \cdot \sin^2(a/2)} + \omega_f(a)$$

Conclude that the *uniform continuity* of the function f on $[0, 2\pi]$ implies that the sequence $\{F_N^f\}$ converges uniformly to f over the interval $[0, 2\pi]$.

A.9 The case when f is real-valued. When f is real-valued the Fourier series takes the form

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cdot \cos kx + \sum_{k=1}^{\infty} b_k \cdot \sin kx$$

Here $a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) \cdot dx$ and when $k \geq 1$ one has

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cdot \cos kx \cdot dx \quad : b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cdot \sin kx \cdot dx$$

Fourier's partial sum functions become

$$S_n(f) = \frac{a_0}{2} + \sum_{k=1}^{k=n} a_k \cdot \cos kx + \sum_{k=1}^{k=n} b_k \cdot \sin kx$$

The Jackson kernel

.

Above we proved that the fejer sums converge unifor myl to f. One may ask if there exists a constant C which is independent of both f and of n such that

$$(*) \qquad \max_\theta ||f(\theta) - F_n^f(\theta)| \leq C \cdot \omega_f(\frac{1}{n})$$

Examples show that (*) does not hold in general. To obtain a uniform constant C one must include an extra factor.

A.10 Exercise. Use (A.7-8) to show that there exists an absolute constant C such that

$$(**) \qquad ||f - \mathcal{F}_n(f)|| \leq C \cdot \omega_f(\frac{1}{n}) \cdot \big(1 + \log^+ \frac{1}{\omega_f(\frac{1}{n})}\big)$$

hold for all continuous 2π-periodic functions f.

To attain (*) D. Jackson introduced a new kernel in his thesis *Über die Genauigkeit der Annährerung stegiger funktionen durch ganze rationala funktionen* from Göttingen in 1911. To each 2π-periodic and continuous function $f(x)$ on the real line and every $n \geq 1$ we set

$$\mathcal{J}_n^f(x) = \frac{3}{2\pi} \cdot \int_{-\infty}^{\infty} f(x + \frac{2t}{n}) \cdot (\frac{\sin t}{t})^4 \cdot dt$$

A.11 Theorem. *The function $\mathcal{J}_n^f(x)$ is a trigonometric polynomial of degree $2n-1$ at most and one has the inequality*

$$\max_x |f(x) - \mathcal{J}_n^f(x)| \leq (1 + \frac{6}{\pi}) \cdot \omega_f(\frac{1}{n})$$

Proof. The variable substitution $t \to nt$ gives

$$(1) \qquad \mathcal{J}_n^f(x) = \frac{3}{2\pi n^3} \cdot \int_{-\infty}^{\infty} f(x + 2t) \cdot (\frac{\sin nt}{t})^4 \cdot dt$$

Since $t \mapsto f(x+2t) \cdot sin^4\, nt$ is π-periodic it follows that (1) is equal to

$$(2) \qquad \frac{3}{2\pi n^3} \cdot \int_0^\pi f(x + 2t) \cdot \sum_{k=-\infty}^{\infty} \frac{\sin^4(nt)}{(k\pi + t)^4} \cdot dt$$

Next, recall from § XX that

$$\frac{1}{\sin^2 z} = \sum_{k=-\infty}^{\infty} \frac{1}{(z + k\pi)^2}$$

Taking a second derivative when $z = t$ is real it follows that

$$(3) \qquad \partial_t^2(\frac{1}{\sin^2 t}) = \frac{1}{6} \cdot \sum_{k=-\infty}^{\infty} \frac{1}{(t + k\pi)^4}$$

Hence we obtain

$$(*) \qquad \mathcal{J}_n^f(x) = \frac{1}{4\pi n^3} \cdot \int_0^\pi f(x + 2t) \cdot \sin^4(nt) \cdot \partial_t^2(\frac{1}{\sin^2 t})\, dt$$

Next, the function

$$\sin^4(nz) \cdot \partial_z^2(\frac{1}{\sin^2 z})$$

is entire and even and the reader may verify that it is a finite sum of entire cosine-functions which implies that the Jackson kernel is expressed by a finite sum of integrals:

$$(4) \qquad \mathcal{J}_f^n(x) = \sum_{k=0}^{2n-1} c_k \int_0^{2\pi} f(u) \cdot \cos\, k(x - u))\, du$$

In particular $\mathcal{J}_f^n(x)$ is a trigonometric polynomial of degree $2n-1$ a most. Integration by parts give the equality

$$(5)\qquad \int_{-\infty}^{\infty}(\frac{\sin nt}{t})^4\,dt = \frac{1}{6}\int_0^{\pi}\sin^4 t\cdot\partial_t^2(\frac{1}{\sin^2 t})\,dt = \frac{4}{3}\int_0^{\pi}\cos^2 t\,dt = \frac{2\pi}{3}$$

Next, we leave it to the reader to verify the inequality

$$(6)\qquad \frac{3}{2\pi}\int_{-\infty}^{\infty}(1+2|t|)\cdot(\frac{\sin t}{t})^4\cdot dt \leq 1+\frac{6}{\pi}$$

From the above where we use (1) and (*) it follows that

$$(7)\qquad \mathcal{J}_n^f(x) - f(x) = \frac{3}{2\pi}\cdot\int_{-\infty}^{\infty}[f(x+\frac{2t}{n}) - f(x)]\cdot(\frac{\sin t}{t})^4\cdot dt$$

Now

$$|f(x+\frac{2t}{n}) - f(x)| \leq \omega_f(\frac{2t}{n}) \leq (2|t|+1)\cdot\omega_f(\frac{1}{n})$$

where the last equality follows from Lemma XX. Hence (7) gives

$$\max_x\,|\mathcal{J}_n^f(x) - f(x)| \leq \omega_f(\frac{1}{n})\cdot\frac{3}{2\pi}\cdot\int_{-\infty}^{\infty}(2|t|+1)\cdot(\frac{\sin t}{t})^4\cdot dt$$

Finally, by (6) the last factor is majorized by $1+\frac{6}{\pi}$ and Jackson's inequality follows.

A.12 A lower bound for polynomial approximation.

Denote by $\mathcal{T}_n$ the linear space of trigonometric polynomials of degree $\leq n$. For a 2π-periodic and continuous function f we put

$$\rho_f(n) = \min_{T\in\mathcal{T}_n}\,||f-T||$$

where $||\cdot||$ denotes the maximum norm over $[0,2\pi]$. We shall establish a lower bound for the ρ-numbers when certain sign-conditions hold for Fourier coefficients. In general, let f be a periodic function and for each positive integer n we find $T\in\mathcal{T}_n$ such that $||f-T|| = \rho_f(n)$. Since Fejer kernels do not increase maximum norms one has

$$(i)\qquad ||F_k^f - F_k^T|| \leq \rho_f(n)$$

for every positive integer k. Apply this with $k=n$ and $k=n+p$ where p is another positive integer. If $T\in\mathcal{T}_n$ the equation from Exercise XX gives

$$(ii)\qquad T = \frac{(n+p)\cdot\mathcal{F}_{n+p}(T) - n\cdot\mathcal{F}_n(T)}{p}$$

Since (i) hold for n, $n+p$ and $||f-T|| \leq \rho_f(n)$, the triangle inequality gives

$$(iii)\qquad ||f - \frac{(n+p)\cdot\mathcal{F}_{n+p}(f) - n\cdot\mathcal{F}_n(f)}{p}|| \leq 2\cdot\frac{n+p}{p}\cdot\rho_f(n)$$

Next, by the formula (§ xx) it follows that (iii) gives

$$||f - \frac{S_n(f)+\cdots S_{n+p-1}(f)}{p}|| \leq 2\cdot\frac{n+p}{p}\cdot\rho_f(n)$$

In particular we take $p=n$ and get the inequality

$$(*)\qquad ||f - \frac{S_n(f)+\cdots S_{2n-1}(f)}{n}|| \leq \frac{4}{n}\cdot\rho_f(n)$$

A.12 A special case. Assume that $f(x)$ is an even function on $[-\pi,\pi]$ which gives a Fourier series:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k\cdot\cos kx$$

A.12 Proposition *Let f be even as above and assume that $a_k \leq 0$ for every $k \geq 1$. Then the following inequality holds for every $n \geq 1$:*

$$f(0) - \frac{S_n(f)(0) + \cdots S_{2n-1}(f)(0)}{n} \leq -\sum_{k=2n}^{\infty} a_k$$

The easy verification is left to the reader. Taking the maximum norm over $[-\pi, \pi]$ it follows from (*) that

$$(**) \qquad \rho_f(n) \geq \frac{n}{4} \cdot \sum_{k=2n}^{\infty} |a_k|$$

holds when the sign conditions on the Fourier coefficients above are satisfied. Notice that (**) means that one has a lower bound for polynomial approximations of f.

A.13 The function $f(x) = \sin|x|$ It is obvious that

$$\omega_f(\frac{1}{n}) = \frac{1}{n}$$

Next, the periodic function $f(x($ is even and hence we only get a cosine-series. For each positive integer m we have:

$$a_k = \frac{2}{\pi} \int_0^{\pi} \sin x \cdot \cos kx \cdot dx$$

To evaluate these integrals we use the trigonometric formula

$$sin\,(k+1)x - \sin(k-1)x = 2\sin x \cdot \cos kx$$

Now the reader can verify that $a_\nu = 0$ when ν is odd while

$$a_{2k} = -\frac{4}{\pi} \cdot \frac{1}{2k^2 - 1}$$

Hence the requested sign conditions hold and (**) entails that

$$\rho_f(n) \geq \frac{n}{\pi} \cdot \sum_{k=n}^{\infty} \frac{1}{2k^2 - 1}$$

Here the right hand side is $\geq \frac{C}{n}$ for a constant C which is independent of n. So this example shows that the inequality (*) in § A.11 is sharp up to a multiple with a fixed constant.

A: The kernels of Dini, Fejer and Jackson

Denote by $C^0_{\text{per}}[0,2\pi]$ the family of complex-valued continuous functions $f(\theta)$ on $[0,2\pi]$ which satisfy $f(0)=f(2\pi)$. The Fourier coefficients of such a function f are defined by:

$$\widehat{f}(n)=\frac{1}{2\pi}\cdot\int_0^{2\pi} e^{-in\phi}f(\phi)\cdot d\phi$$

where n are integers. Fourier's partial sum of degree N is defined by

$$S^f_N(\theta)=\sum_{n=-N}^{n=N}\hat{f}(n)\cdot e^{in\theta} \tag{A.0}$$

A.1.The Dini kernel. If $N\geq 0$ we set

$$D_N(\theta)=\frac{1}{2\pi}\sum_{n=-N}^{n=N}e^{in\theta}$$

A.2 Exercise. Show that the folloeing hold for each $N\geq 0$:

$$S^f_N(\theta)=\int_0^{2\pi} D_N(\theta-\phi)\cdot f(\phi)\cdot d\phi=\int_0^{2\pi} D_N(\phi)\cdot f(\theta+\phi)\cdot d\phi$$

A.3 Proposition. *One has the formula*

$$D_N(\theta)=\frac{1}{2\pi}\cdot\frac{\sin((N+\frac{1}{2})\theta)}{\sin\frac{\theta}{2}} \tag{*}$$

Proof. We have

$$\sum_{n=-N}^{n=N}e^{in\theta}=e^{-iN\theta}\cdot\sum_{n=0}^{n=2N}e^{in\theta}=e^{-iN\theta}\cdot\frac{e^{i(2N+1)\theta}-1}{e^{i\theta}-1}=$$

$$e^{-iN\theta-i\theta/2}\cdot\frac{e^{i(2N+1)\theta}-1}{2i\cdot\sin\theta/2}=\frac{2i\cdot\sin((N+1/2)\theta)}{2i\cdot\sin\theta/2}$$

and (*) follows after division with $2i$.

A.4 The Fejer kernel. For each $N\geq 0$ we set

$$\mathcal{F}_N(\theta)=\frac{D_0(\theta)+\ldots+D_N(\theta)}{2\pi(N+1)}$$

A.5 Proposition *One has the formula*

$$\mathcal{F}_N(\theta)=\frac{1}{2\pi(N+1)}\cdot\frac{1-\cos((N+1)\theta)}{2\cdot\sin^2(\frac{\theta}{2})} \tag{**}$$

Proof. To each $\nu\geq 0$ we have $\sin((\nu+1/2)\theta)=\mathfrak{Im}\big[e^{i(\nu+1/2)\theta)}\big]$. Hence $F_N(\theta)$ is the imaginary part of

$$\frac{1}{2\pi(N+1)}\cdot\frac{e^{i\theta/2}}{\sin(\theta/2)}\cdot\sum_{\nu=0}^{\nu=N}e^{i\nu\theta}$$

Next, we have

$$e^{i\theta/2}\cdot\sum_{\nu=0}^{\nu=N}e^{i\nu\theta}=e^{i\theta/2}\cdot\frac{e^{i(N+1)\theta}-1}{e^{i\theta-1}}=\frac{e^{i(N+1)\theta}-1}{2i\cdot\sin(\theta/2)}$$

Since $i^2=-1$ we see that the imaginary part of the last term is equal to

$$\frac{1-\cos((N+1)\theta)}{2\cdot\sin(\frac{\theta}{2})}$$

and then (**) follows.

A.6 Fejer sums. For each f and every $N \geq 0$ we set

$$F_N^f(\theta) = \int_0^{2\pi} \mathcal{F}_N(\phi) \cdot f(\theta + \phi) \cdot d\phi$$

A.7 An inequality. If $0 < a > 2\pi$ we have the inequality

(i) $$\sin^2(\theta/2) \geq \sin^2(a/2) \quad : \quad a \leq \theta \leq 2\pi - a$$

Let f be given and denote by $M(f)$ the maximum norm of $|f(\theta)|$ over $[0, 2\pi]$. Then (i) gives

$$\int_a^{2\pi - a} \mathcal{F}_N(\phi) \cdot f(\theta + \phi) \cdot d\phi \leq$$

(A.7.1) $$\frac{M}{2\pi(N+1) \cdot \sin^2(a/2)} \int_a^{2\pi - a} (1 - \cos(N\phi)) \cdot d\phi \leq \frac{2M}{(N+1) \cdot \sin^2(a/2)}$$

A.8 Exercise. Given some θ_0 and $0 < a < \pi 0$ we set

$$\omega_f(a) = \max_{|\theta - \theta_0| \leq a} |f(\theta) - f(\theta_0)|$$

Use (A.7.1) to prove that

$$|F_N^f(\theta_0) - f(\theta_0)| \leq \frac{2M}{(N+1) \cdot \sin^2(a/2)} + \omega_f(a)$$

Conclude that the *uniform continuity* of the function f on $[0, 2\pi]$ implies that the sequence $\{F_N^f\}$ converges uniformly to f over the interval $[0, 2\pi]$.

A.9 The case when f is real-valued. When f is real-valued the Fourier series takes the form

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cdot \cos kx + \sum_{k=1}^{\infty} b_k \cdot \sin kx$$

Here $a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) \cdot dx$ and when $k \geq 1$ one has

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cdot \cos kx \cdot dx \quad : b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cdot \sin kx \cdot dx$$

Fourier's partial sum functions become

$$S_n(f) = \frac{a_0}{2} + \sum_{k=1}^{k=n} a_k \cdot \cos kx + \sum_{k=1}^{k=n} b_k \cdot \sin kx$$

The Jackson kernel

.

Above we proved that the Fejer sums converge uniformly to f. One may ask if there exists a constant C which is independent of both f and of n such that

$$(*)\qquad \max_\theta |f(\theta)-F_n^f(\theta)|\leq C\cdot\omega_f(\frac{1}{n})$$

hold for every $n\geq 1$, Examples show that (*) does not hold in general. To obtain a uniform constant C one must include an extra factor.

A.10 Exercise. Use (A.7-8) to show that there exists an absolute constant C such that

$$(**)\qquad ||f-\mathcal{F}_n(f)||\leq C\cdot\omega_f(\frac{1}{n})\cdot\big(1+\log^+\frac{1}{\omega_f(\frac{1}{n})})$$

hold for all continuous 2π-periodic functions f.

To attain (*) D. Jackson introduced a new kernel in his thesis *Über die Genauigkeit der Annährerung stegiger funktionen durch ganze rationala funktionen* from Göttingen in 1911. To each 2π-periodic and continuous function $f(x)$ on the real line and every $n\geq 1$ we set

$$\text{(A.10.1)}\qquad J_n^f(x)=\frac{3}{2\pi}\cdot\int_{-\infty}^\infty f(x+\frac{2t}{n})\cdot(\frac{\sin t}{t})^4\cdot dt$$

As explained in (A.xx) below one has

$$\text{(A.10.2)}\qquad \frac{3}{2\pi}\cdot\int_{-\infty}^\infty(\frac{\sin t}{t})^4\cdot dt=1$$

A.11 Theorem. *The function $J_n^f(x)$ is a trigonometric polynomial of degree $2n-1$ at most and one has the inequality*

$$\max_x |f(x)-J_n^f(x)|\leq(1+\frac{6}{\pi})\cdot\omega_f(\frac{1}{n})$$

The proof requires several steps. To begin with we prove that

$$\text{(i)}\qquad \frac{3}{2\pi}\cdot\int_{-\infty}^\infty(\frac{\sin t}{t})^4\cdot dt=1$$

To get (i) we use that the function $t\mapsto\sin^4 t$ is π-periodic which entails that (i) is equal to

$$\sum_{k=-\infty}^\infty\int_0^\pi\frac{\sin^4 t}{(t+k\pi)^4}\,dt$$

Next, by (xxx) one has the equation

$$\frac{1}{\sin^2 t}=\sum_{k=-\infty}^\infty\frac{1}{(t+k\pi)^2}$$

for every real t Taing a second order derivstive we obtain

$$\partial_t^2(\frac{1}{\sin^2 t})=6\cdot\sum_{k=-\infty}^\infty\frac{1}{(t+k\pi)^4}\implies(i)=\frac{1}{4\pi}\cdot\int_0^\pi\sin^4 t\cdot\partial_t^2(\frac{1}{\sin^2 t})\,dt$$

integration by parts gives

$$(i)=\frac{1}{4\pi}\cdot\int_0^\pi\partial_t^2(\sin^4 t)\cdot\frac{1}{\sin^2 t}\,dt=\frac{1}{4\pi}\int_0^\pi(12\sin^2 t\cdot\cos^2 t-4\cdot\sin^4 t]\cdot\frac{1}{\sin^2 t}\,dt=$$

$$=\frac{1}{4\pi}\int_0^\pi(12-16\cdot\sin^2 t)\,dt=1$$

It follows from (ii) that

$$|f(x) - J_n^f(x)| \leq \int_{-\infty^\infty} |f(x + \frac{2t}{n} - f(x)|\cdot$$

For each postov integer n the unequality in xxx shows that the right hsnd side above is majorised by

$$\omega_f(\frac{1}{n}) \cdot \int_{-\infty^\infty} (1 + 2/t/) \cdot\, dt$$

Proof. The variable substitution $t \to nt$ gives

$$J_n^f(x) = \frac{3}{2\pi n^3} \cdot \int_{-\infty}^{\infty} f(x+2t) \cdot (\frac{\sin nt}{t})^4 \cdot dt \tag{1}$$

Since $t \mapsto f(x+2t) \cdot sin^4\, nt$ is π-periodic it follows that (1) is equal to

$$\frac{3}{2\pi n^3} \cdot \int_0^\pi f(x+2t) \cdot \sum_{k=-\infty}^{\infty} \frac{\sin^4(nt)}{(k\pi + t)^4} \cdot dt \tag{2}$$

Next, recall from § XX that

$$\frac{1}{\sin^2 z} = \sum_{k=-\infty}^{\infty} \frac{1}{(z + k\pi)^2}$$

Taking a second derivative when $z = t$ is real it follows that

$$\partial_t^2(\frac{1}{\sin^2 t}) = \frac{1}{6} \cdot \sum_{k=-\infty}^{\infty} \frac{1}{(t + k\pi)^4} \tag{3}$$

Hence we obtain

$$J_n^f(x) = \frac{1}{4\pi n^3} \cdot \int_0^\pi f(x+2t) \cdot \sin^4(nt) \cdot \partial_t^2(\frac{1}{\sin^2 t})\, dt \tag{4}$$

Integration by parts give the equality

$$\int_{-\infty}^{\infty} (\frac{\sin nt}{t})^4\, dt = \frac{1}{6} \int_0^\pi \sin^4 t \cdot \partial_t^2(\frac{1}{\sin^2 t})\, dt = \frac{4}{3} \int_0^\pi \cos^2 t\, dt = \frac{2\pi}{3} \tag{5}$$

Next, we leave it to the reader to verify the inequality

$$\frac{3}{2\pi} \int_{-\infty}^{\infty} (1 + 2|t|) \cdot (\frac{\sin t}{t})^4 \cdot dt \leq 1 + \frac{6}{\pi} \tag{6}$$

From the above where we use (1) and (*) it follows that

$$\mathcal{J}_n^f(x) - f(x) = \frac{3}{2\pi} \cdot \int_{-\infty}^{\infty} [f(x + \frac{2t}{n}) - f(x)] \cdot (\frac{\sin t}{t})^4 \cdot dt \tag{7}$$

Now

$$|f(x + \frac{2t}{n}) - f(x)| \leq \omega_f(\frac{2t}{n}) \leq (2|t| + 1) \cdot \omega_f(\frac{1}{n})$$

where the last equality follows from Lemma XX. Hence (7) gives

$$\max_x |\mathcal{J}_n^f(x) - f(x)| \leq \omega_f(\frac{1}{n}) \cdot \frac{3}{2\pi} \cdot \int_{-\infty}^{\infty} (2|t| + 1) \cdot (\frac{\sin t}{t})^4 \cdot dt$$

Finally, by (6) the last factor is majorized by $1 + \frac{6}{\pi}$ and Jackson's inequality follows.

A.12 A lower bound for polynomial approximation.

Denote by $\mathcal{T}_n$ the linear space of trigonometric polynomials of degree $\leq n$. For a 2π-periodic and continuous function f we put

$$\rho_f(n) = \min_{T \in \mathcal{T}_n} ||f - T||$$

where $||\cdot||$ denotes the maximum norm over $[0,2\pi]$. We shall establish a lower bound for the ρ-numbers when certain sign-conditions hold for Fourier coefficients. In general, let f be a periodic function and for each positive integer n we find $T\in\mathcal{T}_n$ such that $||f-T||=\rho_f(n)$. Since Fejer kernels do not increase maximum norms one has

(i)
$$||F_k^f-F_k^T||\leq\rho_f(n)$$

for every positive integer k. Apply this with $k=n$ and $k=n+p$ where p is another positive integer. If $T\in\mathcal{T}_n$ the equation from Exercise XX gives

(ii)
$$T=\frac{(n+p)\cdot\mathcal{F}_{n+p}(T)-n\cdot\mathcal{F}_n(T)}{p}$$

Since (i) hold for n, $n+p$ and $||f-T||\leq\rho_f(n)$, the triangle inequality gives

(iii)
$$||f-\frac{(n+p)\cdot\mathcal{F}_{n+p}(f)-n\cdot\mathcal{F}_n(f)}{p}||\leq 2\cdot\frac{n+p}{p}\cdot\rho_f(n)$$

Next, by the formula (§ xx) it follows that (iii) gives

$$||f-\frac{S_n(f)+\cdots S_{n+p-1}(f)}{p}||\leq 2\cdot\frac{n+p}{p}\cdot\rho_f(n)$$

In particular we take $p=n$ and get the inequality

(*)
$$||f-\frac{S_n(f)+\cdots S_{2n-1}(f)}{n}||\leq\frac{4}{n}\cdot\rho_f(n)$$

A.12 A special case. Assume that $f(x)$ is an even function on $[-\pi,\pi]$ which gives a Fourier series:

$$f(x)=\frac{a_0}{2}+\sum_{k=1}^{\infty}a_k\cdot\cos kx$$

A.12 Proposition *Let f be even as above and assume that $a_k\leq 0$ for every $k\geq 1$. Then the following inequality holds for every $n\geq 1$:*

$$f(0)-\frac{S_n(f)(0)+\cdots S_{2n-1}(f)(0)}{n}\leq-\sum_{k=2n}^{\infty}a_k$$

The easy verification is left to the reader. Taking the maximum norm over $[-\pi,\pi]$ it follows from (*) that

(**)
$$\rho_f(n)\geq\frac{n}{4}\cdot\sum_{k=2n}^{\infty}|a_k|$$

holds when the sign conditions on the Fourier coefficients above are satisfied. Notice that (**) means that one has a lower bound for polynomial approximations of f.

Next, the function

$$\sin^4(nz)\cdot\partial_z^2(\frac{1}{\sin^2 z})$$

is entire and even and the reader may verify that it is a finite sum of entire cosine-functions which implies that the Jackson kernel is expressed by a finite sum of integrals:

(4)
$$\mathcal{J}_f^n(x)=\sum_{k=0}^{2n-1}c_k\int_0^{2\pi}f(u)\cdot\cos k(x-u))\,du$$

In particular $\mathcal{J}_f^n(x)$ is a trigonometric polynomial of degree $2n-1$ a most. Integration by parts give the equality

(5)
$$\int_{-\infty}^{\infty}(\frac{\sin nt}{t})^4\,dt=\frac{1}{6}\int_0^{\pi}\sin^4 t\cdot\partial_t^2(\frac{1}{\sin^2 t})\,dt=\frac{4}{3}\int_0^{\pi}\cos^2 t\,dt=\frac{2\pi}{3}$$

Next, we leave it to the reader to verify the inequality

$$\frac{3}{2\pi}\int_{-\infty}^{\infty}(1+2|t|)\cdot\big(\frac{\sin t}{t}\big)^4\cdot dt\leq 1+\frac{6}{\pi} \tag{6}$$

A.13 The function $f(x)=\sin|x|$ It is obvious that

$$\omega_f(\frac{1}{n})=\frac{1}{n}$$

Next, the periodic function $f(x($ is even and hence we only get a cosine-series. For each positive integer m we have:

$$a_k=\frac{2}{\pi}\int_0^{\pi}\sin x\cdot\cos kx\cdot dx$$

To evaluate these integrals we use the trigonometric formula

$$sin\,(k+1)x-\sin(k-1)x=2\sin x\cdot\cos kx$$

Now the reader can verify that $a_\nu=0$ when ν is odd while

$$a_{2k}=-\frac{4}{\pi}\cdot\frac{1}{2k^2-1}$$

Hence the requested sign conditions hold and (**) entails that

$$\rho_f(n)\geq\frac{n}{\pi}\cdot\sum_{k=n}^{\infty}\frac{1}{2k^2-1}$$

Here the right hand side is $\geq\frac{C}{n}$ for a constant C which is independent of n. So this example shows that the inequality (*) in § A.11 is sharp up to a multiple with a fixed constant.