Lecture 1

Frits Beukers

Arithmetic of values of E- and G-function

A general transcendence problem

Let $f(z) \in \overline{\mathbb{Q}}[[z]]$ be power series in z with coefficients in $\overline{\mathbb{Q}}$, with positive radius of convergence ρ. We assume $f(z)$ is not algebraic over $\overline{\mathbb{Q}}(z)$.

Question

Let $\alpha \in \overline{\mathbb{Q}}$ and suppose $0 < |\alpha| < \rho$. Is $f(\alpha)$ transcendental?

A bizarre function

There exist non-algebraic $f \in \mathbb{Q}[[z]]$ with $\rho = \infty$ such that

$$f(\alpha) \in \overline{\mathbb{Q}} \text{ for all } \alpha \in \overline{\mathbb{Q}}$$

A bizarre function

There exist non-algebraic $f \in \mathbb{Q}[[z]]$ with $\rho = \infty$ such that

$$f(\alpha) \in \overline{\mathbb{Q}} \text{ for all } \alpha \in \overline{\mathbb{Q}}$$

Idea of construction:
Enumerate the elements of $\mathbb{Z}[z]$ by $P_1, P_2, \ldots$ and consider

$$f(z) = 1 + \sum_{k=1}^{\infty} c_k z^k P_1(z) \cdots P_k(z)$$

where $c_k \in \mathbb{Q}$ are chosen such that the resulting f has infinite radius of convergence.

A bizarre function

There exist non-algebraic $f \in \mathbb{Q}[[z]]$ with $\rho = \infty$ such that

$$f(\alpha) \in \overline{\mathbb{Q}} \text{ for all } \alpha \in \overline{\mathbb{Q}}$$

Idea of construction:
Enumerate the elements of $\mathbb{Z}[z]$ by $P_1, P_2, \dots$ and consider

$$f(z) = 1 + \sum_{k=1}^{\infty} c_k z^k P_1(z) \cdots P_k(z)$$

where $c_k \in \mathbb{Q}$ are chosen such that the resulting f has infinite radius of convergence.
Question: Do there exist $f \in \mathbb{Z}[[z]]$ with positive radius of converge ρ such that $f(\alpha) \in \overline{\mathbb{Q}}$ for all $\alpha \in \overline{\mathbb{Q}}$ and $|\alpha| < \rho$?

Lindemann-Weierstrass theorem

Around 1882 F.Lindemann proved the transcendence of π. In fact his method yielded more.

Lindemann-Weierstrass

Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be distinct algebraic numbers. Then

$$e^{\alpha_1}, e^{\alpha_2}, \ldots, e^{\alpha_n}$$

are linearly independent over $\overline{\mathbb{Q}}$.

Lindemann-Weierstrass theorem

Around 1882 F.Lindemann proved the transcendence of π. In fact his method yielded more.

Lindemann-Weierstrass

Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be distinct algebraic numbers. Then

$$e^{\alpha_1}, e^{\alpha_2}, \ldots, e^{\alpha_n}$$

are linearly independent over $\overline{\mathbb{Q}}$.

Application: π is transcendental.

Lindemann-Weierstrass theorem

Around 1882 F.Lindemann proved the transcendence of π. In fact his method yielded more.

Lindemann-Weierstrass

Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be distinct algebraic numbers. Then

$$e^{\alpha_1}, e^{\alpha_2}, \ldots, e^{\alpha_n}$$

are linearly independent over $\overline{\mathbb{Q}}$.

Application: π is transcendental.
Proof: Suppose π were algebraic. Take $\alpha_1 = 0, \alpha_2 = \pi i$. Then Lindemann-Weierstrass implies that $1, e^{\pi i}$ are $\overline{\mathbb{Q}}$-linear independent, contradicting $e^{\pi i} = -1$.

Some notations

For any element $\alpha \in \overline{\mathbb{Q}}$ we define $\overline{|\alpha|}$ to be maximum of the absolute values of all conjugates of α. We call it the *size* of α.

Some notations

For any element $\alpha \in \overline{\mathbb{Q}}$ we define $\overline{|\alpha|}$ to be maximum of the absolute values of all conjugates of α. We call it the *size* of α. By $\mathrm{den}(\alpha)$ we denote the *denominator* of α, the smallest positive integer d such that $d\alpha$ is an algebraic integer.

Some notations

For any element $\alpha \in \overline{\mathbb{Q}}$ we define $\overline{|\alpha|}$ to be maximum of the absolute values of all conjugates of α. We call it the *size* of α.
By $\mathrm{den}(\alpha)$ we denote the *denominator* of α, the smallest positive integer d such that $d\alpha$ is an algebraic integer.
For any set of $\alpha_1, \ldots, \alpha_r$ we denote by $\mathrm{den}(\alpha_1, \ldots, \alpha_r)$ the lowest common multiple of the denominators of $\alpha_1, \ldots, \alpha_r$.

E-function definition

Definition

An entire function $f(z)$ given by a powerseries

$$\sum_{k=0}^{\infty} \frac{a_k}{k!} z^k$$

with $a_k \in \overline{\mathbb{Q}}$ for all k, is called an E-function if

1. $f(z)$ satisfies a linear differential equation with coefficients in $\overline{\mathbb{Q}}(z)$.
2. Both $\overline{|a_k|}$ and the common denominator $\mathrm{den}(a_0, \ldots, a_k)$ are bounded by an exponential bound of the form C^k, where $C > 0$ depends only on f.

E-function definition

Definition

An entire function $f(z)$ given by a powerseries

$$\sum_{k=0}^{\infty} \frac{a_k}{k!} z^k$$

with $a_k \in \overline{\mathbb{Q}}$ for all k, is called an E-function if

1. $f(z)$ satisfies a linear differential equation with coefficients in $\overline{\mathbb{Q}}(z)$.
2. Both $\overline{|a_k|}$ and the common denominator $\mathrm{den}(a_0, \ldots, a_k)$ are bounded by an exponential bound of the form C^k, where $C > 0$ depends only on f.

Remark: For any E-function there exists a finite extension K of $\mathbb{Q}$ such that $a_k \in K$ for all k.

E-function definition

Definition

An entire function $f(z)$ given by a powerseries

$$\sum_{k=0}^{\infty} \frac{a_k}{k!} z^k$$

with $a_k \in \overline{\mathbb{Q}}$ for all k, is called an E-function if

1. $f(z)$ satisfies a linear differential equation with coefficients in $\overline{\mathbb{Q}}(z)$.
2. Both $\overline{|a_k|}$ and the common denominator $\mathrm{den}(a_0, \ldots, a_k)$ are bounded by an exponential bound of the form C^k, where $C > 0$ depends only on f.

Remark: For any E-function there exists a finite extension K of $\mathbb{Q}$ such that $a_k \in K$ for all k. In most cases of interest: $K = \mathbb{Q}$.

More general?

Siegel used instead of the exponential bound C^k in condition 2) the (seemingly) less restrictive $c_\epsilon (k!)^\epsilon$ for every $\epsilon > 0$. However, it is conjectured that the bound $(k!)^\epsilon$ together with the condition of a linear differential equation is enough to garantee that we have the exponential bound C^k.

More general?

Siegel used instead of the exponential bound C^k in condition 2) the (seemingly) less restrictive $c_\epsilon (k!)^\epsilon$ for every $\epsilon > 0$. However, it is conjectured that the bound $(k!)^\epsilon$ together with the condition of a linear differential equation is enough to garantee that we have the exponential bound C^k.

Therefore we shall stick to our definition above.

E-function examples

$$\exp(az) = \sum_{k=0}^{\infty} \frac{a^k z^k}{k!}, \ a \in \overline{\mathbb{Q}}^*$$

$$J_0(-z^2) = \sum_{k=0}^{\infty} \frac{z^{2k}}{k!k!} = \sum_{k\geq 0} \binom{2n}{n} \frac{z^{2k}}{(2k)!}$$

E-function examples

$$\exp(az) = \sum_{k=0}^{\infty} \frac{a^k z^k}{k!}, \ a \in \overline{\mathbb{Q}}^*$$

$$J_0(-z^2) = \sum_{k=0}^{\infty} \frac{z^{2k}}{k!k!} = \sum_{k\geq 0} \binom{2n}{n} \frac{z^{2k}}{(2k)!}$$

The corresponding differential equations read

$$y' - ay = 0$$

$$zy'' + y' - 4zy = 0$$

Hypergeometric example

A very general example, the *confluent hypergeometric series*,

$$ {}_pF_q\left(\begin{matrix}\alpha_1,\ldots,\alpha_p\\ \beta_1,\ldots,\beta_q\end{matrix}\,\middle|\, z^{q+1-p}\right) = \sum_{k=0}^{\infty} \frac{(\alpha_1)_k\cdots(\alpha_p)_k}{(\beta_1)_k\cdots(\beta_q)_k k!} z^{(q+1-p)k} $$

where $q \geq p$ (confluence) and $\alpha_i, \beta_j \in \mathbb{Q}$ for all i, j.
$(x)_n$ is the Pochhammer symbol defined by $x(x+1)\cdots(x+n-1)$.

Hypergeometric example

A very general example, the *confluent hypergeometric series*,

$$ {}_pF_q\left(\begin{matrix}\alpha_1,\dots,\alpha_p\\ \beta_1,\dots,\beta_q\end{matrix}\,\middle|\, z^{q+1-p}\right)=\sum_{k=0}^{\infty}\frac{(\alpha_1)_k\cdots(\alpha_p)_k}{(\beta_1)_k\cdots(\beta_q)_k k!}z^{(q+1-p)k} $$

where $q \geq p$ (confluence) and $\alpha_i, \beta_j \in \mathbb{Q}$ for all i, j.
$(x)_n$ is the Pochhammer symbol defined by $x(x+1)\cdots(x+n-1)$.
${}_pF_q$ satisfies a linear differential equation of order $q+1$.

Differential ring structure

The E-functions form a so-called differential ring. More precisely,

Proposition

Let $f(z), g(z)$ be E-functions. Then the following functions are again E-functions

- $f'(z)$
- $f(z) + g(z)$
- $f(z)g(z)$

Differential ring structure

The E-functions form a so-called differential ring. More precisely,

Proposition

Let $f(z), g(z)$ be E-functions. Then the following functions are again E-functions

- $f'(z)$
- $f(z) + g(z)$
- $f(z)g(z)$

Let L be any subfield of $\mathbb{C}$. Let $f \in L[[z]]$. Then the following two statements are equivalent

1. f satisfies a linear differential equation with coefficients in $L(z)$.
2. $L(z)$-vectorspace spanned by f and its derivatives is finite dimensional.

Product of E-functions

We show that if f, g are E-functions then so is fg. There exists a number field K so that $f, g \in K[[z]]$.

Product of E-functions

We show that if f, g are E-functions then so is fg. There exists a number field K so that $f, g \in K[[z]]$.

Differential equation The $K(z)$-vectorspace spanned by fg and its derivatives is contained in the space spanned by $f^{(r)}g^{(s)}$. The latter space is finite-dimensional, hence also the former space.

Products of E-functions, cont'd

We show that if f, g are E-functions then so is fg. There exists a number field K so that $f, g \in K[[z]]$.

Coefficient estimates Write

$$f = \sum_{k \geq 0} f_k \frac{z^k}{k!}, \quad g = \sum_{k \geq 0} g_k \frac{z^k}{k!}, \quad fg = \sum_{k \geq 0} h_k \frac{z^k}{k!}.$$

Then

$$h_k = \sum_{r+s=k}^{k} \binom{k}{r} f_r g_s.$$

Products of E-functions, cont'd

We show that if f, g are E-functions then so is fg. There exists a number field K so that $f, g \in K[[z]]$.
Coefficient estimates Write

$$f = \sum_{k \geq 0} f_k \frac{z^k}{k!}, \quad g = \sum_{k \geq 0} g_k \frac{z^k}{k!}, \quad fg = \sum_{k \geq 0} h_k \frac{z^k}{k!}.$$

Then

$$h_k = \sum_{r+s=k}^{k} \binom{k}{r} f_r g_s.$$

Notice:

1. $\mathrm{den}(h_k)$ divides $\mathrm{den}(f_0, \ldots, f_k)\mathrm{den}(g_0, \ldots, g_k)$
2. $$\overline{|h_k|} \leq 2^k (\max_{0 \leq r \leq k} \overline{|f_r|})(\max_{0 \leq s \leq k} \overline{|g_s|}).$$

First order systems

Let L be any subfield of $\mathbb{C}$. An $n \times n$-system of first order linear differential equations is given by

$$\frac{d}{dz}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

where $A_{ij} \in L(z)$ for all i, j.

First order systems

Let L be any subfield of $\mathbb{C}$. An $n \times n$-system of first order linear differential equations is given by

$$\frac{d}{dz}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

where $A_{ij} \in L(z)$ for all i, j.
We abbreviate by

$$\mathbf{y}' = A\mathbf{y}$$

where A is the $n \times n$-matrix with entries A_{ij}.

First order systems

Let L be any subfield of $\mathbb{C}$. An $n \times n$-system of first order linear differential equations is given by

$$\frac{d}{dz}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

where $A_{ij} \in L(z)$ for all i, j.
We abbreviate by

$$\mathbf{y}' = A\mathbf{y}$$

where A is the $n \times n$-matrix with entries A_{ij}.
Let $T(z)$be the common denominator of the A_{ij}. The zeros of $T(z)$ are called the *singularities* of the system.

From equations to systems

Consider the linear n-th order differential equation

$$y^{(n)} + p_1 y^{(n-1)} + p_2 y^{(n-2)} + \cdots + p_{n-1} y' + p_n y = 0, \ \ p_i \in L(z)$$

From equations to systems

Consider the linear n-th order differential equation

$$y^{(n)} + p_1 y^{(n-1)} + p_2 y^{(n-2)} + \cdots + p_{n-1} y' + p_n y = 0, \ \ p_i \in L(z)$$

Put

$$y_1 = y, \ y_2 = y', \ \ldots, \ y_n = y^{(n-1)}$$

From equations to systems

Consider the linear n-th order differential equation

$$y^{(n)} + p_1 y^{(n-1)} + p_2 y^{(n-2)} + \cdots + p_{n-1} y' + p_n y = 0, \;\; p_i \in L(z)$$

Put

$$y_1 = y, \; y_2 = y', \; \ldots, \; y_n = y^{(n-1)}$$

Note that

$$y_1' = y_2, \; y_2' = y_3, \; \ldots, \; y_{n-1}' = y_n.$$

From equations to systems

Consider the linear n-th order differential equation

$$y^{(n)} + p_1 y^{(n-1)} + p_2 y^{(n-2)} + \cdots + p_{n-1} y' + p_n y = 0, \; p_i \in L(z)$$

Put

$$y_1 = y, \; y_2 = y', \; \ldots, \; y_n = y^{(n-1)}$$

Note that

$$y_1' = y_2, \; y_2' = y_3, \; \ldots, \; y_{n-1}' = y_n.$$

Finally,

$$y_n' = -p_1 y_n - p_2 y_{n-1} - \ldots - p_n y_1.$$

From equations to systems

Consider the linear n-th order differential equation

$$y^{(n)} + p_1 y^{(n-1)} + p_2 y^{(n-2)} + \cdots + p_{n-1} y' + p_n y = 0, \ \ p_i \in L(z)$$

Put

$$y_1 = y, \ y_2 = y', \ \ldots, \ y_n = y^{(n-1)}$$

Note that

$$y_1' = y_2, \ y_2' = y_3, \ \ldots, \ y_{n-1}' = y_n.$$

Finally,

$$y_n' = -p_1 y_n - p_2 y_{n-1} - \ldots - p_n y_1.$$

Rewrite as

$$\frac{d}{dz}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -p_n & -p_{n-1} & -p_{n-2} & \cdots & -p_1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

From systems to equations

Consider

$$\mathbf{y}' = A\mathbf{y}$$

Replace $\mathbf{y}$ by $S\mathbf{y}$ in the system, where $S \in GL(n, L(z))$, we obtain a new system for the new $\mathbf{y}$,

$$\mathbf{y}' = (S^{-1}AS - S^{-1}S')\mathbf{y}.$$

From systems to equations

Consider

$$\mathbf{y}' = A\mathbf{y}$$

Replace $\mathbf{y}$ by $S\mathbf{y}$ in the system, where $S \in GL(n, L(z))$, we obtain a new system for the new $\mathbf{y}$,

$$\mathbf{y}' = (S^{-1}AS - S^{-1}S')\mathbf{y}.$$

Two $n \times n$-systems with coefficient matrices $A, B \in M_n(L(z))$ are called *equivalent over* $L(z)$ if there exists $S \in GL(n, L(z))$ such that $B = S^{-1}AS - S^{-1}S'$.

From systems to equations

Consider

$$\mathbf{y}' = A\mathbf{y}$$

Replace $\mathbf{y}$ by $S\mathbf{y}$ in the system, where $S \in GL(n, L(z))$, we obtain a new system for the new $\mathbf{y}$,

$$\mathbf{y}' = (S^{-1}AS - S^{-1}S')\mathbf{y}.$$

Two $n \times n$-systems with coefficient matrices $A, B \in M_n(L(z))$ are called *equivalent over* $L(z)$ if there exists $S \in GL(n, L(z))$ such that $B = S^{-1}AS - S^{-1}S'$.

Proposition

Any system of linear first order differential equations over $L(z)$ is equivalent over $L(z)$ to a system which comes from a differential equation.

Cyclic vector Lemma

To go from system to equation consider

$$F(z) = P_1 y_1 + \cdots + P_n y_n$$

where $P_i \in \mathbb{C}(z)$.

Cyclic vector Lemma

To go from system to equation consider

$$F(z) = P_1 y_1 + \cdots + P_n y_n$$

where $P_i \in \mathbb{C}(z)$. By differentiation and use of the first order system we find that for every m there exist $P_{m,i}$ such that

$$F^{(m)} = P_{m1} y_1 + \cdots + P_{mn} y_n.$$

Cyclic vector Lemma

To go from system to equation consider

$$F(z) = P_1 y_1 + \cdots + P_n y_n$$

where $P_i \in \mathbb{C}(z)$. By differentiation and use of the first order system we find that for every m there exist $P_{m,i}$ such that

$$F^{(m)} = P_{m1} y_1 + \cdots + P_{mn} y_n.$$

Let

$$\Delta(z) = \begin{vmatrix} P_{01} & \cdots & P_{0n} \\ \vdots & & \vdots \\ P_{n-1,1} & \cdots & P_{n-1,n} \end{vmatrix}$$

If $\Delta(z) \not\equiv 0$ the matrix (P_{ij}) gives us an equivalence from the system to the equation satisfied by F.

Cyclic vector Lemma

To go from system to equation consider

$$F(z) = P_1 y_1 + \cdots + P_n y_n$$

where $P_i \in \mathbb{C}(z)$. By differentiation and use of the first order system we find that for every m there exist $P_{m,i}$ such that

$$F^{(m)} = P_{m1} y_1 + \cdots + P_{mn} y_n.$$

Let

$$\Delta(z) = \begin{vmatrix} P_{01} & \cdots & P_{0n} \\ \vdots & & \vdots \\ P_{n-1,1} & \cdots & P_{n-1,n} \end{vmatrix}$$

If $\Delta(z) \not\equiv 0$ the matrix (P_{ij}) gives us an equivalence from the system to the equation satisfied by F.

Cyclic vector Lemma

There exist $P_1, \ldots, P_n$ such that $\Delta(z) \neq 0$.

Siegel-Shidlovskii theorem

Siegel-Shidlovskii, 1929, 1956

Let $(f_1(z), \ldots, f_n(z)$ be a solution vector of a system of first order equations of the form

$$\mathbf{y}'(z) = A(z)\mathbf{y}(z)$$

and suppose that the $f_i(z)$ are E-functions. Let $T(z)$ be the common denominator of the entries of $A(z)$. Let $\alpha \in \overline{\mathbb{Q}}$ and suppose $\alpha T(\alpha) \neq 0$. Then

$$\mathrm{degtr}_{\overline{\mathbb{Q}}}(f_1(\alpha), f_2(\alpha), \ldots, f_n(\alpha)) = \mathrm{degtr}_{\mathbb{C}(z)}(f_1(z), f_2(z), \ldots, f_n(z))$$

Siegel-Shidlovskii theorem

Siegel-Shidlovskii, 1929, 1956

Let $(f_1(z), \ldots, f_n(z)$ be a solution vector of a system of first order equations of the form

$$\mathbf{y}'(z) = A(z)\mathbf{y}(z)$$

and suppose that the $f_i(z)$ are E-functions. Let $T(z)$ be the common denominator of the entries of $A(z)$. Let $\alpha \in \overline{\mathbb{Q}}$ and suppose $\alpha T(\alpha) \neq 0$. Then

$$\mathrm{degtr}_{\overline{\mathbb{Q}}}(f_1(\alpha), f_2(\alpha), \ldots, f_n(\alpha)) = \mathrm{degtr}_{\mathbb{C}(z)}(f_1(z), f_2(z), \ldots, f_n(z))$$

In particular, if the $f_i(z)$ are algebraically independent over $\mathbb{C}(z)$ then the values at $z = \alpha$ are algebraically independent over $\overline{\mathbb{Q}}$ (or $\mathbb{Q}$, which amounts to the same).

Algebraic relations between E-function

In the 1960's and 70's much energy has gone into showing algebraic independence of (mainly hypergeometrical) E-functions.

Algebraic relations between E-function

In the 1960's and 70's much energy has gone into showing algebraic independence of (mainly hypergeometrical) E-functions. In the 1980's the tool of differential galois theory was used, which clarified very much of the earlier work.

Algebraic relations between E-function

In the 1960's and 70's much energy has gone into showing algebraic independence of (mainly hypergeometrical) E-functions. In the 1980's the tool of differential galois theory was used, which clarified very much of the earlier work.

We display a non-trivial example of E-functions.

Consider

$$f(z) = \sum_{k=0}^{\infty} \frac{((2k)!)^2}{(k!)^2(6k)!} z^k.$$

Then $f(z^4)$ is an E-function satisfying a differential equation of order 5. The differential galois group is $SO(5, \mathbb{C})$. Dimension of its orbits is 4 and we have a quadratic form Q with coefficients in $\mathbb{Q}(z)$ such that

$$Q(f, f', f'', f''', f'''') = 1$$

Example of a relation between E-functions

Explicitly,

$$f(z) = \sum_{k=0}^{\infty} \frac{((2k)!)^2}{(k!)^2(6k)!}(2916z)^k$$

satisfies

$$\mathbf{F}^t \mathscr{Q} \mathbf{F} = (z)$$

where $\mathbf{F} = (f(z), Df(z), D^2 f(z), D^3 f(z), D^4 f(z))^t$ with $D = z\frac{d}{dz}$ and

$$\mathscr{Q} = \begin{pmatrix} z - 324z^2 & -18z & 198z & -486z & 324z \\ -18z & -\frac{10}{9} & \frac{23}{2} & -28 & 18 \\ 198z & \frac{23}{2} & -120 & 297 & -198 \\ -486z & -28 & 297 & -729 & 486 \\ 324z & 18 & -198 & 486 & -324 \end{pmatrix}$$

Principle of Siegel's method

Suppose we have one or more independent linear relations between $f_1(\alpha), \ldots, f_n(\alpha)$ with coefficients in $\overline{\mathbb{Q}}$. Say

$$\beta_{i1} f_1(\alpha) + \beta_{i2} f_2(\alpha) + \cdots + \beta_{in} f_n(\alpha) = 0$$

with $i = 1, \ldots, m$.

Principle of Siegel's method

Suppose we have one or more independent linear relations between $f_1(\alpha), \ldots, f_n(\alpha)$ with coefficients in $\overline{\mathbb{Q}}$. Say

$$\beta_{i1} f_1(\alpha) + \beta_{i2} f_2(\alpha) + \cdots + \beta_{in} f_n(\alpha) = 0$$

with $i = 1, \ldots, m$.
Let K be the number field generated by α, the coefficients of the f_i and the β_{ij}.

Principle of Siegel's method

Suppose we have one or more independent linear relations between $f_1(\alpha), \ldots, f_n(\alpha)$ with coefficients in $\overline{\mathbb{Q}}$. Say

$$\beta_{i1} f_1(\alpha) + \beta_{i2} f_2(\alpha) + \cdots + \beta_{in} f_n(\alpha) = 0$$

with $i = 1, \ldots, m$.
Let K be the number field generated by α, the coefficients of the f_i and the β_{ij}.
Suppose we have $n - m$ additional inequalities of the form

$$|B_{i1} f_1(\alpha) + B_{i2} f_2(\alpha) + \cdots + B_{in} f_n(\alpha)| < \delta$$

with $i = m + 1, \ldots, n$, B_{ij} are algebraic integers in K and such that $n \times n$ matrix with the rows $\beta_{i1}, \ldots, \beta_{in}$ and $B_{i1}, \ldots, B_{in}$ is non-singular.

Principle of Siegel's method

Suppose we have one or more independent linear relations between $f_1(\alpha), \ldots, f_n(\alpha)$ with coefficients in $\overline{\mathbb{Q}}$. Say

$$\beta_{i1} f_1(\alpha) + \beta_{i2} f_2(\alpha) + \cdots + \beta_{in} f_n(\alpha) = 0$$

with $i = 1, \ldots, m$.
Let K be the number field generated by α, the coefficients of the f_i and the β_{ij}.
Suppose we have $n - m$ additional inequalities of the form

$$|B_{i1} f_1(\alpha) + B_{i2} f_2(\alpha) + \cdots + B_{in} f_n(\alpha)| < \delta$$

with $i = m+1, \ldots, n$, B_{ij} are algebraic integers in K and such that $n \times n$ matrix with the rows $\beta_{i1}, \ldots, \beta_{in}$ and $B_{i1}, \ldots, B_{in}$ is non-singular.
Then there exists a constant $c > 0$, independent of the choice of the B_{ij}, such that $1 < c\delta B^{d(n-m)-1}$, where d is the degree of K and B the maximum of the $\overline{|B_{ij}|}$.

The arithmetic inequality

So we have

$$\begin{pmatrix} \beta_{11} & \dots & \beta_{1n} \\ \vdots & & \vdots \\ \beta_{m1} & \dots & \beta_{mn} \\ B_{m+1,1} & \dots & B_{m+1,n} \\ \vdots & & \vdots \\ B_{n1} & \dots & B_{nn} \end{pmatrix} \begin{pmatrix} f_1(\alpha) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ f_n(\alpha) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta_{m+1} \\ \vdots \\ \delta_n \end{pmatrix} \quad |\delta_j| < \delta.$$

The arithmetic inequality

So we have

$$\begin{pmatrix} \beta_{11} & \dots & \beta_{1n} \\ \vdots & & \vdots \\ \beta_{m1} & \dots & \beta_{mn} \\ B_{m+1,1} & \dots & B_{m+1,n} \\ \vdots & & \vdots \\ B_{n1} & \dots & B_{nn} \end{pmatrix} \begin{pmatrix} f_1(\alpha) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ f_n(\alpha) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta_{m+1} \\ \vdots \\ \delta_n \end{pmatrix} \quad |\delta_j| < \delta.$$

Using Cramer's rule $\Delta f_j(\alpha) = \Delta_j, \quad j = 1, \dots, n$
where Δ is the coefficient determinant and Δ_j is the determinant after replacing the j-th column with the right hand column.
Suppose $f_1(\alpha) \neq 0$.
Then $|\Delta| \cdot |f_1(\alpha)| = |\Delta_1| \leq c_1 B^{n-m-1}\delta$
Furthermore, $|\Delta| \geq \overline{|\Delta|}^{-(d-1)} \geq (n! B^{n-m})^{-(d-1)}$.

The arithmetic inequality

So we have

$$\begin{pmatrix} \beta_{11} & \dots & \beta_{1n} \\ \vdots & & \vdots \\ \beta_{m1} & \dots & \beta_{mn} \\ B_{m+1,1} & \dots & B_{m+1,n} \\ \vdots & & \vdots \\ B_{n1} & \dots & B_{nn} \end{pmatrix} \begin{pmatrix} f_1(\alpha) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ f_n(\alpha) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta_{m+1} \\ \vdots \\ \delta_n \end{pmatrix} \quad |\delta_j| < \delta.$$

Using Cramer's rule $\Delta f_j(\alpha) = \Delta_j, \quad j = 1, \dots, n$
where Δ is the coefficient determinant and Δ_j is the determinant after replacing the j-th column with the right hand column.
Suppose $f_1(\alpha) \neq 0$.
Then $|\Delta| \cdot |f_1(\alpha)| = |\Delta_1| \leq c_1 B^{n-m-1}\delta$
Furthermore, $|\Delta| \geq \overline{|\Delta|}^{-(d-1)} \geq (n! B^{n-m})^{-(d-1)}$.
Our estimate follows.

Siegel construction

Fix N sufficiently large and choose $\epsilon > 1$. Using linear algebra and Siegel's Lemma, construct polynomials $P_1, \ldots, P_n \in K[z]$ of degrees $\leq N$ such that

$$P_1 f_1 + P_2 f_2 + \cdots + P_n f_n = O(z^{N(n-\epsilon)}).$$

Siegel construction

Fix N sufficiently large and choose $\epsilon > 1$. Using linear algebra and Siegel's Lemma, construct polynomials $P_1, \ldots, P_n \in K[z]$ of degrees $\leq N$ such that

$$P_1 f_1 + P_2 f_2 + \cdots + P_n f_n = O(z^{N(n-\epsilon)}).$$

Rewrite in vector notation

$$\mathbf{P} \cdot \mathbf{f} = O(z^{N(n-\epsilon)}).$$

Siegel construction

Fix N sufficiently large and choose $\epsilon > 1$. Using linear algebra and Siegel's Lemma, construct polynomials $P_1, \ldots, P_n \in K[z]$ of degrees $\leq N$ such that

$$P_1 f_1 + P_2 f_2 + \cdots + P_n f_n = O(z^{N(n-\epsilon)}).$$

Rewrite in vector notation

$$\mathbf{P} \cdot \mathbf{f} = O(z^{N(n-\epsilon)}).$$

Differentiate,

$$\mathbf{P}' \cdot \mathbf{f} + \mathbf{P} \cdot A\mathbf{f} = O(z^{N(n-\epsilon)-1}).$$

Siegel construction

Fix N sufficiently large and choose $\epsilon > 1$. Using linear algebra and Siegel's Lemma, construct polynomials $P_1, \ldots, P_n \in K[z]$ of degrees $\leq N$ such that

$$P_1 f_1 + P_2 f_2 + \cdots + P_n f_n = O(z^{N(n-\epsilon)}).$$

Rewrite in vector notation

$$\mathbf{P} \cdot \mathbf{f} = O(z^{N(n-\epsilon)}).$$

Differentiate,

$$\mathbf{P}' \cdot \mathbf{f} + \mathbf{P} \cdot A\mathbf{f} = O(z^{N(n-\epsilon)-1}).$$

Denote derivation by D, then we get

$$((D + A^t)\mathbf{P}) \cdot \mathbf{f} = O(z^{N(n-\epsilon)-1}).$$

Siegel construction

Fix N sufficiently large and choose $\epsilon > 1$. Using linear algebra and Siegel's Lemma, construct polynomials $P_1, \ldots, P_n \in K[z]$ of degrees $\leq N$ such that

$$P_1 f_1 + P_2 f_2 + \cdots + P_n f_n = O(z^{N(n-\epsilon)}).$$

Rewrite in vector notation

$$\mathbf{P} \cdot \mathbf{f} = O(z^{N(n-\epsilon)}).$$

Differentiate,

$$\mathbf{P}' \cdot \mathbf{f} + \mathbf{P} \cdot A\mathbf{f} = O(z^{N(n-\epsilon)-1}).$$

Denote derivation by D, then we get

$$((D + A^t)\mathbf{P}) \cdot \mathbf{f} = O(z^{N(n-\epsilon)-1}).$$

The coefficients of new polynomial relation are then given by $T(z)(D + A^t)\mathbf{P}$.

Shidlovskii's idea

We need n independent approximations and hope that they can be found among

$$\mathbf{P},\ (D+A^t)\mathbf{P},\ (D+A^t)^2\mathbf{P}\ ,\ldots,\ (D+A^t)^{n-1}\mathbf{P}.$$

Denote the determinant of the $n \times n$-matrix formed by these vectors by $\Delta(z)$.

Shidlovskii's idea

We need n independent approximations and hope that they can be found among

$$\mathbf{P},\ (D+A^t)\mathbf{P},\ (D+A^t)^2\mathbf{P}\ ,\ldots,\ (D+A^t)^{n-1}\mathbf{P}.$$

Denote the determinant of the $n \times n$-matrix formed by these vectors by $\Delta(z)$.
We want $\Delta(z) \not\equiv 0$.

Shidlovskii's idea

We need n independent approximations and hope that they can be found among

$$\mathbf{P},\ (D+A^t)\mathbf{P},\ (D+A^t)^2\mathbf{P}\ ,\dots,\ (D+A^t)^{n-1}\mathbf{P}.$$

Denote the determinant of the $n \times n$-matrix formed by these vectors by $\Delta(z)$.
We want $\Delta(z) \not\equiv 0$.

Shidlovskii's Lemma

Suppose $\epsilon < 1$. Then $\Delta(z) \not\equiv 0$.

Last skirmishes

We specialise the above approximations by setting $z = \alpha$ to obtain the desired extra inequalities

$$|B_{i1} f_1(\alpha) + \cdots + B_{in} f_n(\alpha)| < \delta.$$

Last skirmishes

We specialise the above approximations by setting $z = \alpha$ to obtain the desired extra inequalities

$$|B_{i1}f_1(\alpha) + \cdots + B_{in}f_n(\alpha)| < \delta.$$

Proposition

There exists $\gamma > 0$,independent of N or ϵ such that the rank of the vectors

$$((D + A^t)^r \mathbf{P})(\alpha), \ \ r = 0, 1, \ldots, N\epsilon + \gamma$$

is precisely n.