Tsinghua Lectures on Hypergeometric Functions (Unfinished and comments are welcome)

Gert Heckman
Radboud University Nijmegen
G.Heckman@math.ru.nl

December 8, 2015

Contents

Preface

The Euler–Gauss hypergeometric function

$$F(\alpha,\beta,\gamma;z)=\sum_{k=0}^{\infty}\frac{\alpha(\alpha+1)\cdots(\alpha+k-1)\beta(\beta+1)\cdots(\beta+k-1)}{\gamma(\gamma+1)\cdots(\gamma+k-1)k!}z^k$$

was introduced by Euler in the 18^{th} century, and was well studied in the 19^{th} century among others by Gauss, Riemann, Schwarz and Klein. The numbers α,β,γ are called the parameters, and z is called the variable.

On the one hand, for particular values of the parameters this function appears in various problems. For example

$$\begin{aligned}
(1-z)^{-\alpha} &= F(\alpha,1,1;z)\\
\arcsin z &= 2zF(1/2,1,3/2;z^2)\\
K(z) &= \frac{\pi}{2}F(1/2,1/2,1;z^2)\\
P_n^{(\alpha,\beta)}(z) &= \frac{\alpha(\alpha+1)\cdots(\alpha+n)}{n!}F(-n,\alpha+\beta+n+1;\alpha+1|\frac{1-z}{2})
\end{aligned}$$

with $K(z)$ the Jacobi elliptic integral of the first kind given by

$$K(z)=\int_0^1\frac{dx}{\sqrt{(1-x^2)(1-z^2x^2)}}\ ,$$

and $P_n^{(\alpha,\beta)}(z)$ the Jacobi polynomial of degree n, normalized by

$$P_n^{(\alpha,\beta)}(1)=\binom{\alpha+n}{n}\ .$$

On the other hand, the hypergeometric differential equation (of which $F(\alpha,\beta;\gamma|z)$ is a solution) served as a guiding example for the general theory of ordinary differential equations in a complex domain. For example, the calculation of the monodromy of the hypergeometric equation led Riemann to formulate the so called Riemann–Hilbert problem, later reformulated by Hilbert in his famous list of 1900 as Problem 21.

A natural generalization of the Euler–Gauss hypergeometric function is the Clausen–Thomae hypergeometric function

$$F(\alpha;\beta|z)=\sum_{k=0}^{\infty}\frac{(\alpha_1)_k(\alpha_2)_k\cdots(\alpha_n)_k}{(\beta_1)_k(\beta_2)_k\cdots(\beta_n)_k}z^k$$

with $(\lambda)_k = \Gamma(\lambda+k)/\Gamma(\lambda) = \lambda(\lambda+1)\cdots(\lambda+k-1)$ the so called Pochhammer symbol. The numbers $\alpha = (\alpha_1, \cdots, \alpha_n)$ are called the numerator parameters and $\beta = (\beta_1, \cdots, \beta_n)$ the denominator parameters. Usually $\beta_n = 1$ so that $(\beta_n)_k = k!$ and therefore the Euler–Gauss hypergeometric function has numerator parameters $(\alpha_1, \alpha_2) = (\alpha, \beta)$ and denominator parameter $\beta_1 = \gamma$. Many (but not all) results of the Euler–Gauss hypergeometric function can be generalized for the Clausen–Thomae hypergeometric function. For very particular values of the parameters the Clausen–Thomae hypergeometric function appeared in modern mathematics in the context of mirror symmetry for Calabi–Yau threefolds.

After a fairly detailed treatment of these two classical hypergeometric functions of the 19^{th} century we discuss a multivariable analogue of the Euler–Gauss hypergeometric function: the hypergeometric function

$$F(\lambda, k; t)$$

associated with a root system R. These functions generalize the Euler–Gauss hypergeometric function (for the rank one root system) and the elementary spherical functions on a real semisimple Lie group (for particular parameter values). They were introduced and studied in a collaboration of Eric Opdam and the author of these lectures in the eighties and nineties of the last century. These special functions are intimately connected with the Calogero–Moser system of n points on a circle, under influence of an inverse square potential. The classical integrability of this system was conjectured by Calogero and proved by Moser. The root system hypergeometric functions appear as the simultaneous eigenfunctions of the Schrödinger operator and its conserved operators for the quantum integrable system. In order to make these lecture notes self contained the basic properties of root systems and Weyl groups are included.

Some time ago I read a nice paper by Dyson entitled ”Birds and Frogs” [12]. He discusses vividly the two extreme archetypes of mathematicians. On the one hand there are the birds. Like eagles they fly high up in the air and have a magnificient view of the mathematical landscape. They see the great analogies in mathematics for example between geometry and number theory or geometry and mathematical physics. Examples of birds are Atiyah, Grothendieck, Harish–Chandra, Langlands and Yau. On the other hand there are the frogs. They live down in the mud, and are eager to spot some precious stone hidden under the dirt that the birds might miss. Examples of frogs are Coxeter, Dyson (in his own opinion), Macdonald and Selberg. Some truely great mathematicians like Deligne, Mumford and Serre unite both aspects. In all modesty and without any comparison with these great mathematicians I am a frog at heart.

These notes are written for a series of lectures at the Tsinghua University of Beijing in the fall of 2011. Most likely they should appeal to an audience of frogs. I would like to thank the Mathematical Sciences Center for their hospitality. In particular I am greatful to Professors Yau and Poon for the invitation to come to Tsinghua. In addition I would like to thank the students for their questions and patience. Finally I want to thank Miss Li Fei for helping me around in China, and Elisabeth Giljam and Eduard Looijenga for making my stay at Tsinghua in many aspects a wonderful experience.

1 Linear differential equations

1.1 The local existence problem

We shall write $\partial = d/dz$ where z is the standard coordinate in the complex plane. Let us consider the two linear ordinary differential equations

$$(\partial^n + a_1\partial^{n-1} + \cdots + a_{n-1}\partial + a_n)f = 0$$

$$(\partial + A)F = 0$$

with scalar coefficients $a_1, \cdots, a_n$ and matrix coefficient $A = (a_{ij})_{1\leq i,j\leq n}$ holomorphic on some domain $Z \subset \mathbb{C}$.

The first linear differential equation is a scalar equation of order n: the coefficients $a_1(z), \cdots, a_n(z)$ are holomorphic functions on Z, and we shall seek holomorphic solutions $f(z)$ on suitable open subsets of Z. The second linear differential equation is a first order matrix equation: the entries $a_{ij}(z)$ of the matrix $A(z)$ are holomorphic functions on Z, and we shall seek vector valued holomorphic solutions

$$F(z) = (f_1(z), \cdots, f_n(z))^t$$

on suitable open subsets of Z. The (local) existence problem of higher order scalar equations can be reduced to the (local) existence problem of first order matrix equations.

Theorem 1.1. *Suppose holomorphic functions $a_1(z), \cdots, a_n(z)$ have been given on a domain $Z \subset \mathbb{C}$. Define the matrix valued holomorphic function $A(z)$ on Z by*

$$\begin{pmatrix} 0 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 \\ a_n & a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_2 & a_1 \end{pmatrix}$$

If the vector valued function $F(z) = (f_1(z), \cdots, f_n(z))^t$ is a solution of

$$(\partial + A)F = 0$$

then $f(z) = f_1(z)$ is a solution of

$$(\partial^n + a_1\partial^{n-1} + \cdots + a_{n-1}\partial + a_n)f = 0$$

and $f_{j+1}(z) = \partial f_j(z)$ for $j = 1, \cdots, n-1$.

Proof. With the matrix valued function $A(z)$ as above and the vector valued function $F(z) = (f_1(z), \cdots, f_n(z))^t$ the equation $(\partial + A)F = 0$ amounts to

$$\begin{pmatrix} \partial f_1 \\ \partial f_2 \\ \vdots \\ \partial f_{n-1} \\ \partial f_n \end{pmatrix} + \begin{pmatrix} -f_2 \\ -f_3 \\ \vdots \\ -f_n \\ a_n f_1 + \cdots + a_1 f_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

which in turn is equivalent to the equations

$$f_1 = f, f_2 = \partial f, \cdots, f_n = \partial^{n-1} f$$

$$(\partial^n + a_1 \partial^{n-1} + \cdots + a_{n-1}\partial + a_n) f = 0$$

which proves the theorem. □

For $r > 0$ let $D_r = \{z \in \mathbb{C}; |z| < r\}$ be the open disc around $z = 0$ with radius r. In this section we shall carry out a local analysis for the domain $Z = D_r$. Consider the matrix equation

$$(\partial + A)F = 0$$

with $A = (a_{ij})$ and $a_{ij} = a_{ij}(z)$ holomorphic on D_r. Do there exist solutions $F = (f_1, \cdots, f_n)^t$ with $f_j = f_j(z)$ holomorphic on D_r, and if yes how many? For this purpose develop $A(z)$ in a power series

$$A(z) = \sum_0^\infty A_k z^k$$

and substitute a formal power series

$$F(z) = \sum_0^\infty F_k z^k$$

with $F_k \in \mathbb{C}^n$ undetermined coefficients.

Proposition 1.2. *The formal power series* $F(z) = \sum_0^\infty F_k z^k$ *is a formal solution of* $(\partial + A)F = 0$ *with* $A(z) = \sum_0^\infty A_k z^k$ *if and only if*

$$(k+1)F_{k+1} + \sum_{l=0}^{k} A_{k-l} F_l = 0$$

for all $k \in \mathbb{N} = \{0, 1, 2, \cdots\}$.

Proof. Substitution in $(\partial + A)F = 0$ gives

$$\sum_{k=1}^{\infty} kF_k z^{k-1} + \sum_{k,l\geq 0} A_k F_l z^{k+l} =$$

$$\sum_{k=0}^{\infty}((k+1)F_{k+1} + \sum_{l=0}^{k} A_{k-l}F_l)z^k = 0$$

and the proposition is clear. ∎

On the level of formal power series we get:

On the level of formal power series we get:
$F_0 \in \mathbb{C}^n$ is undetermined and can be freely choosen,
$F_1 = -A_0F_0$,
$F_2 = -(A_1F_0 + A_0F_1)/2 = (A_0^2 - A_1)F_0/2$,
$F_3 = -(A_2F_0 + A_1F_1 + A_0F_2)/3 = (-A_0^3 + A_0A_1 + 2A_1A_0 - 2A_2)F_0/6$,
$\cdots$

So given F_0 the F_k with $k \geq 1$ can be explicitly computed via the recurrence relation. Using Theorem 1.1 we obtain the following result.

Corollary 1.3. *The n^{th} order scalar equation*

$$(\partial^n + a_1\partial^{n-1} + \cdots + a_{n-1}\partial + a_n)f = 0$$

on the disc D_r has a unique formal power series solution $f(z) = \sum f_k z^k$ for freely chosen $f_0, \cdots, f_{n-1} \in \mathbb{C}$.

Our next goal is to show that these formal power series solutions are in fact convergent power series. The following lemma is familiar from complex function theory.

Lemma 1.4. *A formal power series $\sum a_k z^k$ (with coefficients in a Banach space) is convergent on the disc D_r with radius $r > 0$ if and only if for each $\rho \in (0,1)$ there exists a constant $M_\rho \geq 0$ such that*

$$|a_k| \leq M_\rho(\rho r)^{-k}$$

for all $k \in \mathbb{N}$.

Theorem 1.5. *If the coefficients of the matrix equation $(\partial + A)F = 0$ are convergent on the disc D_r then the formal power series solution $\sum F_k z^k$ with F_0 undetermined and F_{k+1} given by the recurrence relation of Proposition 1.2 also converges on the disc D_r.*

Proof. The power series $\sum A_k z^k$ converges on D_r, and therefore we have an estimate of the form (switch from M_ρ to $M_\rho(\rho r)^{-1}$)

$$\forall \rho \in (0,1)\ \exists M_\rho \geq 0\ ;\ |A_k| \leq M_\rho(\rho r)^{-k-1}\ \forall k \in \mathbb{N}.$$

We claim that this implies an estimate for F_k of the form

$$|F_k| \leq M_\rho(M_\rho+1)\cdots(M_\rho+k-1)(\rho r)^{-k}|F_0|/k!\ \forall k \in \mathbb{N}.$$

We prove this claim by induction on $k \in \mathbb{N}$. The case $k=0$ is trivial. Using the recurrence relation and the induction hypothesis we get

$$|F_{k+1}| \leq (k+1)^{-1}\{\sum_{l=0}^{k} M_\rho(\rho r)^{-k+l-1} \cdot \frac{\Gamma(M_\rho+l)}{\Gamma(M_\rho)l!}(\rho r)^{-l}|F_0|\}.$$

Using the formula (easily proved by induction on k)

$$\sum_{l=0}^{k} \frac{\Gamma(M+l)}{\Gamma(M)l!} = \frac{\Gamma(M+k+1)}{\Gamma(M+1)k!}$$

we arrive at the estimate

$$\begin{aligned} |F_{k+1}| &\leq \frac{1}{(k+1)} \cdot \frac{M_\rho\Gamma(M_\rho+k+1)}{\Gamma(M_\rho+1)k!}(\rho r)^{-k-1}|F_0| \\ &= \frac{\Gamma(M_\rho+k+1)}{\Gamma(M_\rho)(k+1)!}(\rho r)^{-(k+1)}|F_0| \end{aligned}$$

which proves our claim.

For each $M \geq 0$ (even for $M \in \mathbb{C}$) the binomial series

$$(1-z)^{-M} = \sum_{0}^{\infty} \frac{\Gamma(M+k)}{\Gamma(M)k!} z^k$$

is convergent on the unit disc D_1. Hence by Lemma 1.4 we get the estimate

$$\forall \sigma \in (0,1)\ \exists N_{\rho,\sigma} > 0\ ;\ \frac{\Gamma(M_\rho+k)}{\Gamma(M_\rho)k!} \leq N_{\rho,\sigma}\sigma^{-k}\ \forall k \in \mathbb{N}.$$

So finally we arrive at

$$\forall \rho, \sigma \in (0,1)\ \exists N_{\rho,\sigma} > 0\ ;\ |F_k| \leq |F_0|N_{\rho,\sigma}(\rho\sigma r)^{-k}\ \forall k \in \mathbb{N},$$

which in turn implies

$$\forall \rho \in (0,1)\ \exists L_\rho > 0\ ;\ |F_k| \leq L_\rho(\rho r)^{-k}\ \forall k \in \mathbb{N}.$$

Indeed just take $L_\rho = |F_0|N_{\sqrt{\rho},\sqrt{\rho}}$. Now apply Lemma 1.4. ☐

Corollary 1.6. *Let* $(\partial + A)F = 0$ *be a first order matrix equation with coefficients* $A(z) = (a_{ij}(z))_{1\le i,j\le n}$ *holomorphic in a domain* $Z \subset \mathbb{C}$. *For every point* $z_0 \in Z$ *and* $F_0 \in \mathbb{C}^n$ *there is a unique local holomorphic solution* $F(z)$ *around* z_0 *with initial value* $F(z_0) = F_0$.

Corollary 1.7. *Let* $(\partial^n + a_1\partial^{n-1} + \cdots + a_{n-1}\partial + a_n)f = 0$ *be an* n^{th} *order scalar equation with coefficients* $a_1(z), \cdots, a_n(z)$ *holomorphic in a domain* $Z \subset \mathbb{C}$. *For every point* $z_0 \in Z$ *and complex numbers* $f_0, \cdots, f_{n-1}$ *there is a unique local holomorhic solution* $f(z)$ *around* z_0 *with initial conditions*

$$f(z_0) = f_0, \partial f(z_0) = f_1, \cdots, \partial^{n-1} f(z_0) = f_{n-1}.$$

Example 1.8. *The second order differential equation* $(\partial^2 + (1/z)\partial)f = 0$ *on the punctured complex plane* $\mathbb{C}^\times$ *has*

$$f(z) = \log z = \log(1 + (z-1)) = (z-1) - (z-1)^2/2 + (z-1)^3/3 + \cdots$$

as unique local holomorphic solution around $z = 1$ *with* $f(1) = 0, \partial f(1) = 1$. *The differential equation provides the analytic continuation*

$$\log z = \int_1^z \frac{d\zeta}{\zeta}$$

with the line integral taken along a curve in $\mathbb{C}^\times$ *from* 1 *to* z.

Remark 1.9. *Suppose that the coefficients of the linear differential equation* $(\partial + A)F = 0$ *in a domain* Z *also depend in a holomorphic way on a complex parameter* α, *so* $A = A(\alpha, z)$ *with* α *a parameter and* z *the variable of the differential equation, so* $\partial = d/dz$. *Suppose that the power series*

$$A(\alpha, z) = \sum_0^\infty A_k(\alpha) z^k$$

converges on D_r *in a locally uniform way in* α, *so the constant* $M_\rho = M_\rho(\alpha)$ *in Lemma 1.4 is locally independent of* α. *The estimates in Theorem 1.5 are also locally uniform in* α, *so the power series* $\sum F_k(\alpha) z^k$ *on* D_r *converges in a locally uniform way in* α. *Hence for an initial value* $F_0(\alpha) \in \mathbb{C}^n$ *that is holomorphic in* α *the unique solution* $F(\alpha, z)$ *of* $(\partial + A)F = 0$ *with initial value* $F(\alpha, 0) = F_0(\alpha)$ *depends also in a holomorphic way on* α.

1.2 The fundamental group

Let Z be a connected topological space. The example to have in mind is a domain Z in $\mathbb{C}$.

Definition 1.10. *A path in Z is a continuous map $\gamma : [0,1] \to Z, t \mapsto \gamma(t)$. The point $\gamma(0)$ is called the begin point and the point $\gamma(1)$ the end point of γ. If begin and end point of γ coincide then γ is called a loop with base point $\gamma(0) = \gamma(1)$.*

Definition 1.11. *Let γ_1 and γ_2 be two paths in Z with equal begin points $\gamma_1(0) = \gamma_2(0)$ and equal end points $\gamma_1(1) = \gamma_2(1)$. The paths γ_1 and γ_2 are called homotopic if there exists a continuous map $h : [0,1] \times [0,1] \to Z, (s,t) \mapsto h(s,t)$ such that*

$$h(0,t) = \gamma_1(t) \ , \ h(1,t) = \gamma_2(t) \ \forall t \in [0,1] \ ,$$

$$h(s,0) = \gamma_1(0) = \gamma_2(0) \ , \ h(s,1) = \gamma_1(1) = \gamma_2(2) \ \forall s \in [0,1] \ .$$

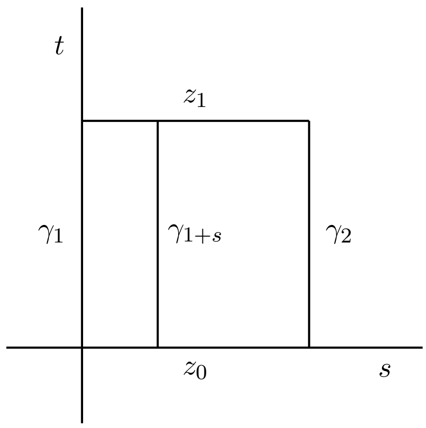
The map h is called the homotopy between the paths γ_1 and γ_2.

In other words the two paths γ_1 and γ_2 are homotopic if there exists a one parameter continuous family (with parameter $s \in [0,1]$) of paths

$$\gamma_{1+s} : [0,1] \to Z$$

$$\gamma_{1+s}(0) = \gamma_1(0) = \gamma_2(0) \ , \ \gamma_{1+s}(1) = \gamma_1(1) = \gamma_2(1) \ \forall s \in [0,1] \ .$$

The link with our previous notation is $\gamma_{1+s}(t) = h(s,t)$. If $\gamma_1(0) = \gamma_2(0) = z_0$ and $\gamma_1(1) = \gamma_2(1) = z_1$ then we draw the following schematic picture.



We shall write $\gamma_1 \sim \gamma_2$ if the paths γ_1 and γ_2 in Z with equal begin points and equal end points are homotopic. It is easy to show that being homotopic is an equivalence relation. The equivalence class of a path $\gamma : [0,1] \to Z$ is denoted by $[\gamma]$.

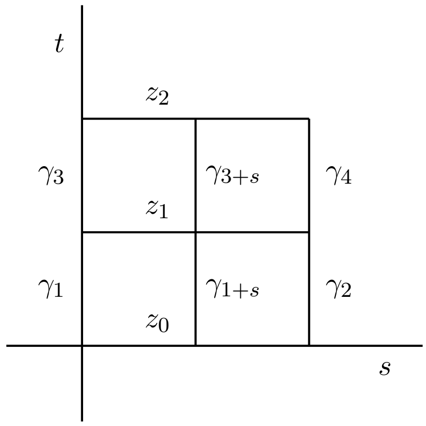
Definition 1.12. *Let $z_0, z_1, z_2 \in Z$ be three points and let $\gamma_1, \gamma_2 : [0,1] \to Z$ be two paths in Z with $\gamma_1(0) = z_0, \gamma_1(1) = \gamma_2(0) = z_1, \gamma_2(1) = z_2$. We define a new path $\gamma_2\gamma_1 : [0,1] \to Z$ by*

$$\gamma_2\gamma_1(t) = \gamma_1(2t) \; \forall t \in [0, 1/2] \; ,$$

$$\gamma_2\gamma_1(t) = \gamma_2(2t-1) \; \forall t \in [1/2, 1] \; .$$

The path $\gamma_2\gamma_1$ is called the product of γ_2 and γ_1, and is always taken in the order: start with γ_1 and then follow with γ_2.

It is easy to show that if $\gamma_1 \sim \gamma_2$ and $\gamma_3 \sim \gamma_4$ and the end point z_1 of γ_1, γ_2 coincides with the begin point z_1 of γ_3, γ_4 then $\gamma_3\gamma_1 \sim \gamma_4\gamma_2$. Here is a schematic picture of the homotopy.



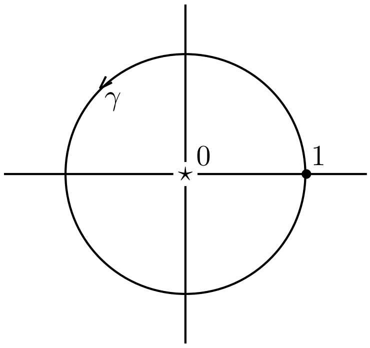
Hence the product $[\gamma_2][\gamma_1]$ of the homotopy classes of paths γ_2 and γ_1 as in Definition 1.12 is well defined. We leave it as an exercise to show that the product of paths is associative on homotopy classes of paths.

Theorem 1.13. *For $z_0 \in Z$ a fixed point let $\Pi_1(Z, z_0)$ denote the collection of homotopy classes of loops in Z with base point (i.e. begin and end point) z_0. The product rule on paths in Z according to Definition 1.12 defines a group structure on $\Pi_1(Z, z_0)$. The unit element is represented by the constant path $\epsilon(t) = z_0 \; \forall t \in [0,1]$ at z_0. The inverse $[\gamma]^{-1}$ of $[\gamma] \in \Pi_1(Z, z_0)$ is represented by the loop $\gamma^{-1}(t) = \gamma(1-t) \; \forall t \in [0,1]$.*

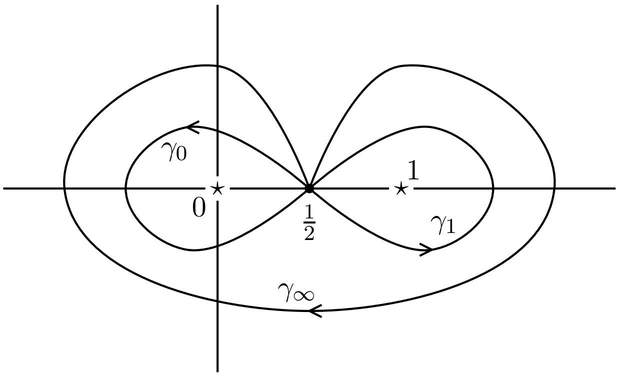
Definition 1.14. *The group $\Pi_1(Z, z_0)$ is called the fundamental group of the connected topological space Z with base point z_0.*

Elements of $\Pi_1(Z, z_0)$ are homotopy classes of loops, but sometimes one refers to the elements of $\Pi_1(Z, z_0)$ simply as loops (based at z_0). Even worse, in the notation one simply writes $\gamma \in \Pi_1(Z, z_0)$ rather than $[\gamma] \in \Pi_1(Z, z_0)$.

Example 1.15. *Let $Z = \mathbb{C}^\times = \mathbb{C} - \{0\}$ and $z_0 = 1$. Let $\gamma(t) = exp(2\pi i t)$ for $t \in [0, 1]$. Then $\Pi_1(Z, z_0)$ is a cyclic group with generator γ.*



Example 1.16. *Let $Z = \mathbb{P} - \{0, 1, \infty\} = \mathbb{C} - \{0, 1\}$ with $\mathbb{P} = \mathbb{C} \cup \{\infty\}$ the projective line and take $z_0 = 1/2$. Choose loops $\gamma_0, \gamma_1, \gamma_\infty$ around the points $0, 1, \infty$ respectively as in the picture.*



It is easy to see that $\gamma_\infty \gamma_1 \gamma_0 = 1$ in $\Pi_1(Z, 1/2)$. It turns out that $\Pi_1(Z, 1/2)$ is isomorphic to the group on three generators $\gamma_0, \gamma_1, \gamma_\infty$ with the single relation $\gamma_\infty \gamma_1 \gamma_0 = 1$.

Remark 1.17. *For two points $z_0, z_1 \in Z$ the choice of a path δ from z_0 to z_1 gives an isomorphism $\phi_{[\delta]} : \Pi_1(Z, z_0) \to \Pi_1(Z, z_1)$ by $\phi_{[\delta]}([\gamma]) = [\delta][\gamma][\delta]^{-1}$.*

Since $\phi_{[\delta_2]}([\gamma]) = [\delta_2\delta_1^{-1}]\phi_{[\delta_1]}([\gamma])[\delta_1\delta_2^{-1}]$ *the base point free fundamental group* $\Pi_1(Z)$ *as a group is only defined up to inner automorphisms. In turn the Abelianized fundamental group*

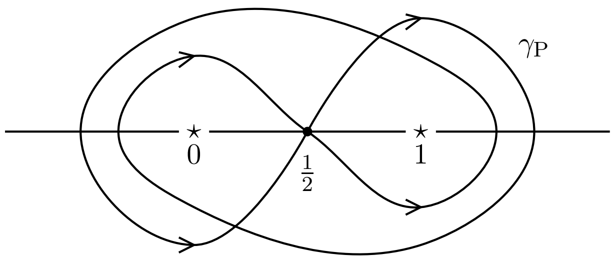
$$\Pi_1(Z)^{\mathrm{Abel}} = \Pi_1(Z)/[\Pi_1(Z), \Pi_1(Z)]$$

is a canonically defined Abelian group, and is called the first homology group of the space Z*, denoted* $H_1(Z)$*. By abuse of notation the class* $[\gamma] \in H_1(Z)$ *is called the cycle of the loop* $[\gamma] \in \Pi_1(Z, z_0)$*.*

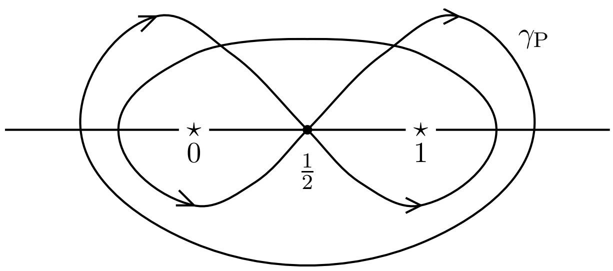
Example 1.18. *The Pochhammer contour* $\gamma_{\mathrm{P}} \in \Pi_1(\mathbb{P} - \{0, 1, \infty\}, \frac{1}{2})$ *is defined by*

$$\gamma_{\mathrm{P}} = [\gamma_0, \gamma_1] = \gamma_0\gamma_1\gamma_\infty$$

with $[\cdot, \cdot]$ *for the commutator in the fundamental group* $\Pi_1(\mathbb{P} - \{0, 1, \infty\}, \frac{1}{2})$*. The second equality follows from the topological relation* $\gamma_\infty\gamma_1\gamma_0 = 1$*. One can draw two pictures of the Pochhammer contour both exhibiting reflection symmetry. One is the holomorphic symmetry* $z \mapsto (1 - z)$ *of point reflection in* $z = \frac{1}{2}$



and the other is the antiholomorphic symmetry $z \mapsto (1 - \overline{z})$ *of line reflection in* $\Re z = \frac{1}{2}$



The Pochhammer contour was introduced independently by Jordan in 1887 *and Pochhammer in* 1890 *[22], [29]. The Pochammer contour is nontrivial in homotopy, but the associated cycle is trivial in homology.*

1.3 The monodromy representation

Suppose G is a group and V is a finite dimensional vector space over the complex numbers $\mathbb{C}$. Let $\mathrm{GL}(V)$ denote the group of invertible linear operators on V. A representation (π, V) of G on V is a homomorphism $\pi : G \to \mathrm{GL}(V)$. If (π_1, V_1) and (π_2, V_2) are two representations of a group G, then a linear map $A \in \mathrm{Hom}(V_1, V_2)$ is called an intertwiner from (π_1, V_1) to (π_2, V_2) if

$$A\pi_1(g) = \pi_2(g)A \ \forall g \in G \ .$$

The intertwiners from (π_1, V_1) to (π_2, V_2) form a linear subspace of the vector space $\mathrm{Hom}(V_1, V_2)$ of linear maps from V_1 to V_2, denoted by $\mathrm{Hom}(V_1, V_2)^G$. A bijective intertwiner $A \in \mathrm{Hom}(V_1, V_2)^G$ is called an equivalence between (π_1, V_1) and (π_2, V_2). It is easy to check that equivalence of representations of a group G is an equivalence relation on the set of representations (π, V) of G.

Given a representation (π, V) of G a linear subspace $U \subset V$ is called invariant if $\pi(g)u \in U \ \forall g \in G \ \forall u \in U$. In this case we denote $\pi_U(g) = \pi(g)|_U$ and call (π_U, U) a subrepresentation of (π, V). A representation (π, V) of G is called irreducible if the only invariant linear subspaces of V are the trivial ones 0 and V. Given a representation (π, V) of G a Hermitian form $\langle \cdot, \cdot \rangle$ on V (which by definition is linear in the first argument, and antilinear in the second argument) is called invariant if

$$\langle \pi(g)u, \pi(g)v \rangle = \langle u, v \rangle \ \forall u, v \in V \ .$$

The kernel of an invariant Hermitian form is easily seen to be an invariant linear subspace. In particular a nonzero invariant Hermitian form on an irreducible representation space V is nondegenerate. A representation (π, V) of G is called unitary if there exists a positive definite invariant Hermitian form on V.

Suppose $Z \subset \mathbb{C}$ is a domain. Suppose we are given an n^{th} order linear differential equation

$$(\partial^n + a_1\partial^{n-1} + \cdots + a_{n-1}\partial + a_n)f = 0$$

with coefficients $a_1, \cdots, a_n$ holomorhic in Z. Fix a base point $z_0 \in Z$ and let V_0 be the linear space of local holomorphic solutions around z_0. We know that the dimension of V_0 is equal to n. Suppose γ is a path in Z with begin point z_0 and end point z_1, and let V_1 be the linear space of local holomorphic solutions around z_1. Analytic continuation of local solutions along γ depends only on the homotopy class $[\gamma]$ of γ in Z. Therefore we have defined a linear monodromy operator

$$M([\gamma]) : V_0 \to V_1 .$$

The monodromy operator corresponding to the product of two paths is clearly equal to the product of the two monodromy operators corresponding to the individual paths. Restricting to loops in Z with base point z_0 therefore defines the monodromy representation

$$M : \Pi_1(Z, z_0) \to \mathrm{GL}(V_0) .$$

If $f_1, \cdots, f_n$ is a basis of the local solution space V_0 around z_0 then as usual the monodromy matrix in this basis is defined by $M([\gamma])f_j = \sum m_{ij}([\gamma])f_i$ and $\Pi_1(Z, z_0) \ni [\gamma] \mapsto m_{ij}([\gamma]) \in \mathrm{GL}(n, \mathbb{C})$ is the corresponding monodromy matrix representation. The monodromy representation is a powerful (in general transcendental) invariant of a linear differential equation, and was introduced by Riemann.

For a first order matrix linear differential equation

$$(\partial + A)F = 0$$

of size n by n in a domain Z the monodromy is defined likewise. If γ is a path in Z with begin point z_0 and end point z_1 and F is a local solution around z_0 then $M([\gamma])F$ is a local solution around z_1, obtained by analytic continuation of F along $[\gamma]$. If V_0 is the local solution space of $(\partial + A)F = 0$ around z_0 then we get a mondromy representation $M : \Pi_1(Z, z_0) \to \mathrm{GL}(V_0)$ and after a choice of basis $F_1, \cdots, F_n$ of V_0 we get the associated monodromy matrix representation $\Pi_1(Z, z_0) \ni [\gamma] \mapsto m_{ij}([\gamma]) \in \mathrm{GL}(n, \mathbb{C})$ with $M([\gamma])F_j = \sum m_{ij}([\gamma])F_i$. If the scalar and matrix equations

$$(\partial^n + a_1\partial^{n-1} + \cdots + a_{n-1}\partial + a_n)f = 0$$

$$(\partial + A)F = 0$$

are related by Theorem 1.1, so that

$$F(z) = (f_1(z), \cdots, f_n(z))^t \ , \ f(z) = f_1(z) \ , \ f_{j+1}(z) = \partial f_j(z)$$

for $j = 1, \cdots, n-1$, then $M([\gamma])F$ corresponds likewise to $M([\gamma])f$. Indeed the operator ∂ commutes with monodromy.

1.4 Regular singular points

Suppose $Z \subset \mathbb{C}$ is a domain with base point z_0. Consider the linear system of differential equations

$$(\partial + A)F = 0$$

with $A = (a_{ij})_{1\le i,j\le n}$ and $a_{ij} = a_{ij}(z)$ holomorphic functions on Z. Suppose we choose a basis $F_1, \cdots, F_n$ of the local solution space V_0 around z_0. If we write $F_j = (f_{1j}, \cdots, f_{nj})^t$ then $F = (f_{ij})$ is called a local solution matrix around z_0. Let $\gamma \in \Pi_1(Z, z_0)$ be a loop in Z based at z_0 and let $M = M([\gamma]) : V_0 \to V_0$ be the corresponding monodromy operator. The monodromy matrix $(m_{jk})_{1\le j,k\le n}$ with respect to the basis $F_1, \cdots, F_n$ is defined by the usual relations

$$M(F_k) = \sum m_{jk} F_j \ .$$

Under the monodromy operator M the matrix entry f_{ik} of the local solution matrix $F = (F_1, \cdots, F_n)$ transforms into $M(f_{ik}) = \sum m_{jk} f_{ij}$. In other words we get

$$M(F) = FM$$

so that the monodromy operator M acts on the local solution matrix (f_{ij}) by multiplication on the right with the monodromy matrix (m_{jk}).

Example 1.19. *Let $\theta = z\partial = z\frac{d}{dz}$ and consider the linear system*

$$(\partial + B/z)F = 0 \ \Leftrightarrow (\theta + B)F = 0$$

with $B = (b_{ij})_{1\le i,j\le n} \in \mathrm{Mat}(n, \mathbb{C})$ a scalar matrix. For the domain Z we take $\mathbb{C}^\times$, say with base point $z_0 = 1$. As local solution matrix around $z_0 = 1$ we can take

$$F(z) = z^{-B} = e^{-B\log z}$$

which defines a single valued solution matrix on $\mathbb{C} - (-\infty, 0]$ by taking the branch $\log 1 = 0$. If $\gamma(t) = e^{2\pi i t}$ for $t \in [0,1]$ then the monodromy operator $M = M([\gamma])$ has monodromy matrix

$$M = e^{-2\pi i B} \ .$$

Let us now take for the domain Z the punctured disc

$$D^\times_{r_0} = \{z \in \mathbb{C}; 0 < |z| < r_0\}$$

for some $r_0 > 0$ with base point $z_0 = r_0/2$, and consider the linear system $(\partial + A)F = 0$ with coefficients holomorphic on $D^\times_{r_0}$. Let $F = (F_1, \cdots, F_n) = (f_{ij})$ be a local solution matrix around $z_0 = r_0/2$. Let γ be the loop $\gamma(t) = r_0 e^{2\pi i t}/2$ and let M be the matrix of the monodromy operator $M([\gamma])$.

Proposition 1.20. *The exponential map* $\exp : \mathrm{Mat}(n, \mathbb{C}) \to \mathrm{GL}(n, \mathbb{C})$ *is a surjection.*

Proof. This follows from the Jordan decomposition. □

Choose $B \in \mathrm{Mat}(n, \mathbb{C})$ with $M = e^{-2\pi i B}$. The matrix B is not unique, but can be fixed uniquely by the requirement $0 \le \Re\lambda < 1$ for each eigenvalue λ of B. Consider the matrix valued holomorphic function

$$G(z) = F(z)z^B = F(z)e^{B\log z}$$

around $z_0 = r_0/2$ with the branch fixed by $\log(r_0/2) > 0$. Because $F(z)$ and z^B are nonsingular we have $G(z) \in \mathrm{GL}(n, \mathbb{C})$. Both $F(z)$ and z^B have analytic continuation along loops γ in $D^\times_{r_0}$ based at z_0, and therefore also the product $G(z)$ has analytic continuation. For the monodromy operator $M = M([\gamma])$ along γ we find

$$M(G(z)) = M(F(z))M(z^B) = F(z)Mz^Be^{2\pi iB} = G(z)$$

because $z^Be^{2\pi iB} = e^{2\pi iB}z^B$ and $M = e^{-2\pi iB}$. Therefore the function $G(z)$ has trivial monodromy, so is univalued and holomorphic on $D^\times_{r_0}$. Hence the function $G(z)$ has a Laurent series expansion

$$G(z) = \sum_{k=-\infty}^{\infty} G_k z^k$$

with coefficients matrices $G_k \in \mathrm{Mat}(n, \mathbb{C})$, which converges on $D^\times_{r_0}$. The original local solution matrix $F(z)$ is therefore of the form

$$F(z) = G(z)z^{-B}$$

with $G : D^\times_{r_0} \to \mathrm{GL}(n, \mathbb{C})$ univalued and holomorphic. The multivalued behaviour of $F(z)$ is just a consequence of the factor z^{-B}.

Definition 1.21. *The solutions of* $(\partial + A)F = 0$ *have moderate growth near the singular point* $z = 0$ *if for each sector*

$$\{z \in D^\times_{r_0}; \theta_1 < \arg z < \theta_2\}$$

with $\theta_1 < \theta_2 < \theta_1 + 2\pi$ *and each holomorphic solution* $F(z)$ *on this sector there exist constants* $C > 0$ *and* $D \in \mathbb{R}$ *with*

$$|F(re^{i\theta})| \le Cr^D$$

on this sector.

It is clear that the solution matrix $F(z)$ has moderate growth near $z = 0$ if and only if the matrix function $G(z)$ has a pole or a removable singularity at $z = 0$.

Definition 1.22. *The linear system $(\partial + A)F = 0$ on the punctured disc $D^{\times}_{r_0}$ has a regular singular point at $z = 0$ if $z \mapsto zA(z)$ is holomorphic at $z = 0$, or equivalently if the linear system has the form $(\theta + B)F = 0$ with $z \mapsto B(z) = zA(z)$ holomorphic at $z = 0$. Here we always denote $\theta = z\partial$.*

Theorem 1.23. *If the linear system $(\partial + A)F = 0$ has a regular singularity at $z = 0$ then all solutions have moderate growth at $z = 0$.*

We first prove a lemma.

Lemma 1.24. *If $(a, b) \ni r \mapsto F(r) \in \mathbb{R}^n - \{0\}$ is smooth then*

$$|\frac{d}{dr}|F(r)|| \leq |\frac{dF}{dr}(r)| \; .$$

Proof. Suppose $F(r) = (f_1(r), \cdots, f_n(r))^t$. Then we get

$$\frac{d}{dr}|F(r)| = \frac{d}{dr}(\sum_1^n f_j(r)^2)^{\frac{1}{2}} = (\sum_1^n f_j(r)^2)^{-\frac{1}{2}}(\sum_1^n f_j(r)\frac{df_j}{dr}(r))$$

which in absolute value is $\leq |\frac{dF}{dr}(r)|$ by the Cauchy inequality. ☐

We now come to the proof of the above theorem.

Proof. For $0 < r < r_1 < r_0$ we have

$$\log|F(re^{i\theta})| - \log|F(r_1e^{i\theta})| \leq \int_r^{r_1} |\frac{d}{ds}\log|F(se^{i\theta})||ds$$

$$= \int_r^{r_1} |F(se^{i\theta})|^{-1}|\frac{d}{ds}|F(se^{i\theta})||ds \leq \int_r^{r_1} |F(se^{i\theta})|^{-1}|\frac{d}{ds}F(se^{i\theta})|ds$$

$$= \int_r^{r_1} |F(se^{i\theta})|^{-1}|\partial F(se^{i\theta})|ds \leq M\int_r^{r_1} s^{-1}ds = M\log\frac{r_1}{r}$$

with $M = \max\{|zA(z)|; |z| \leq r_1\} < \infty$ by assumption. Because the natural logarithm is monotonically increasing we find

$$|F(re^{i\theta})| \leq (\frac{r_1}{r})^M|F(r_1e^{i\theta})| \; ,$$

such that for $C = \max\{r_1^M|F(r_1e^{i\theta})|; \theta_1 \leq \theta \leq \theta_2\}\}$ and $D = -M$ the required inequality of Definition 1.21 is obtained. ☐

The converse of the above theorem is not true, in the sense that for a linear system having a regular singularity at $z = 0$ it is not a necessary condition for the solutions having moderate growth at $z = 0$.

Example 1.25. *For $n = 2$ consider the linear system $(\theta + B)F = 0$ on $\mathbb{C}^\times$ with coefficients matrix*

$$B = \begin{pmatrix} 0 & -z^k \\ 0 & k \end{pmatrix}$$

for some $k \in \mathbb{Z}$ and $F = (f_1, f_2)^t$. Spelled out the linear system becomes

$$z\partial f_1 - z^k f_2 = 0 \ , \ z\partial f_2 + k f_2 = 0 \ .$$

The second equation (after multiplication by z^{k-1}) becomes $\partial(z^k f_2) = 0$. Hence $f_2 = c_2 z^{-k}$ for some integration constant c_2. Substitution in the first equation gives $z\partial f_1 - c_2 = 0$, and therefore $f_1 = c_1 + c_2 \log z$ for a second integration constant c_1. Hence the general solution becomes

$$F(z) = (f_1(z), f_2(z))^t = c_1(1, 0)^t + c_2(\log z, z^{-k})^t \ .$$

These functions have moderate growth for all $k \in \mathbb{Z}$, but clearly for $k \leq -1$ the linear system is not regular singular at $z = 0$. Therefore being regular singular of $(\partial + A)F = 0$ at $z = 0$ is a sufficient, but not a necessary condition for the solutions having moderate growth around $z = 0$.

Let us now consider a linear system with a regular singularity at $z = 0$, so a linear system of the form

$$(\partial + A)F = 0$$

or equivalently

$$(\theta + B)F = 0$$

with $\theta = z\partial$ and $B = zA$ holomorphic on the disc $D_{r_0} = \{|z| < r_0\}$ for some $r_0 > 0$. Hence the power series

$$B(z) = \sum_0^\infty B_k z^k$$

with $B_k \in \mathrm{Mat}(n, \mathbb{C})$ converges on D_r. The Frobenius method consists in the substition of a formal series

$$F(z) = z^s \sum_0^\infty F_k z^k = \sum_0^\infty F_k z^{s+k}$$

with $s \in \mathbb{C}$ undetermined and $F_k \in \mathbb{C}^n$ undetermined. We have the following analogue of Proposition 1.2.

Proposition 1.26. *The formal series* $F(z) = z^s \sum_0^\infty F_k z^k$ *is a solution of* $(\theta + B)F = 0$ *with* $B(z) = \sum_0^\infty B_k z^k$ *if and only if*

$$(s + B_0)F_0 = 0$$

and

$$(s + k + 1 + B_0)F_{k+1} + \sum_{l=0}^{k} B_{k+1-l}F_l = 0$$

for all $k \in \mathbb{N}$.

Proof. This is a direct computation using $\theta(z^{s+k}) = (s + k)z^{s+k}$. □

Definition 1.27. *The characteristic equation* $\det(s + B_0) = 0$ *is called the indicial equation and the roots of the indicial equation are called the exponents of the differential equation* $(\theta + B)F = 0$ *at* $z = 0$.

Corollary 1.28. *Consider the linear system* $(\theta + B)F = 0$ *with a regular singularity at* $z = 0$. *If* s *is an exponent but* $(s+k+1)$ *is not an exponent for all* $k \in \mathbb{N}$ *then there exists for each* $F_0 \in \mathrm{Ker}(s+B_0)$ *a unique formal solution* $F(z) = z^s \sum_0^\infty F_k z^k$ *with* $F_{k+1} \in \mathbb{C}^n$ *given by the recurrence relations in Proposition 1.26.*

Proof. This is clear from the recurrence relations in Proposition 1.26 because $(s + k + 1 + B_0)$ is invertible for all $k \in \mathbb{N}$. □

Theorem 1.29. *The formal solution* $F(z) = z^s \sum_0^\infty F_k z^k$ *with* $F_k \in \mathbb{C}^n$ *given by Proposition 1.26 converges on* D_{r_0}.

Proof. Because $B(z) = \sum_0^\infty B_k z^k$ converges on D_{r_0} we know by Lemma 1.4 $\forall \rho \in (0, 1)$ the existence of a constant $M_\rho \geq 0$ such that

$$|B_k| \leq M_\rho(\rho r_0)^{-k} \; \forall k \in \mathbb{N} \; .$$

Because $(s + k + 1)$ is not an exponent for $k \in \mathbb{N}$ there exists a constant $K \geq 1$ such that

$$|(s + k + 1 + B_0)^{-1}| \leq K(k + 1)^{-1} \; .$$

Using this inequality one can show by induction on k (as in the proof of Theorem 1.5) that

$$|F_k| \leq M_\rho(M_\rho + 1) \cdots (M_\rho + k - 1)K^k(\rho r_0)^{-k}|F_0|/k!$$

for all $k \in \mathbb{N}$. Hence the formal series $F(z) = z^s \sum_0^\infty F_k z^k$ converges on the disc $|z| < r_0/K$. But then the series also converges on D_{r_0} as solution of the differential equation $(\theta + B)F = 0$. □

Corollary 1.30. *Consider a linear system* $(\theta + B)F = 0$ *on* $D_{r_0}^{\times}$ *with a regular singularity at* $z = 0$*. Suppose the exponents* $s_1, \cdots, s_n$ *at* $z = 0$ *are modulo* $\mathbb{Z}$ *distinct:* $s_i - s_j \notin \mathbb{Z}$ *for* $i \neq j$*. Then the* n *solutions*

$$F_j(z) = z^{s_j}(F_{j0} + F_{j1}z + F_{j2}z^2 + \cdots)$$

with $0 \neq F_{j0} \in \mathrm{Ker}(s_j + B_0)$ *for* $j = 1, \cdots, n$ *are a basis of the local solution space, say around* $z = r_0/2$ *(with* $z^s = e^{s \log z}$ *and* $\Im(\log(r_0/2)) = 0$*).*

Proposition 1.31. *Consider the linear system* $(\theta + B)F = 0$ *with a regular singularity at* $z = 0$*. The matrix* $e^{-2\pi i B_0}$ *lies in the closure of the conjugation orbit of the monodromy matrix* M_0*. In particular* $e^{-2\pi i B_0}$ *and* M_0 *have the same characteristic polynomial.*

Proof. In polar coordinates $z = re^{i\theta}$ we integrate $(\theta + B)F = 0$ along circles $r =$ constant. Because $\theta = z\partial = -id/d\theta$ we get

$$(\frac{d}{d\theta} + iB_0 + rC(r, \theta))F(r, \theta) = 0$$

for $0 \leq r < r_0, 0 \leq \theta \leq 2\pi$ and $C(r, \theta) = r^{-1}i(B(re^{i\theta}) - B_0)$. Let now $F(r, \theta)$ be the solution matrix with initial value $F(r, 0) = I_n$ for all $r \in [0, r_0)$. Then $M(r) = F(r, 2\pi)$ is the monodromy matrix obtained by analytic continuation along paths $\gamma_r(t) = re^{2\pi i t}$ with time t and fixed $r \in (0, r_0)$. Hence $M(r)$ is conjugated with M_0 for all $r \in (0, r_0)$. The function $C(r, \theta)$ is continuous for $(r, \theta) \in [0, r_0) \times [0, 2\pi]$. Hence the solution $F(r, \theta)$ with continuous initial value $F(0, \theta)$ for $\theta \in [0, 2\pi]$ is also continuous for $(r, \theta) \in [0, r_0) \times [0, 2\pi]$. Hence $M(0) = F(0, 2\pi)$ is equal to the limit of $M(r)$ for $r \downarrow 0$. However $M(0) = e^{-2\pi i B_0}$ by direct integration, since $rC(r, \theta)$ vanishes in the limit for $r \downarrow 0$. This proves the proposition. □

Corollary 1.32. *A matrix in* $\mathrm{GL}_n(\mathbb{C})$ *is called regular if the conjugacy class has maximal dimension* $n(n-1)$*, or equivalently if the centralizer in* $\mathrm{GL}_n(\mathbb{C})$ *has minimal dimension* n*. If the matrix* $e^{-2\pi i B_0}$ *is regular then* M *and* $e^{-2\pi i B_0}$ *are conjugated.*

Example 1.33. *Consider the linear system* $(\theta + B)F = 0$ *for* $n = 2$ *and with coefficients matrix*

$$B = \begin{pmatrix} 0 & -z^k \\ 0 & k \end{pmatrix}$$

for some $k \in \mathbb{Z}$ *discussed in Example 1.25. Relative to the solution matrix*

$$F = \begin{pmatrix} 1 & \log z \\ 0 & z^{-k} \end{pmatrix}$$

the monodromy matrix around $z = 0$ is given by

$$M = \begin{pmatrix} 1 & 2\pi i \\ 0 & 1 \end{pmatrix}$$

while for $k \geq 1$ the matrix $e^{2\pi i B_0}$ is the identity matrix. Hence the matrices M and $e^{-2\pi i B_0}$ need not be conjugated. The conjugacy class of M consists of all regular unipotent matrices, and the identity matrix lies in the closure of this orbit.

Definition 1.34. *A first order system of differential equations $(\partial + A)F = 0$ on the complement of a finite set S in the projective line $\mathbb{P}$ is called Fuchsian if all its singular points are regular singular.*

A first order system $(\partial + A)F = 0$ is regular at $z = \infty$ if and only $A(z) = O(z^{-2})$ for $z \to \infty$. Hence the general form of a first order Fuchsian system with regular singularities at $z_1, \cdots, z_m$ in the complex plane $\mathbb{C}$ and regular at $z = \infty$ is

$$(\partial + A)F = 0 \quad , \quad A(z) = \frac{A_1}{z - z_1} + \cdots + \frac{A_m}{z - z_m}$$

for certain matrices $A_1, \cdots, A_m \in \mathrm{Mat}(n, \mathbb{C})$ with $A_1 + \cdots + A_m = 0$. If the latter condition is dropped then $z = \infty$ becomes also a regular singular point.

1.5 The theorem of Fuchs

Consider the n^{th} order scalar differential equation

$$(\partial^n + a_1 \partial^{n-1} + \cdots + a_{n-1}\partial + a_n)f = 0$$

with coefficients holomorphic on the puntured disc $D_{r_0}^{\times}$ for some $r_0 > 0$.

Lemma 1.35. *If $\theta = z\partial$ then $z^k \partial^k = \theta(\theta - 1) \cdots (\theta - k + 1)$ for $k \in \mathbb{N}$.*

Proof. By induction on k we have

$$z^{k+1}\partial^{k+1} = z\theta(\theta - 1) \cdots (\theta - k + 1)\partial = (\theta - 1)(\theta - 2) \cdots (\theta - k) z\partial$$

because $z\theta = (\theta - 1)z \Leftrightarrow \theta z = z(\theta + 1)$ by the Leibniz product rule. ☐

Multiplying the above differential equation by z^n we can rewrite this equation in the form

$$(\theta^n + b_1\theta^{n-1} + \cdots + b_{n-1}\theta + b_n)f = 0$$

with the transition from the functions $\{1, za_1, z^2a_2, \cdots, z^na_n\}$ to the new coefficients $\{1, b_1, b_2, \cdots, b_n\}$ given by an integral unitriangular matrix (with unitriangular meaning lower triangular with 1 on the diagonal). Note that the collection of unitriangular matrices in $\mathrm{GL}_n(\mathbb{Z})$ is a group.

Definition 1.36. *The point $z = 0$ is a regular singular point of the above n^{th} order scalar differential equation if $z^j a_j$ is holomorphic at $z = 0$ $\forall j$ or equivalently if b_j is holomorphic at $z = 0$ $\forall j$.*

The next result is called the theorem of Fuchs. It marks an important difference between first order matrix systems and n^{th} order scalar differential equations.

Theorem 1.37. *The point $z = 0$ is a regular singularity of the above n^{th} order scalar differential equation if and only if all solutions around $z = 0$ have moderate growth.*

Proof. Suppose $z = 0$ is a regular singular point of the n^{th} order scalar differential equation, so $b_1, \cdots, b_n$ are holomorphic around $z = 0$ $\forall j$. We associate to the n^{th} order scalar differential equation a first order matrix system

$$(\theta + B)F = 0$$

with the coefficient matrix B given by

$$\begin{pmatrix} 0 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 \\ b_n & b_{n-1} & b_{n-2} & b_{n-3} & \cdots & b_2 & b_1 \end{pmatrix}$$

and

$$F = (f, \theta f, \theta^2 f, \cdots, \theta^{n-1} f)^t\ .$$

This F is a solution of this first order matrix system if and only if the first coordinate f of F is a solution of the n^{th} order scalar differential equation. Hence f has moderate growth around $z = 0$ by Theorem 1.23.

Conversely, suppose that all solutions of the n^{th} order scalar differential equation have moderate growth around $z = 0$. We prove the statement by induction on the order n of the scalar equation. There always exists a multivalued holomorphic solution on $\mathbb{D}_{r_0}^{\times}$ ($r_0 > 0$ sufficiently small) of the form

$$f_0(z) = z^s(1 + O(z)) \ , \ z \to 0$$

with exponent $s \in \mathbb{C}$. Indeed, just take a suitably normalized eigenvector of the monodromy operator $M(\gamma(t) = r_0 e^{2\pi i t}/2)$ in the local solution space around $z_0 = r_0/2$. Consider the linear differential operators

$$D = \theta^n + b_1\theta^{n-1} + \cdots + b_{n-1}\theta + b_n$$

and

$$E = f_0^{-1} \circ D \circ f_0 = \theta^n + c_1\theta^{n-1} + \cdots + c_{n-1}\theta + c_n \ .$$

Here f_0 stands for the 0^{th} order linear differential operator of multiplication by f_0. Because $f_0^{-1} \circ \theta \circ f_0 = \theta + \theta(f_0)/f_0$ with $\theta(f_0)/f_0$ (univalued) holomorphic around $z = 0$, we conclude that $b_1, \cdots, b_n$ are holomorphic around $z = 0$ if and only if $c_1, \cdots, c_n$ are holomorphic around $z = 0$. Moreover $E(1) = 0$ hence $c_n = 0$. In other words E factorizes as

$$E = F\theta$$

with $F = \theta^{n-1} + c_1\theta^{n-2} + \cdots c_{n-1}$. The solutions g of $E(g) = 0$ and h of $F(h) = 0$ are related by $h = \theta g$. The solutions g of $E(g) = 0$ are of the form $g = f/f_0$ with f a solution of $Df = 0$. The solutions f of $D(f) = 0$ have moderate growth around $z = 0$ by assumption. Hence the solutions g of $E(g) = 0$ have moderate growth around $z = 0$, but then also the solutions h of $F(h) = 0$ have moderate growth around $z = 0$. By induction on the the order n of the scalar equation we can assume that $c_1, \cdots, c_{n-1}$ are holomorphic around $z = 0$. Because $c_n = 0$ is holomorphic as well we conclude that $b_1, \cdots, b_n$ are holomorphic around $z = 0$. This completes the proof of the theorem of Fuchs. □

Definition 1.38. *If the n^{th} order scalar linear differential equation*

$$Df = 0 \ , \ D = \theta^n + b_1\theta^{n-1} + \cdots + b_{n-1}\theta + b_n$$

on $D_{r_0}^{\times}$ has a regular singularity at $z = 0$ then the degree n polynomial equation

$$s^n + b_1(0)s^{n-1} + \cdots + b_{n-1}(0)s + b_n(0) = 0$$

is called the indicial equation and its roots are called the exponents of $Df = 0$ at $z = 0$.

A solution of $Df = 0, D = \theta^n + b_1\theta^{n-1} + \cdots + b_{n-1}\theta + b_n$ around the regular singular point $z = 0$ of the form

$$f(z) = z^s \sum_0^\infty f_k z^k$$

with $f_k \in \mathbb{C}, f_0 \neq 0$ is called a formal solution with exponent s. Such a formal solution is only possible if s is a root of the indicial equation. The coefficients $f_k \in \mathbb{C}, k \in \mathbb{Z}$ of such a solution are again given by recurrence relations, and these have a unique solution (for given f_0) if $s + k$ is not an exponent $\forall k \in \mathbb{Z}, k \geq 1$. If the n exponents at $z = 0$ are all distinct modulo $\mathbb{Z}$ then there exists a basis of formal solutions with these exponents. Using Theorem 1.29 it follows that these formal solutions have a positive radius of convergence.

The eigenvalues of the monodromy operator around $z = 0$ are of the form $e^{2\pi i s}$ with s an exponent at $z = 0$. However, just like in Proposition 1.31 the Jordan normal form of the monodromy operator around $z = 0$ can not in general be deduced from the indicial equation, namely in case some exponents coincide modulo $\mathbb{Z}$.

Definition 1.39. *An n^{th} order scalar linear differential equation on the projective line $\mathbb{P}$ minus the singular points of the form*

$$Df = 0 \ , \ D = \partial^n + a_1\partial^{n-1} + \cdots + a_{n-1}\partial + a_n$$

with rational coefficients $a_1, \cdots, a_n \in \mathbb{C}(z)$ is called a Fuchsian equation if all its singular points (possibly including $z = \infty$) are regular singular.

In order to analyze the behaviour of an n^{th} order scalar linear differential equation at $z = \infty$ one makes the substitution $w = z^{-1}$ and considers the behaviour of the transformed equation at $w = 0$. The same strategy works for first order matrix systems. Remark that $\theta = zd/dz = -wd/dw$.

1.6 The Riemann–Hilbert problem

The general form of a first order Fuchsian system with regular singularities at $z_1, \cdots, z_m$ in the complex plane $\mathbb{C}$ and regular at $z = \infty$ is

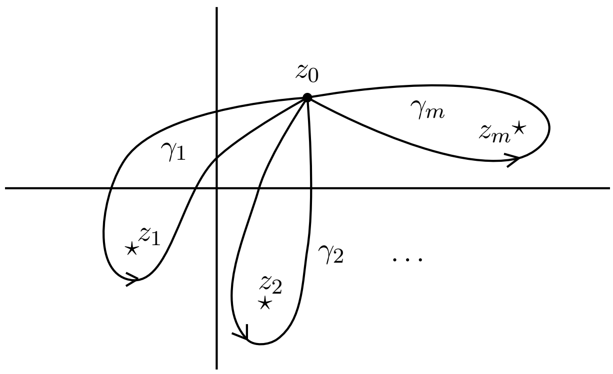
$$(\partial + A)F = 0 \ , \ A(z) = \frac{A_1}{z - z_1} + \cdots + \frac{A_m}{z - z_m}$$

for certain matrices $A_1, \cdots, A_m \in \mathrm{Mat}(n, \mathbb{C})$ with $A_1 + \cdots + A_m = 0$. If we choose an additional base point z_0 in $\mathbb{C}$ then the monodromy group

$\Pi_1(\mathbb{P} - \{z_1, z_2, \cdots, z_m\}, z_0)$ is the group with generators $\gamma_1, \gamma_2, \cdots, \gamma_m$ as in the picture below, and the single relation $\gamma_m \cdots \gamma_2 \gamma_1 = 1$. The monodromy representation

$$M : \Pi_1(\mathbb{P} - \{z_1, z_2, \cdots, z_m\}, z_0) \to \mathrm{GL}(n, \mathbb{C})$$

is known, once we can compute the monodromy matrices $M(\gamma_i) \in \mathrm{GL}(n, \mathbb{C})$ for $i = 1, 2, \cdots, m$.



The calculation of the monodromy is easy in case $n = 1$, because the group $\mathrm{GL}(1, \mathbb{C}) = \mathbb{C}^\times$ is Abelian. Indeed, in this case $A_1 = a_1, \cdots, A_m = a_m$ are just scalars, and the general solution of the above Fuchsian differential equation becomes $c(z - z_1)^{-a_1} \cdots (z - z_n)^{-a_m}$ and so $M(\gamma_j) = e^{-2\pi i a_j}$. The case $n \geq 2$ and $m = 2$ is again easy to solve, because the fundamental group $\Pi_1(\mathbb{C}^\times, 1) \cong \mathbb{Z}$ is Abelian. But for $n \geq 2$ and $m \geq 3$ both groups $\mathrm{GL}(n, \mathbb{C})$ and $\Pi_1(\mathbb{P} - \{z_1, z_2, \cdots, z_m\}, z_0)$ are no longer Abelian, and the calculation of the monodromy representation seems to be an intractible transcendental problem.

However, in 1857 Riemann did show that for $n = 2$ and $m = 3$, which amounts to the case of the Euler–Gauss hypergeometric equation, one can describe the monodromy representation in terms of the local exponents in an algebraic way. On the basis of this example Riemann expected (unpublished, but in his Nachlass) that all representations of $\Pi_1(\mathbb{P} - \{z_1, \cdots, z_m\}, z_0)$ of dimension n might occur as monodromy representations of a first order Fuchsian system of rank n with regular singular points at $z_1, \cdots, z_m$. The fact that both Fuchsian systems of rank n with m prescribed regular singular points and a representations of dimension n of the fundamental group of their complement have the same number of moduli, namely $n^2(m - 2) + 1$, is at least an indication that this question might have a positive answer.

A modulus is an essential parameter of a problem, so after dividing out symmetries. In the above case Fuchsian systems are taken up to isomorphism and monodromy representations are taken up to equivalence.

Hilbert formulated this question in his famous list of mathematical problems from 1900 as Problem 21, and the question became known as the Riemann–Hilbert problem. The answer to the question was shown to be essentially yes by Josip Plemelj in 1905. For the modern solution of the problem one has to speak the language of holomorphic vector bundles and regular singular connections.

A Riemann surface S is just a one dimensional complex manifold. Basic examples are the Riemann sphere $\mathbb{P} = \mathbb{C} \sqcup \{\infty\}$, the m-punctured Riemann sphere $\mathbb{P} - \{z_1, \cdots, z_m\}$, the unit disc $\mathbb{D}$ and the punctured unit disc $\mathbb{D}^\times$.

Definition 1.40. *A holomorphic vector bundle of rank n on a Riemann surface S is a holomorphic map $p : V \to S$ such that*

- $\forall\, z \in S$, *the fiber $p^{-1}(z)$ is a complex vector space of dimension n,*
- *each $z \in S$ has an open neighborhood U and a biholomorphic map $\varphi_U : p^{-1}(U) \to U \times \mathbb{C}^n$, called a local trivialization over U, such that*

$$\begin{array}{ccc} p^{-1}(U) & \xrightarrow{\varphi_U} & U \times \mathbb{C}^n \\ & {\scriptstyle p} \searrow \quad \swarrow {\scriptstyle \mathrm{pr}_U} & \\ & U & \end{array}$$

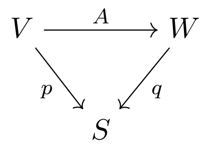
  with pr_U projection on the first factor. The map $\varphi_U : p^{-1}(U) \to U \times \mathbb{C}^n$ is called a local trivialization over U, and

- *the induced map $\varphi_z : p^{-1}(z) \to \mathbb{C}^n$ is a linear isomorphism.*

We shall write shortly vector bundle for holomorphic vector bundle. The simplest example of a vector bundle on S is the trivial vector bundle on S $\mathrm{pr}_S : S \times \mathbb{C}^n \to S$ of rank n. Vector bundles of rank one are called line bundles. Any construction of linear algebra on vector spaces, like the direct sum $\oplus$, the tensor product $\otimes$ and the dual vector space, can be performed likewise with vector bundles. For example, the dual of the tangent line bundle $T(S)$ of S is the cotangent line bundle $\Omega(S) = T^*(S)$ of S. If $p : V \to S$ is a vector bundle then $\Omega(V) = \Omega(S) \otimes V$ is the vector bundle of differentials with values in V.

Definition 1.41. *A morphism of vector bundles $p : V \to S$ and $q : W \to S$ on S is a holomorphic map $A : V \to W$, such that*

- *the diagram*



  is commutative, and

- *for all $z \in S$, the restriction*

$$A_z : p^{-1}(z) \to q^{-1}(z)$$

  is a linear map.

Let $p : V \to S$ be a vector bundle on S. If $\{U_i; i \in I\}$ is an open covering of S, such that the restriction of V on each U_i admits a trivialization $\varphi_i : p^{-1}(U_i) \to U_i \times \mathbb{C}^n$, then for each $i, j \in I$ we define the transition function

$$\varphi_{ij} : U_i \cap U_j \to \mathrm{GL}(n, \mathbb{C})$$

by $v = \varphi_{ij}(z)w$ if and only if $\varphi_i^{-1}(z, v) = \varphi_j^{-1}(z, w)$ for $z \in U_i \cap U_j$ and $v, w \in \mathbb{C}^n$. For $i, j, k \in I$ the transition functions satisfy $\varphi_{ij}(z)\varphi_{jk}(z) = \varphi_{ik}(z)$ for $z \in U_i \cap U_j \cap U_k$ and $\varphi_{ii}(z) = \mathrm{Id}_n$ for $z \in U_i$, which are called the cocycle relations. The vector bundle can be recovered from its transition functions by

$$V = \{\sqcup_{i \in I}(U_i \times \mathbb{C}^n)\}/\sim$$

with $(U_i \times \mathbb{C}^n) \ni (z, v) \sim (z, w) \in (U_j \times \mathbb{C}^n)$ for $z \in U_i \cap U_j$ and $v, w \in \mathbb{C}^n$ if and only if $v = \varphi_{ij}(z)w$. In other words, a general vector bundle is built up from locally trivial bundles by gluing.

Example 1.42. *The Riemann sphere $\mathbb{P}$ is covered by the two open sets $U_0 = \mathbb{P} - \{\infty\}$ and $U_\infty = \mathbb{P} - \{0\}$. If we take $\varphi_{0\infty}(z) = z^m$ for $z \in \mathbb{C}^\times$ and $m \in \mathbb{Z}$, then the corresponding line bundle on $\mathbb{P}$ is denoted $L(m)$.*

Definition 1.43. *If $p : V \to S$ is a vector bundle on S, then a holomorphic map $s : S \to V$ with $p \circ s = \mathrm{Id}_S$ is called a (global) section. The set of sections $\Gamma(V)$ has the structure of a complex vector space and a module over the algebra of holomorphic functions on S.*

Consider the example of the line bundle $L(m) \to \mathbb{P}$. Let $z \in \mathbb{C}$ be the natural coordinate on U_0, and $w \in \mathbb{C}$ the natural coordinate on U_∞, related by $z = 1/w$ on the intersection. For $k \in \mathbb{Z}$ the meromorphic section w^k on U_∞ is identified with the meromorphic section z^{m-k} on U_0. We see that the vector space $\Gamma(L(m))$ of global holomorphic sections has dimension 0 for $m < 0$ and dimension $m+1$ for $m \geq 0$.

Because a vector bundle $p : V \to S$ might have no nonzero sections it is useful to consider for each open subset U of S the local sections $\Gamma(V|_U)$. For U sufficiently small these will be infinite dimensional vector spaces. All these spaces together form the sheaf of local sections of V, and will be denoted $\mathcal{V}$. The sheaf of local sections of the trivial line bundle $\mathrm{pr}_S : \mathcal{O}(S) = S \times \mathbb{C} \to S$ is called the structure sheaf of S, and is denoted $\mathcal{O}_S$. Likewise, the sheaf of local sections of the cotangent line bundle $\Omega(S) \to S$ is denoted Ω_S. The global sections of $\Omega(S) \to S$ are the holomorphic differentials on S. The sheaf of local sections of $\Omega(V) \to S$ is denoted $\Omega(\mathcal{V})$.

Definition 1.44. *A (holomorphic) connection ∇ on a (holomorphic) vector bundle $V \to S$ is a linear map $\nabla : \mathcal{V} \to \Omega(\mathcal{V})$ satisfying the Leibniz rule*

$$\nabla(fs) = df \otimes s + f\nabla(s)$$

for all $f \in \mathcal{O}_S$ and all $s \in \mathcal{V}$.

If S is an open subset of $\mathbb{C}$, then any connection ∇ on the trivial vector bundle $\mathrm{pr}_S : S \times \mathbb{C}^n \to S$ is of the form $\nabla = d + A(z)dz$ for some holomorphic map $A : S \to \mathrm{Mat}(n, \mathbb{C})$. Indeed, it is obvious that $\nabla = d$ is a connection on the trivial bundle, and the difference $\nabla_1 - \nabla_2$ of two connections is $\mathcal{O}_S$-linear by definition.

Let (V, ∇) be a vector bundle with connection on S. A local section $s \in \mathcal{V}$ is called horizontal if $\nabla(s) = 0$. If $\gamma : [0,1] \to S$ is a curve in S with begin point z_0 and end point z_1, then parallel transport along γ (using horizontal sections) induces a linear isomorphism $M : V_0 \to V_1$ with V_0, V_1 the fibers of V over z_0, z_1 respectively. But these are just new words for familiar concepts: a horizontal local section $s(z) = (z, f(z))$ on U in a local trivialization $p^{-1}(U) \cong U \times \mathbb{C}^n$ with $\nabla = d + A(z)dz$ is just a solution of $(\partial + A)f = 0$, and parallel transport of horizontal sections is just analytic continuation of solutions of $(\partial + A)f = 0$.

Theorem 1.45. *If $p : V \to S$ is a vector bundle on S with connections ∇_1 and ∇_2, then (V, ∇_1) and (V, ∇_2) are isomorphic as vector bundles with connection (the vector bundle isomorphism maps horizontal sections to horizontal sections) if and only if their monodromy representations*

$M_1 : \Pi_1(S, z_0) \to \mathrm{GL}(V_0)$ and $M_2 : \Pi_1(S, z_0) \to \mathrm{GL}(V_0)$ are equivalent. Moreover, each matrix representation $M : \Pi_1(S, z_0) \to \mathrm{GL}(n, \mathbb{C})$ occurs as the monodromy representation of some vector bundle with connection (V, ∇) of rank n on S.

Proof. Suppose $A : V_0 \to V_0$ is a linear intertwining isomorphism, that is $AM_1(\gamma) = M_2(\gamma)A$ for all $\gamma \in \Pi_1(S, z_0)$. Parallel transport of horizontal sections gives locally around z_0 the desired isomorphism, and because the monodromy representations are intertwined by A the isomorphism is even globally defined.

If $\tilde{S} \to S$ is the universal covering relative to the base point $z_0 \in S$, then $\Pi_1(S, z_0)$ acts freely on $\tilde{S}$ with quotient S. Indeed, if $\tilde{S} \ni \tilde{z} \mapsto z \in S$ is represented by a curve $\delta : [0, 1] \to S$ with $\delta(0) = z_0$ and $\delta(1) = z$, then we define $\gamma \cdot \tilde{z} = \delta\gamma^{-1}$ for $\gamma \in \Pi_1(S, z_0)$. Now the fundamental group $\Pi_1(S, z_0)$ acts on the trivial vector bundle with connection $(\tilde{S} \times \mathbb{C}^n, \nabla = d)$ by $\gamma \cdot (\tilde{z}, v) = (\gamma \cdot \tilde{z}, M(\gamma)v)$, and the quotient bundle by the action of $\Pi_1(S, z_0)$ gives the desired vector bundle with connection, with monodromy the given monodromy representation. □

Definition 1.46. *Suppose $V \to S$ is a vector bundle on a Riemann surface S, and let $z \in S$. A holomorphic connection ∇ on the restriction of V to $S - \{z\}$ is called regular singular in z if for each neighborhood U of z and each local section $s \in \Gamma(U, V|_U)$ the section $\nabla(s)$ has at most a simple pole in z.*

Consider the line bundle $L(m)$ on $\mathbb{P}$, given by the open covering $U_0 = \mathbb{C}$ with coordinate z and $U_\infty = \mathbb{P} - \{0\} \cong \mathbb{C}$ with coordinate $w = 1/z$ and transition function $\varphi_{0\infty}(z) = z^m$ for $z \in \mathbb{C}^\times$. The trivial connection $\nabla_\infty = d$ over U_∞ becomes $\nabla_0 = d - (m/z)dz$ over U_0, and z^m is a horizontal section over U_0. This connection ∇ on $L(m)$ has a regular singular point at $z = 0$ in case $m \neq 0$. The monodromy is trivial, and for that reason $z = 0$ is called an apparent singularity.

Theorem 1.47. *Suppose S is Riemann surface with m distinct marked point $z_1, \cdots, z_m$ and a given base point $z_0 \in S - \{z_1, \cdots, z_m\}$. Any matrix representation $M : \Pi_1(S - \{z_1, \cdots, z_m\}, z_0) \to \mathrm{GL}(n, \mathbb{C})$ occurs as the monodromy representation of a vector bundle with connection (V, ∇) on S with regular singular points at $z_1, \cdots, z_m$.*

Proof. Let $Z = S - \{z_1, \cdots, z_m\}$. Choose loops $\gamma_j \in \Pi_1(Z, z_0)$ by choosing a path from z_0 to nearby z_j, then making a positive loop around z_j inside a small neighborhood $U_j \cong \mathbb{D}$ of z_j in S, with $z_i \in U_j$ only if $i = j$, and then returning back to z_0 along the original path. Hence $U_j^\times = U_j - \{z_j\}$ is biholomorphic to the punctured unit disc $\mathbb{D}^\times$.

Let $M : \Pi_1(Z, z_0) \to \mathrm{GL}(n, \mathbb{C})$ be a representation, and let us write $M_j = M(\gamma_j) \in \mathrm{GL}(n, \mathbb{C})$. By Theorem 1.45 there is a vector bundle with connection (V°, ∇°) on Z with monodromy the prescribed representation. The restriction of (V°, ∇°) to $U_j^\times \cong \mathbb{D}^\times$ has monodromy M_j. If we choose a matrix $A_j \in \mathrm{Mat}(n, \mathbb{C})$ with $M_j = e^{-2\pi i A_j}$ then clearly the differential equation $(\partial + A_j/z)F = 0$ on $\mathbb{D}$ with regular singular point at $z = 0$ also has monodromy for the standard generator of $\Pi_1(\mathbb{D}^\times, 1/2)$ equal to M_j. By Theorem 1.45 the restriction of this system to $\mathbb{D}^\times$ is isomorphic as vector bundles with connection to the restriction of (V°, ∇°) to $U_j^\times$. By transport of structures we can extend the vector bundle V° over Z to a vector bundle V over S, and the holomorphic connection ∇° on V° to a meromorphic connection on V with regular singular points at $z_1, \cdots, z_m$. □

What have we done? We have given a couple of slick fairly abstract definitions, namely of a vector bundle with connection with regular singular points, and in this generalized setting the Riemann–Hilbert problem is almost trivially solved. Remark that our construction is not quite canonical, because of the ambiguity in the choice of the matrix A_j with $M_j = e^{-2\pi i A_j}$. The Riemann–Hibert problem asks for a regular singular system of the form

$$(\partial + A)F = 0 \ , \ A(z) = \frac{A_1}{z - z_1} + \cdots + \frac{A_m}{z - z_m}$$

in case $S = \mathbb{P}$, for certain matrices $A_1, \cdots, A_m \in \mathrm{Mat}(n, \mathbb{C})$ with $\sum A_j = 0$, and with prescribed monodromy. This goal would be achieved if the vector bundle V in the above theorem in case $S = \mathbb{P}$ is (holomorphically) trivial. We have the following result of Birkhoff and Grothendieck.

Theorem 1.48. *Any holomorphic vector bundle of rank n on the Riemann sphere $\mathbb{P}$ is isomorphic to a direct sum of line bundles $L(m_1) \oplus \cdots \oplus L(m_n)$ for certain integers $m_1, \cdots, m_n$, which are unique up to permutations.*

This implies that the desired regular singular system as above can be found, except that $\sum A_j$ need not be 0, but is a diagonal matrix with integer entries. So the point $z = \infty$ need not be a regular point, but will be an additional regular singular point with trivial local monodromy, so an apparent singularity.

Can we get rid of the apparent singularity at $z = \infty$ using the ambiguity of our construction? In case one of the singular points z_j has semisimple monodromy the answer is always yes. For $n = 2$ the answer is also always yes, as shown by Dekkers [9] in his Nijmegen PhD, written under Levelt (whose work on the Clausen–Thomae hypergeometric function will be discussed in a later chapter [26]). But for $n \geq 3$ it was shown by Bolibruch that there are examples for which the apparent singularity at $z = \infty$ can not be removed [1]. The above approach has been generalized to the case of several variables by Deligne [10].

1.7 Exercises

Exercise 1.1. *Show the formula*

$$\sum_{l=0}^{k} \frac{\Gamma(M+l)}{\Gamma(M)l!} = \frac{\Gamma(M+k+1)}{\Gamma(M+1)k!}$$

by induction on k.

Exercise 1.2. *Suppose we have given a second order differential equation of the form $(\partial^2 + a_1\partial + a_2)f = 0$ with coefficients a_1, a_2 holomorphic on $\mathbb{C}^\times$, and suppose that $f(z) = \log z = \log r + i\theta$ is a local solution around $z = 1$. Show that $f(z) \equiv 1$ is also a local solution around $z = 1$. Conclude that $a_1 = 1/z, a_2 \equiv 0$. Is the same conclusion valid if we only know that the coefficients a_1, a_2 are holomorphic on $\Re z > 0$, while still assuming that $f(z) = \log z$ is a local solution around $z = 1$?*

Exercise 1.3. *Show that the Euler-Gauss hypergeometric equation*

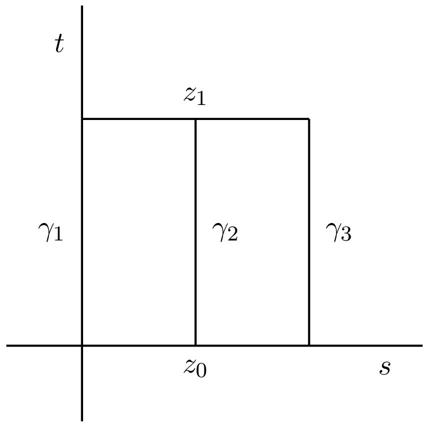
$$[z(1-z)\partial^2 + (\gamma - (\alpha+\beta+1)z)\partial - \alpha\beta)]f = 0$$

can be transformed to the form

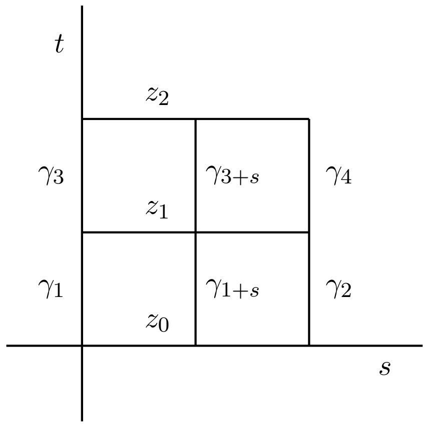
$$[(\theta+\gamma-1)\theta - z(\theta+\alpha)(\theta+\beta)]f = 0$$

with $\theta = z\partial$. Hint: Use that $z^2\partial^2 = \theta(\theta-1)$.

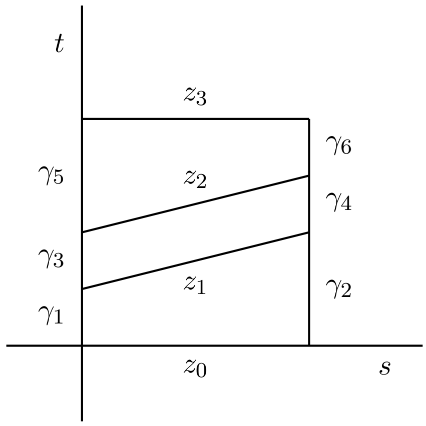
Exercise 1.4. *Show that homotopy equivalence for paths in Z is an equivalence relation. Here is a schematic picture of the argument.*



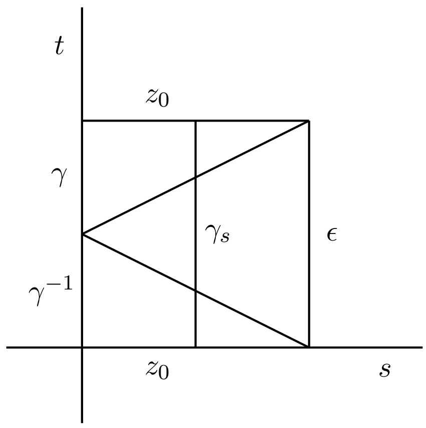
Exercise 1.5. *Show that if the end points of* $\gamma_1 \sim \gamma_2$ *coincide with the begin points of* $\gamma_3 \sim \gamma_4$ *then* $\gamma_3\gamma_1 \sim \gamma_4\gamma_2$.



Exercise 1.6. *Suppose that* $\gamma_1, \cdots, \gamma_6$ *are paths in* Z *such that* $\gamma_1 \sim \gamma_2$, $\gamma_3 \sim \gamma_4$, $\gamma_5 \sim \gamma_6$ *and the products* $\gamma_5(\gamma_3\gamma_1)$ *and* $(\gamma_6\gamma_4)\gamma_2$ *are well defined. In other words we assume that the begin points of* γ_1, γ_2 *equal* z_0, *the end points of* γ_1, γ_2 *and the begin points of* γ_3, γ_4 *equal* z_1, *the end points of* γ_3, γ_4 *and the begin points of* γ_5, γ_6 *equal* z_2, *and finally the end points of* γ_5, γ_6 *equal* z_3. *Show that* $\gamma_5(\gamma_3\gamma_1) \sim (\gamma_6\gamma_4)\gamma_2$. *In turn this implies that the group law on the fundamental group is associative. A picture of the homotopy is given by the picture below.*



Exercise 1.7. *Show that in the notation of Theorem 1.13 we have* $\epsilon\gamma \sim \gamma\epsilon \sim \gamma$ *and* $\gamma\gamma^{-1} \sim \gamma^{-1}\gamma \sim \epsilon$.



Here $\gamma_s(t)$ *is equal to* $\gamma^{-1}(2t)$ *for* $t \in [0, (1-s)/2]$, *is constant equal to* $\gamma^{-1}(1-s) = \gamma(s)$ *for* $t \in [(1-s)/2, (1+s)/2]$, *and is equal to* $\gamma(2t-1)$ *for* $t \in [(1+s)/2, 1]$

Exercise 1.8. *Let* $\gamma : [0,1] \to Z$ *be a path in* Z *with begin point* z_0 *and end point* z_1. *Show that conjugation by* $[\gamma]$ *induces an isomorphism* $\Pi_1(Z, z_0) \to \Pi_1(Z, z_1)$.

Exercise 1.9. *Compute the monodromy representation for the second order linear equation*

$$(z\partial^2 + \partial)f = 0$$

on the domain $Z = \mathbb{C}^\times$ *relative to the basis of solution* $f_1(z) = 1, f_2(z) = \log z$ *around* $z = 1$.

Exercise 1.10. *Prove Proposition 1.20. Give a counterexample for the failure of the proposition if we replace the field* $\mathbb{C}$ *by* $\mathbb{R}$.

Exercise 1.11. *Consider the linear system* $(\partial + A)F = 0$ *with coefficient matrix* A *holomorphic on the domain* Z. *Suppose* $F = (F_1, \cdots, F_n)$ *is a local solution matrix. Show that* $f = \det(F)$ *is a local solution of the first order scalar equation*

$$(\partial + \operatorname{tr}(A))f = 0 \ .$$

In particular if $\operatorname{tr}(A) = 0$ *then the monodromy group is contained in the special linear group* $\mathrm{SL}_n(\mathbb{C})$.

Exercise 1.12. *Let* $(\partial^n + a_1\partial^{n-1} + \cdots + a_{n-1}\partial + a_n)f = 0$ *be a scalar linear equation of order* n. *Show that for solutions* $f_1, \cdots, f_n$ *the Wronskian* $W(f_1, \cdots, f_n)$ *as defined by the determinant*

$$\begin{vmatrix} f_1 & f_2 & f_3 & \cdots & f_n \\ \partial f_1 & \partial f_2 & \partial f_3 & \cdots & \partial f_n \\ \partial^2 f_1 & \partial^2 f_2 & \partial^2 f_3 & \cdots & \partial^2 f_n \\ \vdots & \vdots & \vdots & & \vdots \\ \partial^{n-1} f_1 & \partial^{n-1} f_2 & \partial^{n-1} f_3 & \cdots & \partial^{n-1} f_n \end{vmatrix}$$

is a solution of the differential equation $(\partial + a_1)\, W(f_1, \cdots, f_n) = 0$. *In turn conclude that the Wronskian vanishes if and only if the solutions* $f_1, \cdots, f_n$ *are linearly dependent.*

Exercise 1.13. *Show that the Euler–Gauss hypergeometric equation (introduced in Exercise 1.3)* $[z(1-z)\partial^2 + (\gamma - (\alpha + \beta + 1)z)\partial - \alpha\beta)]f(z) = 0$ *or equivalently* $[(\theta + \gamma - 1)\theta - z(\theta + \alpha)(\theta + \beta)]f(z) = 0$ *is a Fuchsian equation on* $\mathbb{P}$ *with regular singular points at* $z = 0, 1, \infty$. *Show that the exponents are* $0, 1 - \gamma$ *at* $z = 0$, *and* $0, \gamma - (\alpha + \beta)$ *at* $z = 1$, *and* α, β *at* $z = \infty$.

Exercise 1.14. *Compute the singular points and local exponents of the Fuchsian equation*

$$[z(z-a)(z-b)\partial^2 + (3z^2/2 - az - bz + ab/2)\partial + (Az + B)]f = 0$$

for $a, b, A, B \in \mathbb{C}$ *with* $ab(a - b) \neq 0$.

Exercise 1.15. *(for those who have seen elliptic functions) Fix* $\tau \in \mathbb{C}$ *with* $\Im(\tau) > 0$. *Let* $\Lambda = \mathbb{Z} + \mathbb{Z}\tau$ *be the lattice in* $\mathbb{C}$ *generated by* 1 *and* τ, *and let* $\mathbb{E} = \mathbb{C}/\Lambda$ *be the corresponding genus one Riemann surface. Addition on* $\mathbb{C}$ *induces a holomorphic group structure on* E, *denoted* $+$, *and* $(\mathbb{E}, +)$ *is called an elliptic curve. The associated Weierstrass elliptic function* $\mathfrak{p}(z)$, *defined by*

$$\mathfrak{p}(z) = \frac{1}{z^2} + \sum_{\lambda \in \Lambda, \lambda \neq 0} \left\{ \frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right\} ,$$

induces a holomorphic map $\mathfrak{p} : \mathbb{E} \to \mathbb{P}$, *which is generically* $2:1$ *and has the* 4 *points of subgroup* $\{z \in \mathbb{E}; 2z = 0\}$ *as ramification points. For details we refer to the text book of Whittaker and Watson [45].*

The Lamé equation is the second order scalar equation on $\mathbb{E}$ *given by*

$$(\partial^2 - n(n+1)\mathfrak{p}(z) - b)f(z) = 0$$

with $n \in \mathbb{N}, b \in \mathbb{C}$ *and as usual* $\partial = d/dz$ *with* z *the canonical coordinate on* $\mathbb{E}$ *coming from* $\mathbb{C}$. *Compute the the singular points and local exponents of the Lamé equation. Show that the Lamé equation is invariant under the automorphism* $z \mapsto -z$, *and conclude that the Lamé equation is the pull back under* $\mathfrak{p} : \mathbb{E} \to \mathbb{P}$ *of a second order scalar Fuchsian equation on* $\mathbb{P}$. *Determine the singular points and local exponents of the latter equation.*

Exercise 1.16. *Show that for a Fuchsian differential equation*

$$(\partial^n + a_1\partial^{n-1} + \cdots + a_{n-1}\partial + a_n)f = 0$$

of order n *with* $m \geq 2$ *regular singular points (including* ∞*) the sum of all local exponents at the* m *singular points is equal* $(m-2)n(n-1)/2$. *Hint: Show that for a singular point* z *the sum of the exponents at* z *is equal to minus the residue of* a_1 *at* z *plus* $\pm n(n-1)/2$ *(with* $+$ *is* $z \neq \infty$ *and* $-$ *if* $z = \infty$*).*

Exercise 1.17. *Consider the Clausen-Thomae hypergeometric equation*

$$[(\theta + \beta_1 - 1) \cdots (\theta + \beta_n - 1) - z(\theta + \alpha_1) \cdots (\theta + \alpha_n)]f(z) = 0$$

with so called numerator parameters $\alpha = (\alpha_1, \cdots, \alpha_n)$ *and denominator parameters* $\beta = (\beta_1, \cdots, \beta_n)$. *Show that* $z = 0, 1, \infty$ *are the only singular points. Show that they are regular singular with exponents* $1-\beta_1, \cdots, 1-\beta_n$ *at* $z = 0$, *exponents* $\alpha_1, \cdots, \alpha_n$ *at* $z = \infty$ *and exponents* $0, 1, \cdots, (n-2)$ *and* $\gamma = \sum_1^n (\beta_j - \alpha_j) - 1$ *at* $z = 1$.

Exercise 1.18. *Show that the point $z = \infty$ is a regular point of the second order linear differential equation*

$$[\partial^2 + a_1(z)\partial + a_2(z)]f(z) = 0$$

if and only if $a_1(z) = 2z^{-1} + O(z^{-2}), a_2(z) = O(z^{-4})$ for $z \to \infty$. Conclude that the most general second order Fuchsian equation on $\mathbb{P}$ with m distinct regular singular points at $z_1, \cdots, z_m \in \mathbb{C}$ has the form

$$[\partial^2 + \frac{G_{m-1}(z)}{F(z)}\partial + \frac{G_{2m-4}(z)}{F(z)^2}]f(z) = 0$$

with $F(z) = (z - z_1)\cdots(z - z_m)$ and $G_{m-1}, G_{2m-4} \in \mathbb{C}[z]$ polynomials in z of degrees $= (m-1)$ and $\leq (2m-4)$ respectively with leading coefficient of G_{m-1} equal to 2.

Exercise 1.19. *Show that the most general form of a second order Fuchsian equation on $\mathbb{P}$ with m distinct regular singular points $z_1, \cdots, z_m \in \mathbb{C}$ and exponents α_j, β_j at $z = z_j$, which satisfy $\sum_1^m(\alpha_j + \beta_j) = (m-2)$, is of the form*

$$[\partial^2 + \{\sum_1^m \frac{1 - \alpha_j - \beta_j}{(z - z_j)}\}\partial + \frac{1}{F(z)}\{\sum_1^m \frac{F_j(z_j)\alpha_j\beta_j}{(z - z_j)} + G_{m-4}(z)\}]f(z) = 0$$

with $F(z) = (z - z_1)\cdots(z - z_m)$ and $F_j(z) = F(z)/(z - z_j)$. Finally $G_{m-4} \in \mathbb{C}[z]$ is a polynomial of degree $\leq (m-4)$. The $(m-3)$ coefficients of G_{m-4} are called the accesory parameters.

Exercise 1.20. *Prove that for $m = 3$ the above differential equation is completely determined by the three singular points z_1, z_2, z_3 and the exponents at z_1, z_2, z_3 (which are restricted to sum up to 1). The explicit form becomes*

$$[\partial^2 + \{\sum_1^3 \frac{1 - \alpha_j - \beta_j}{(z - z_j)}\}\partial + \frac{1}{F(z)}\{\sum_1^3 \frac{F_j(z_j)\alpha_j\beta_j}{(z - z_j)}\}]f(z) = 0$$

with $F(z) = (z - z_1)(z - z_2)(z - z_3)$ and is called the Riemann–Papperitz hypergeometric equation [31], [28]. A differential equation with no accesory parameters is called a rigid equation, and the Riemann–Papperitz hypergeometric equation is a standard example of a rigid equation.

2 The Euler–Gauss hypergeometric function

2.1 The hypergeometric function of Euler–Gauss

The Euler–Gauss hypergeometric equation, introduced by Euler in the 18^{th} century, is the second order linear differential equation on the projective line $\mathbb{P} = \mathbb{C} \sqcup \{\infty\}$ of the form

$$[z(1-z)\partial^2 + (\gamma - (\alpha+\beta+1)z)\partial - \alpha\beta)]f = 0 \ ,$$

or equivalently

$$[(\theta+\gamma-1)\theta - z(\theta+\alpha)(\theta+\beta)]f = 0 \ .$$

Here as before $\partial = d/dz$, $\theta = z\partial$ and z is a complex variable. The numbers α, β, γ are called the parameters of the hypergeometric equation. It is a Fuchsian equation with regular singular points at $z = 0, 1, \infty$. The local exponents of the hypergeometric equation are given by the so called Riemann scheme

0	1	∞
0	0	α
$1-\gamma$	$\gamma-(\alpha+\beta)$	β

The first line gives the three singular points and the next two lines the exponents at the three singular points.

The Euler–Gauss hypergeometric function with parameters $\alpha, \beta, \gamma \in \mathbb{C}$ (but $\gamma \notin -\mathbb{N}$) is defined as the power series

$$F(\alpha,\beta,\gamma;z) = {}_2F_1(\alpha,\beta,\gamma;z) = \sum_{k=0}^{\infty} \frac{(\alpha)_k(\beta)_k}{(\gamma)_k \ k!} z^k$$

with

$$(\alpha)_k = \alpha(\alpha+1)\cdots(\alpha+k-1) = \Gamma(\alpha+k)/\Gamma(\alpha)$$

the Pochhammer symbol. Its domain of convergence is equal to the unit disc $\mathbb{D}$, unless α or β is a negative integer, in which case the series terminates and converges on all of $\mathbb{C}$. It is the unique holomorphic solution of the hypergeometric differential equation around $z = 0$ (an easy verification), normalized to be 1 at $z = 0$. In other words, the hypergeometric function is the normalized solution of the hypergeometric equation around $z = 0$ with exponent 0.

Besides the differential equation and the power series there is yet a third way of defining the hypergeometric function by means of a contour integral, obtained by Euler in 1748.

Theorem 2.1. *For* $0 < \Re(\beta) < \Re(\gamma)$ *the hypergeometric function is given by the Euler integral*

$$F(\alpha,\beta,\gamma;z) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)}\int_0^1 t^{\beta-1}(1-t)^{\gamma-\beta-1}(1-zt)^{-\alpha}dt$$

with $z \in \mathbb{D}$.

Proof. Note that the condition $0 < \Re(\beta) < \Re(\gamma)$ ensures the convergence of the integral. Moreover the integral defines an analytic continuation from $\mathbb{D}$ to $\mathbb{C} - [1,\infty)$. The theorem is an immediate consequence of the binomial series

$$(1-w)^{-\alpha} = \sum_{k=0}^{\infty} (\alpha)_k w^k/k!$$

for $w \in \mathbb{D}$ and the Euler Beta integral

$$\mathrm{B}(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$$

for $\Re(\alpha), \Re(\beta) > 0$. Details are left to the reader. □

A direct corollary of the Euler integral is the exact evaluation of the Gauss hypergeometric series at $z = 1$, a result of Gauss from 1812.

Theorem 2.2. *If* $\Re(\gamma-\alpha-\beta) > 0$ *then*

$$F(\alpha,\beta,\gamma;1) = \frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)}$$

which is called the Gauss summation formula.

Proof. Using the Euler integral formula we obtain

$$F(\alpha,\beta,\gamma;1) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)}\int_0^1 t^{\beta-1}(1-t)^{\gamma-\alpha-\beta-1}dt$$

which is valid for $\Re\beta > 0$ and $\Re(\gamma-\alpha-\beta) > 0$. So the Gauss summation formula is clear from the Euler Beta integral formula. □

The hypergeometric equation

$$[z(1-z)\partial^2 + (\gamma-(\alpha+\beta+1)z)\partial - \alpha\beta)]f = 0$$

is the unique second order Fuchsian equation with regular singular points at $\{0,1,\infty\}$ and with the given Riemann scheme. This fundamental insight of Riemann was verified in Exercise 1.19. Hence the hypergeometric series $F(\alpha-\gamma+1,\beta-\gamma+1,2-\gamma;z)$ is a solution of the hypergeometric equation with Riemann scheme

0	1	∞
0	0	$\alpha - \gamma + 1$
$\gamma - 1$	$\gamma - (\alpha + \beta)$	$\beta - \gamma + 1$

Therefore the function $z^{1-\gamma}F(\alpha - \gamma + 1, \beta - \gamma + 1, 2 - \gamma; z)$ is a solution of a hypergeometric equation with Riemann scheme

0	1	∞
0	0	α
$1 - \gamma$	$\gamma - (\alpha + \beta)$	β

which is the Riemann scheme of our original hypergeometric equation. This shows that the formal series

$$z^{1-\gamma}F(\alpha - \gamma + 1, \beta - \gamma + 1, 2 - \gamma; z)$$

is the upto scalar unique solution of the original hypergeometric equation around $z = 0$ with exponent $(1 - \gamma)$. By a similar reasoning (going back to Riemann in 1857) we obtain the following result of Kummer from 1836.

Theorem 2.3. *The solution space of the hypergeometric equation*

$$[z(1-z)\partial^2 + (\gamma - (\alpha + \beta + 1)z)\partial - \alpha\beta)]f = 0$$

has a basis of the form

$$F(\alpha, \beta, \gamma; z)$$

$$z^{1-\gamma}F(\alpha - \gamma + 1, \beta - \gamma + 1, 2 - \gamma; z)$$

around the point $z = 0$,

$$F(\alpha, \beta, \alpha + \beta - \gamma + 1; 1 - z)$$

$$(1 - z)^{\gamma-\alpha-\beta}F(\gamma - \alpha, \gamma - \beta, \gamma - \alpha - \beta + 1; 1 - z)$$

around the point $z = 1$,

$$(-z)^{-\alpha}F(\alpha, \alpha - \gamma + 1, \alpha - \beta + 1; 1/z)$$

$$(-z)^{-\beta}F(\beta, \beta - \gamma + 1, \beta - \alpha + 1; 1/z)$$

around the point $z = \infty$. Here the parameters α, β, γ are restricted such that the various hypergeometric series are well defined. For example, the first solution round $z = 0$ is defined for $\gamma \notin -\mathbb{N}$, the second solution around $z = 0$ is defined for $(2 - \gamma) \notin -\mathbb{N} \Leftrightarrow \gamma \notin \mathbb{N} + 2$, while they are linearly independent if $\gamma \notin \mathbb{Z}$. These solutions of the hypergeometric equation are called Kummer solutions.

Proof. The reader is invited to do the singular point $z = \infty$ as an exercise. Because of the symmetry $\alpha \leftrightarrow \beta$ one of the two is sufficient. $\square$

The following result goes under the name Kummer continuation formula.

Theorem 2.4. *Analytic continuation in the variable z from 0 to $-\infty$ along the negative real axis yields the identity*

$$F(\alpha, \beta, \gamma; z) = \frac{\Gamma(\gamma)\Gamma(\beta - \alpha)}{\Gamma(\beta)\Gamma(\gamma - \alpha)}(-z)^{-\alpha}F(\alpha, \alpha - \gamma + 1, \alpha - \beta + 1; 1/z)$$
$$+\frac{\Gamma(\gamma)\Gamma(\alpha - \beta)}{\Gamma(\alpha)\Gamma(\gamma - \beta)}(-z)^{-\beta}F(\beta, \beta - \gamma + 1, \beta - \alpha + 1; 1/z)$$

under the parameter restrictions $\gamma \notin -\mathbb{N}$ and $(\alpha - \beta) \notin \mathbb{Z}$.

Proof. Because of the Kummer solutions of the hypergeometric equation around $z = \infty$ and the symmetry $\alpha \leftrightarrow \beta$ we have a formula

$$F(\alpha, \beta, \gamma; z) = c(\alpha, \beta, \gamma)(-z)^{-\alpha}F(\alpha, \alpha - \gamma + 1, \alpha - \beta + 1; 1/z)$$
$$+c(\beta, \alpha, \gamma)(-z)^{-\beta}F(\beta, \beta - \gamma + 1, \beta - \alpha + 1; 1/z)$$

by analytic continuation along the negative real axis. Here $\gamma \notin -\mathbb{N}$ in order that the solution on the left hand side is well defined, and $\alpha - \beta \notin \mathbb{Z}$ in order that the two solutions on the right hand side are well defined and linearly independent. The coefficient $c(\alpha, \beta, \gamma)$ is holomorphic in the parameters as long as $(\alpha - \beta) \notin \mathbb{Z}$ and $\gamma \notin -\mathbb{N}$.

For $\Re(\alpha - \beta) < 0$ we multiply the above relation by $(-z)^{\alpha}$ and take the limit for $z \to \infty$ along the negative real axis. Using the Euler integral for $F(\alpha, \beta, \gamma; z)$ this yields

$$c(\alpha, \beta, \gamma) = \lim_{z \to \infty} \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma - \beta)} \int_0^1 t^{\beta - 1}(1 - t)^{\gamma - \beta - 1}(t - 1/z)^{-\alpha} dt$$
$$= \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma - \beta)} \int_0^1 t^{\beta - \alpha - 1}(1 - t)^{\gamma - \beta - 1} dt = \frac{\Gamma(\gamma)\Gamma(\beta - \alpha)}{\Gamma(\beta)\Gamma(\gamma - \alpha)}$$

by the Euler Beta integral. The theorem follows by analytic coninuation in the parameters. $\square$

The Kummer continuation formula gives the asymptotic behavior of the hypergeometric series $F(\alpha, \beta, \gamma; z)$ along the negative real axis. It is an important ingredient for the solution of the singular Sturm-Liouville problem (using Weyl-Titchmarsh theory [42]) of the hypergeometric operator on the unbounded interval $(-\infty, 0]$. In turn Harish-Chandra used this formula to derive the Plancherel formula for noncompact Riemannian symmetric spaces of rank one [17], [18].

2.2 The monodromy according to Schwarz–Klein

The question posed and solved by Schwarz in 1873 was: For which of the parameter values $\alpha, \beta, \gamma \in \mathbb{Q}$ are the solutions of the hypergeometric equation algebraic functions of its variable z?

The essential ingredient for the proof is the concept of monodromy, that was introduced by Riemann in his fundamental paper from 1857 on the hypergeometric equation [31]. It turns out that the solutions of the hypergeometric equation are algebraic if and only if the monodromy group of this equation is finite. Schwarz gave a beautiful alternative description of the (projective) monodromy group of the hypergeometric equation using the reflection principle, that he invented exactly for this purpose [33]. Subsequently Klein extended the work of Schwarz to deal not only with finite monodromy groups acting on the (elliptic) Riemann sphere, but also with infinite monodromy groups acting on the (parabolic) Euclidean plane and the (hyperbolic) Poincaré disc [24]. In turn this gave a boost to the theory of automorphic forms and functions.

The local exponents of the hypergeometric equation are given by the Riemann scheme

0	1	∞
0	0	α
$1-\gamma$	$\gamma-(\alpha+\beta)$	β

The exponent differences at the three singular points $0, 1, \infty$ are defined up to a sign choice, and given by

$$\kappa = \pm(\gamma - 1)\ ,\ \lambda = \pm(\alpha + \beta - \gamma)\ ,\ \mu = \pm(\alpha - \beta)$$

respectively. If we shift the parameters (α, β, γ) by integers (such a shift will be called a contiguity), then the (κ, λ, μ) are shifted by integers with even sum. Conversely, any shift of (κ, λ, μ) by integers with even sum arises from a contiguity.

Let us assume that the parameters (α, β, γ) and hence also (κ, λ, μ) are real numbers. By suitable contiguity we may assume $-1 < \kappa, \lambda, \mu \leq 1$ and by suitable sign choices we get $0 \leq \kappa, \lambda, \mu \leq 1$. After suitable permutation we may assume $0 \leq \kappa \leq \lambda \leq \mu \leq 1$, and hence also $\kappa + \lambda \leq \kappa + \mu \leq \lambda + \mu$. In case $\lambda + \mu > 1$ we replace (κ, λ, μ) by $(\kappa, 1 - \lambda, 1 - \mu)$ through contiguity and sign choices. The conclusion is that through contiguity and sign choices we can assume

$$0 \leq \kappa, \lambda, \mu \quad \text{and} \quad \kappa + \lambda, \kappa + \mu, \lambda + \mu \leq 1$$

and such a triple (κ, λ, μ) is called reduced. In fact, reduced triples are a normal formal for $(\kappa, \lambda, \mu) \in \mathbb{R}^3$ for contiguity and sign choices.

Theorem 2.5. *For $(\kappa, \lambda, \mu) \in \mathbb{R}^3$ the matrix*

$$G = \begin{pmatrix} 2 & -2\cos(\kappa\pi) & -2\cos(\mu\pi) \\ -2\cos(\kappa\pi) & 2 & -2\cos(\lambda\pi) \\ -2\cos(\mu\pi) & -2\cos(\lambda\pi) & 2 \end{pmatrix}$$

up to conjugation by the Klein four group $\mathrm{V}_4 = \{\mathrm{diag}(\pm 1, \pm 1, \pm 1)\} \cap \mathrm{SL}(3, \mathbb{R})$ is an invariant of triples $(\kappa, \lambda, \mu) \in \mathbb{R}^3$ up to contiguity and sign choices. The determinant of G is given by

$$\det G = -2 \prod \cos((\kappa \pm \lambda \pm \mu)\pi/2)$$

as a product of 4 factors.

Assume from now on that $(\kappa, \lambda, \mu) \in \mathbb{R}^3$ is reduced. For $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ in $\mathbb{R}^3$ denote $\langle x, y \rangle = xGy^{\mathrm{t}}$ as scalar product on $\mathbb{R}^3$. Then this scalar product is Euclidean for $\kappa + \lambda + \mu > 1$ and Lorentzian for $\kappa + \lambda + \mu < 1$.

Assume from now on that $\kappa, \lambda, \mu > 0$ and $\kappa + \lambda + \mu < 1$. Then the dual basis $\varepsilon_1, \varepsilon_2, \varepsilon_3$ of the standard basis $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$ of $\mathbb{R}^3$ lies in a single connected component $\mathbb{R}^3_+$ of $\{x \in \mathbb{R}^3; \langle x, x \rangle < 0\}$, and the closed convex cone D spanned by $\varepsilon_1, \varepsilon_2, \varepsilon_3$ lies in $\mathbb{R}^3_+ \sqcup \{0\}$. Let $s_i(x) = x - \langle x, e_i \rangle e_i$ be the orthogonal reflection with mirror the plane spanned by the two ε_j for $j \neq i$. Let W be the subgroup of the Lorentz group $\mathrm{O}(\mathbb{R}^3, \langle \cdot, \cdot \rangle)$ generated by the reflections s_1, s_2, s_3 in the faces of D. Then the union over $w \in W$ of the cones wD is equal to $\mathbb{R}^3_+ \sqcup \{0\}$, and under the Schwarz conditions

$$\kappa = 1/k, \lambda = 1/l, \mu = 1/m \quad \text{for } k, l, m \in \mathbb{Z} \text{ and } \geq 2$$

the cones wD for $w \in W$ form a regular tesselation of $\mathbb{R}^3_+ \sqcup \{0\}$.

Proof. The first statement is obvious, and the calculation of $\det G$ is a straightforward calculation. From this determinant formula it follows that for a reduced triple (κ, λ, μ) we have $\det G = 0$ if and only if $\kappa + \lambda + \mu = 1$, and $\kappa + \lambda + \mu > 1$ or < 1 if and only if the scalar product is Euclidean or Lorentzian respectively.

The last paragraph is the three dimensional version of a theorem of Tits for the so called geometric construction of Coxeter groups, see for example the lecture notes [19] on my website. □

Now let us pick two linearly independent solutions f_1, f_2 on the upper half plane $\mathbb{H} = \{\Im(z) > 0\}$, and consider the projective evaluation map (also called the Schwarz map)

$$\mathrm{Pev} : \mathbb{H} \to \mathbb{P}\ ,\ \mathrm{Pev}(z) = f_1(z)/f_2(z)$$

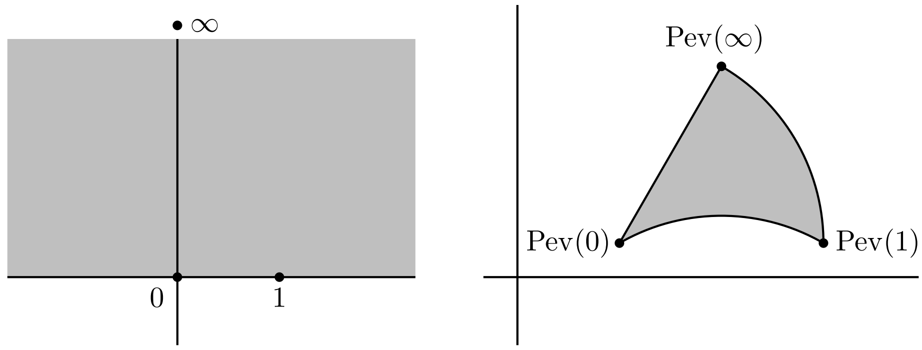
with $\mathbb{P}$ the complex projective line. Because of the ambiguity of the base choice f_1, f_2 the Schwarz map is only canonical upto action of $\mathrm{Aut}(\mathbb{P})$. We claim that the the Schwarz map Pev maps the upper half plane $\mathbb{H}$ conformally onto the interior of a triangle with sides circular arcs, and with angles $\kappa\pi$, $\lambda\pi$ and $\mu\pi$ at the vertices $\mathrm{Pev}(0)$, $\mathrm{Pev}(1)$ and $\mathrm{Pev}(\infty)$ respectively. This circular triangle is called the Schwarz triangle of the hypergeometric equation.

The Schwarz map is conformal because its derivative

$$\partial(\mathrm{Pev}) = \frac{\partial(f_1)f_2 - f_1\partial(f_2)}{f_2^2}$$

vanishes nowhere. Indeed the numerator is the Wronskian, which does not vanish, because f_1, f_2 are linearly independent solutions on $\mathbb{H}$. In order to understand the image of the Schwarz map we look at its behaviour on the real axis as boundary of $\mathbb{H}$.

For example, for the boundary interval $(0, 1)$ we can choose the solutions f_1, f_2 to be real on $(0, 1)$. This is possible because the hypergeometric equation is a real differential equation, since the parameters α, β, γ were assumed to be real numbers. In that case the image of the interval $(0, 1)$ under the Schwarz map is a real interval. For a general choice of f_1, f_2 the image of $(0, 1)$ is the transform under an element of $\mathrm{Aut}(\mathbb{P})$, so a fractional linear transformation, of a real interval, and therefore equal to a real interval or a circular arc.



The angles of the Schwarz triangle at the vertices Pev(0), Pev(1) and Pev(∞) are equal to $\kappa\pi$,$\lambda\pi$ and $\mu\pi$ respectively. For example, near the origin 0 let us choose the solutions f_1, f_2 of the form

$$f_1(z) = (1 + \cdots)\ ,\ f_2(z) = z^{1-\gamma}(1 + \cdots)$$

which in turn implies that

$$f_1(z)/f_2(z) = z^{\kappa}(1 + \cdots)$$

which indeed gives an angle $\kappa\pi$ at the vertex Pev(0) of the Schwarz triangle. For a general choice of f_1, f_2 this angle $\kappa\pi$ is conserved by some fractional linear transformation.

By continuity we can extend the Schwarz map

$$\mathrm{Pev} : \mathbb{H} \sqcup (-\infty, 0) \sqcup (0, 1) \sqcup (1, \infty) \to \mathbb{P}$$

with image the Schwarz triangle minus its vertices. The key step in the argument of Schwarz is the beautiful insight that the analytic continuation of Pev is given by the reflection principle. Indeed, there are three possibilities for analytic continuation from the upper half plane $\mathbb{H}$ to the lower half plane $-\mathbb{H}$, namely through the intervals $(-\infty, 0)$, $(0, 1)$ and $(1, \infty)$. The analytic continuation of the Schwarz map is obtained by reflecting the Schwarz triangle in the corresponding sides. Now we can iterate the above construction with the new triangle, which allows one to understand the full analytic continuation of the Schwarz map, step by step reflecting in sides of circular triangles. The domain of this full analytic continuation is the universal covering space $\widetilde{Z}$ of $Z = \mathbb{P} - \{0, 1, \infty\}$, say relative to the base point $z_0 = 1/2$, and we write

$$\widetilde{\mathrm{Pev}} : \widetilde{Z} \to \mathbb{P}$$

for the analytic continuation of the Schwarz map as a univalued map.

The range of this map can get messy, as the triangles start overlapping. However in case the Schwarz triangle is dihedral, which means that

$$\kappa = 1/k,\ \lambda = 1/l,\ \mu = 1/m$$

for some integers $k, l, m \geq 2$, we do get a regular tesselation by congruent images of the Schwarz triangle. These conditions on the parameters are called the Schwarz integrality conditions. The range $\mathbb{G}$ of this tesselation is equal to

$$\mathbb{G} = \mathbb{P},\ \mathbb{C},\ \mathbb{D}$$

upto an action of Aut($\mathbb{P}$), depending on whether the angle sum of the Schwarz triangle

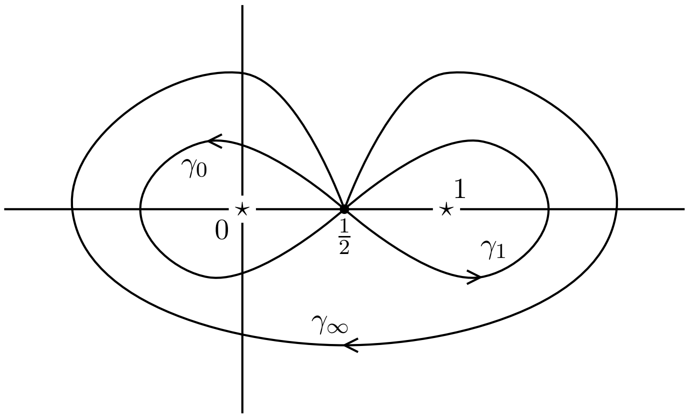
$$(\kappa + \lambda + \mu)\pi = (1/k + 1/l + 1/m)\pi$$

is greater than π, equal to π, or smaller than π respectively. Here

$$\mathbb{D} = \{w \in \mathbb{C}; |w| < 1\}$$

denotes the unit disc. In this last case the disc $\mathbb{D}$ is bounded by a circle (Klein's Orthogonalkreis) which is orthogonal to the three circles bounding the Schwarz triangle. Hence the range $\mathbb{G}$ of the Schwarz map equals the Riemann sphere $\mathbb{P}$, the Euclidean plane $\mathbb{C}$ or the Poincaré disc $\mathbb{D}$ respectively. Note that in all three cases $\mathbb{G}$ is simply connected.

The image of the analytically continued Schwarz map $\widetilde{\text{Pev}} : \widetilde{Z} \to \mathbb{G}$ is equal to $\mathbb{G}$ minus all vertices of the triangular tesselation. These vertices can be filled in by the following construction.



The fundamental group Π of $Z = \mathbb{P} - \{0, 1, \infty\}$ with base point $z_0 = 1/2$ has three generators $g_0 = [\gamma_0], g_1 = [\gamma_1], g_\infty = [\gamma_\infty]$ with a single relation $g_\infty g_1 g_0 = 1$ as indicated in the above picture. Under the above Schwarz integrality conditions the Schwarz map factors through the intermediate covering

$$\widehat{\text{Pev}} : \widehat{Z} = \Pi(k, l, m)\backslash\widetilde{Z} \to \mathbb{G}$$

with $\Pi(k, l, m)$ be the normal subgroup of Π generated by g_0^k, g_1^l and g_∞^m. The group

$$\Gamma = \Gamma(k, l, m) \simeq \Pi/\Pi(k, l, m)$$

is the projective monodromy group of the hypergeometric equation. At this level we can lift the compactification $Z \hookrightarrow Z^+ = \mathbb{P}$ to a partial compactification $\widehat{Z} \hookrightarrow \widehat{Z}^+$, resulting in a commutative diagram

$$\begin{array}{ccccc} \widehat{Z} & \longrightarrow & \widehat{Z}^+ & \xrightarrow{\widehat{\mathrm{Pev}}^+} & \mathbb{G} \\ \downarrow & & \downarrow & & \downarrow \\ Z = \mathbb{P} - \{0,1,\infty\} & \longrightarrow & Z^+ = \mathbb{P} & \xrightarrow{\mathrm{Pev}^+} & \Gamma\backslash\mathbb{G} \end{array}$$

The left vertical arrow is an unramified Γ-covering, while the middle vertical arrow is a ramified covering, branched of orders k, l and m above the points 0, 1 and ∞ respectively. The extended Schwarz map

$$\widehat{\mathrm{Pev}}^+ : \widehat{Z}^+ \to \mathbb{G}$$

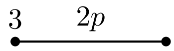
becomes an unramified covering. Since $\widehat{Z}^+$ is connected and $\mathbb{G}$ simply connected we conclude that the Schwarz map yields a conformal isomorphism between $\widehat{Z}^+$ and $\mathbb{G}$. In other words the projective monodromy group $\Gamma(k,l,m) \cong \Pi/\Pi(k,l,m)$ acts on $\mathbb{G}$ with quotient $\Gamma(k,l,m)\backslash\mathbb{G} \cong \mathbb{P}$. This quotient map is given by the inverse of the Schwarz map, and is ramified above $0,1,\infty$ of orders k,l,m respectively. The projective monodromy group $\Gamma(k,l,m)$ is a subgroup of $\mathrm{Aut}(\mathbb{P})$, and is called the Schwarz triangle group. The group $W(k,l,m)$ generated by the reflections in the sides of the Schwarz triangle is called the Coxeter triangle group. It consists of holomorphic and antiholomorphic transformations of $\mathbb{P}$. The Schwarz triangle group is the index two subgroup of the Coxeter triangle group, consisting of even products of reflections in the sides of the Schwartz triangle.

Algebraic hypergeometric functions appear in case the monodromy group is finite, and for rational parameters α,β,γ this is equivalent with the projective monodromy group $\Gamma(k,l,m) \simeq \Pi/\Pi(k,l,m)$ being finite. For the integers $k,l,m \geq 2$ this amounts to $1/k + 1/l + 1/m > 1$. In that case the order n of $\Gamma(k,l,m)$ is given by $1/k + 1/l + 1/m - 1 = 2/n$ as there are $2n$ triangles of area $(1/k + 1/l + 1/m - 1)\pi$ needed to tesselate the unit sphere of area 4π. The results of this section are essentially due to Schwarz [33] and Klein [24], based on earlier ideas of Riemann [31].

Example 2.6. *For $p = 3,4,5$ the hypergeometric function*

$$F((p+6)/(12p), (p-6)/(12p), 2/3; z)$$

is an algebraic function with projective monodromy group Γ equal to the three Platonic rotation groups A_4,S_4,A_5 of tetrahedron, octahedron, icosahedron *respectively. Indeed the exponent differences are* $1/3, 1/2, 1/p$ *at the points* $0, 1, \infty$ *respectively. The monodromy around* 0 *and* 1 *is a complex reflection of order* 3 *and* 2 *respectively. The linear monodromy group is the finite complex reflection group with Coxeter diagram*



in the notation of Coxeter [8],[35]. The order of the group $\Gamma(3, 2, p)$ *is equal to* $12p/(6-p)$ *for* $p = 3, 4, 5$ *and indeed* $12, 24, 60$ *is the order of* A_4,S_4,A_5.

2.3 The Euler integral revisited

Let us fix four rational parameters $\mu_0, \mu_1, \mu_2, \mu_3 \in (0, 1)$ with $\sum \mu_j = 2$. In addition choose four distinct complex variables z_0, z_1, z_2, z_3. If we clear denominators and write $\mu_j = m_j/m$ with $m, m_j \in \mathbb{N}$ (so $\sum m_j = 2m$) and $\gcd(m, m_0, m_1, m_2, m_3) = 1$ then the multivalued differential

$$\omega = \frac{dx}{y} \ \ , \ \ y = \prod (x - z_j)^{\mu_j}$$

on $\mathbb{P} - \{z_0, z_1, z_2, z_3\}$ becomes a univalued holomorphic differential on the Riemann surface

$$C : \ y^m = \prod (x - z_j)^{m_j}$$

lying above $\mathbb{P}$ as a m-fold ramified covering via the map $(x, y) \mapsto x$. This covering map is just the quotient map for the action of the group C_m of the order m roots of unity on C (by multiplication in the variable y). Upto a multiplicative scalar the holomorphic differential ω on C is unique characterized by the transformation behaviour $\omega \mapsto \zeta^{-1}\omega$ if $y \mapsto \zeta y$ for $\zeta \in C_m$. Integrals of the form

$$\pi = \int_{z_i}^{z_j} \omega$$

along suitable curves on C (whose projection on $\mathbb{P}$ apart from begin and end points avoids z_0, z_1, z_2, z_3) are called period integrals.

An element of $\mathrm{Aut}(\mathbb{P})$ transforms the quadruple z_0, z_1, z_2, z_3 and the corresponding Riemann surface C into isomorphic objects. Without loss of generality we can take $z_0 = 0, z_1 = z, z_2 = 1, z_3 = \infty$ with $z \in \mathbb{P} - \{0, 1, \infty\}$. If we integrate from 0 to 1 then the period integral becomes

$$\int_0^1 \omega \ , \ \omega = \frac{dx}{x^{\mu_0}(x - z)^{\mu_1}(x - 1)^{\mu_2}}$$

and apart from Γ-factors, a factor $(-z)^{\mu_1}(-1)^{\mu_2}$ and a substition $z \mapsto 1/z$ this becomes the Euler integral with parameters

$$\mu_0 = 1-\beta,\ \mu_1 = \alpha,\ \mu_2 = 1+\beta-\gamma,\ \mu_3 = \gamma-\alpha$$

as functions of the parameters α, β, γ in the Euler integral. In turn we get

$$(1-\mu_0-\mu_1) = (\beta-\alpha), (1-\mu_0-\mu_2) = (\gamma-1), (1-\mu_0-\mu_3) = (\alpha+\beta-\gamma)\ .$$

Hence the Schwarz integrality conditions

$$\kappa = |\gamma-1|, \lambda = |\alpha+\beta-\gamma|, \mu = |\alpha-\beta| \in 1/\mathbb{N}$$

amount in the new parameters $\mu_0, \cdots, \mu_3$ to

$$(1-\mu_i-\mu_j) \in 1/\mathbb{N}$$

for all $i \neq j$ with $\mu_i + \mu_j < 1$.

Let us assume that the Schwarz integrality conditions do hold. If we write

$$\pi_1(z) = \int_0^1 \omega\ ,\ \pi_2(z) = \int_1^z \omega$$

then the Schwarz projective evaluation map

$$z \mapsto \mathrm{Per}(z) = \pi_1(z)/\pi_2(z)$$

becomes a period map, which we emphasize by writing Per instead of Pev. The Schwarz–Klein theory of the previous section has the following modular interpretation.

Theorem 2.7. *For $\underline{m} = (m_0, \cdots, m_3)$ a quadruple of positive relatively prime integers with $\sum m_j = 2m$ let $\mathcal{M}(\underline{m})$ denote the moduli space of isomorphism classes of curves of the form $y^m = \prod (x-z_j)^{m_j}$ with an ordering on the four points $(z_0, \cdots, z_3)$. If the Schwartz conditions*

$$|1-(m_i+m_j)/m| = 1/p_{ij}\ ,\ p_{ij} \in \{2, 3, \cdots, \infty\}\ \ \forall i \neq j$$

do hold then we have an injective locally biholomorphic period map

$$\mathrm{Per} : \mathcal{M}(\underline{m}) \to \Gamma(k, l, m)\backslash\mathbb{G}$$

with $k = p_{02}, l = p_{03}, m = p_{01}$. The image of the period map is obtained from $\mathbb{G}$ by deleting the three vertices of the Schwartz triangle.

Example 2.8. *In the setting of Example 2.6 with* $p \in \{2, 3, \cdots\}$ *and*

$$\mu_0 = \frac{7p+6}{12p}, \mu_1 = \frac{p-6}{12p}, \mu_2 = \frac{5p+6}{12p}, \mu_3 = \frac{11p-6}{12p}$$

we consider the coarse moduli space of curves

$$\mathcal{M}(p) = \{y^{12p} = (x-z_0)^{7p+6}(x-z_1)^{p-6}(x-z_2)^{5p+6}(x-z_3)^{11p-6}\}/_{\cong}$$

with an ordering on the four points $(z_0, \cdots, z_3) = (0, z, 1, \infty)$. *Here we assume that* p *is not divisible by* 2 *or* 3 *so that the exponents are relatively prime. Indeed a common divisor divides* $\gcd(p+6, 2p, 12)$ *which equals* 1 *if* p *is not divisible by* 2 *or* 3. *This assumption implies that the above curve is connected. The Schwarz–Klein theorem gives a period isomorphism of orbifolds*

$$\mathrm{Per} : \mathcal{M}(p) \to \Gamma(3, 2, p)\backslash\mathbb{G}$$

with $\mathbb{G} = \mathbb{P}$ *if* $p = 5$ *and* $\mathbb{G} = \mathbb{D}$ *for* $p \geq 7$.

Example 2.9. *On can even allow* $p = \infty$ *corresponding to the moduli space*

$$\mathcal{M}(\infty) = \{y^{12} = (x-z_0)^7(x-z_1)(x-z_2)^5(x-z_3)^{11}\}/_{\cong}$$

of a particular class of curves of genus $g = 11$ *(using the additivety of the Euler characteristic). In turn we have the period isomorphism*

$$\mathrm{Per} : \mathcal{M}(\infty) \to \Gamma\backslash\mathbb{H}$$

with $\Gamma = \Gamma(3, 2, \infty) \cong \mathrm{PSL}(2, \mathbb{Z})$ *the modular group. Using the isomorphism* $\mathcal{M}(\infty) \ni (0, z, 1, \infty) \leftrightarrow z \in \mathbb{P} - \{0, 1, \infty\}$ *the inverse map*

$$J : \Gamma\backslash\mathbb{H} \to \mathbb{C}$$

is the modular invariant of Klein. After rescaling by 12^3 *we get the famous modular function* j *with Fourier expansion on the cusp [34]*

$$j(q) = 12^3 J(q) = \sum_{n=-1}^{\infty} c(n)q^n \ , \ q = e^{2\pi i\tau}$$

with $\tau \in \mathbb{H}$ *and* $c(n) \in \mathbb{Z}$ *and* $c\_1 = 1, c_0 = 744, c_1 = 196884, c_2 = 21493760, \cdots$.

2.4 Exercises

Exercise 2.1. *Check the Kummer relations in Theorem 2.3 around the point* $z = \infty$.

Exercise 2.2. *Show the Kummer relation*

$$F(\alpha, \beta, \gamma; z) = (1 - z)^{\gamma-\alpha-\beta} F(\gamma - \alpha, \gamma - \beta, \gamma; z)$$

using Riemann schemes.

Exercise 2.3. *Show using Riemann schemes that*

$$F(\alpha, \beta, 1 + \alpha - \beta; z) = (1 - z)^{-\alpha} F(\tfrac{1}{2}\alpha, \tfrac{1}{2} + \tfrac{1}{2}\alpha - \beta, 1 + \alpha - \beta; \frac{-4z}{(1 - z)^2})$$

and conclude that

$$F(\alpha, \beta, 1 + \alpha - \beta; -1) = \frac{\Gamma(1 + \alpha - \beta)\Gamma(1 + \frac{1}{2}\alpha)}{\Gamma(1 + \alpha)\Gamma(1 + \frac{1}{2}\alpha - \beta)}$$

which is a result of Kummer.

Exercise 2.4. *Show that for given* $2 \le k \le l \le m$ *the projective monodromy group* $\Gamma(k, l, m) = \Pi/\Pi(k, l, m)$ *is finite if and only if* $k = l = 2, m \ge 2$ *or* $k = 2, l = 3, m = 3, 4, 5$. *Show that* $\Gamma(2, 2, m)$ *is isomorphic to the dihedral group* D_m *of order* $2m$.

Exercise 2.5. *Show that the order* N *of the finite group* $\Gamma(3, 2, p)$ *for* $p = 3, 4, 5$ *is given by* $N = 12p/(6 - p)$, *which is the order of the rotation symmetry groups of the tetrahedron, octahedron and icosahedron respectively.*

Exercise 2.6. *The modular group* $\mathrm{PSL}_2(\mathbb{Z})$ *acts on the upper halfplane* $\mathbb{H}$ *by fractional linear transformations. Using that* $\mathrm{PSL}_2(\mathbb{Z})$ *is generated by the two transformations*

$$S : z \mapsto -1/z \ , \ T : z \mapsto z + 1$$

show that $\mathrm{PSL}_2(\mathbb{Z}) \cong \Gamma(3, 2, \infty)$. *Hint: Consider the Schwarz triangle*

$$\{z; -1/2 \le \Re z \le 0, |z| \ge 1\}$$

with vertices in the extended upper half plane $\mathbb{H} \sqcup \mathbb{Q} \sqcup \{\infty\}$ *at the points* $\omega = (-1 + \sqrt{-3})/2$, $i = \sqrt{-1}$ *and* ∞ *with corresponding angles* $\pi/3$, $\pi/2$ *and* 0.

Exercise 2.7. *For (M, g) a Riemannian manifold the Laplace–Beltrami operator Δ is a second order linear differential operator on M, given in local coordinates $(x_1, \cdots, x_n)$ with Riemannian metric $ds^2 = \sum g_{ij} dx_i dx_j$ by the expression*

$$\Delta(f) = g^{-1/2} \sum \partial_i (g^{1/2} g^{ij} \partial_j f)$$

for all smooth functions f on M. Here $\partial_i = \partial/\partial x_i$ are the partial derivatives, g^{ij} is the inverse matrix of g_{ij} and $g = \det(g_{ij})$. Let μ be the Riemannian measure on M, given locally by $d\mu(x) = g^{1/2} dx_1 \cdots dx_n$.

Show that the Laplace–Beltrami operator Δ on M is a second order linear differential operator, with leading symbol the dual Riemannian metric $\sum g^{ij} \partial_i \partial_j$, which is symmetric for the Hermitian inner product $\langle f_1, f_2 \rangle = \int_M f_1(x) \overline{f_2(x)} d\mu(x)$, and satisfies $\Delta(1) = 0$. In fact these three properties characterize the Laplace–Beltrami operator: second order terms are given by the Riemannian metric, first order terms follow by the requirement of being symmetric, and the constant term follows from $\Delta(1) = 0$.

Exercise 2.8. *Let $S^n = \{x \in \mathbb{R}^{n+1}; x_1^2 + \cdots + x_{n+1}^2 = 1\}$ be the unit sphere of dimension $n \geq 1$, and consider $S^1 \hookrightarrow S^n$ via*

$$(\sin\varphi, \cos\varphi) \mapsto (0 \cdots, 0, \sin\varphi, \cos\varphi).$$

A smooth function f on S^n is called zonal, if the value $f(x_1, \cdots, x_{n+1})$ only depends on the last coordinate x_{n+1}. In other, words a zonal function f on S^n is determined by its restriction $\varphi \mapsto \tilde{f}(\varphi)$ to S^1, which is still invariant under $\varphi \mapsto -\varphi$. The radial part $\tilde{\Delta}$ of the Laplace–Beltrami operator Δ for zonal functions on S^n is defined as the second order linear differential operator on S^1, characterized by $\tilde{\Delta}(\tilde{f}) = \widetilde{\Delta(f)}$ for all zonal functions f on S^n. Using the previous exercise show that

$$\tilde{\Delta} = (\sin\varphi)^{1-n} \circ \frac{\mathrm{d}}{\mathrm{d}\varphi} \circ (\sin\varphi)^{n-1} \circ \frac{\mathrm{d}}{\mathrm{d}\varphi} = \frac{\mathrm{d}^2}{\mathrm{d}\varphi^2} + (n-1)\frac{\cos\varphi}{\sin\varphi}\frac{\mathrm{d}}{\mathrm{d}\varphi}$$

acting on functions on S^1, which are even (invariant under $\varphi \mapsto -\varphi$).

Exercise 2.9. *Show that the zonal eigenvalue problem for Δ on S^n in the form*

$$\tilde{\Delta}(\tilde{f}) - s(n-1-s)\tilde{f} = 0$$

can be transformed via the substitution $z = \sin^2(\varphi/2), f(z) = \tilde{f}(\varphi)$ into the hypergeometric equation with parameters $\alpha = s, \beta = n-1-s, \gamma = n/2$. Conclude that the spectrum of the Laplace–Beltrami operator on the unit *sphere* S^n *is equal to* $\{k(n-1-k); k = 0, -1, -2, \cdots\}$. *Using abstract harmonic analysis the eigenspaces can be shown to be irreducible unitary representations of the orthogonal group* $\mathrm{O}(\mathbb{R}^{n+1})$.

Exercise 2.10. *Carry out the same analysis for Lobachevsky space, which is the upper sheet* $H^n = \{x \in \mathbb{R}^{n,1}; x_{n+1}^2 = 1 + x_1^2 + \cdots + x_n^2, x_{n+1} > 0\}$ *of the hyperboloid, using* $H^1 \hookrightarrow H^n$ *via*

$$(\sinh t, \cosh t) \mapsto (0 \cdots, 0, \sinh t, \cosh t)$$

as before. Here $\mathbb{R}^{n,1}$ *is just equal to* $\mathbb{R}^{n+1}$ *as real vector space, but with the Lorentzian scalar product* $\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n - x_{n+1}y_{n+1}$. *In this way the embedding* $H^n \hookrightarrow \mathbb{R}^{n,1}$ *turns* H^n *into a Riemannian manifold of constant negative sectional curvature* -1. *Again the zonal eigenvalue problem for* Δ *on* H^n *in the form*

$$\tilde{\Delta}(\tilde{f}) + s(n-1-s)\tilde{f} = 0$$

can be transformed via the substitution $z = -\sinh^2(t/2), f(z) = \tilde{f}(t)$ *into the hypergeometric equation with parameters* $\alpha = s, \beta = n-1-s, \gamma = n/2$. *Using Weyl–Titchmarch theory for the hypergeometric equation on* $(-\infty, 0]$ *the outcome is that the "tempered" spectrum of* Δ *on* H^n *is equal to*

$$\{-s(n-1-s); \Re(s) = (n-1)/2\} = (-\infty, -(n-1)^2/4] \ .$$

In this way Harish-Chandra determined the explicit form of the Plancherel theorem for Lobachevsky space H^n. *Using abstract harmonic analysis the eigenspaces can be shown to be irreducible unitary (infinite dimensional) representations of the Lorentz group* $\mathrm{O}(\mathbb{R}^{n,1})$.

3 The Clausen–Thomae hypergeometric function

3.1 The hypergeometric function of Clausen–Thomae

The generalized hypergeometric series was introduced for $n = 3$ by Clausen [7] and in general by Thomae [41]. Let

$$\alpha = (\alpha_1, \cdots, \alpha_n), \beta = (\beta_1, \cdots, \beta_n) \in \mathbb{C}^n$$

be complex parameters with $\beta_n = 1$, and let us assume that $\beta_j \notin -\mathbb{N}$ for all j. The power series

$$F(\alpha; \beta|z) = {}_nF_{n-1}(\alpha; \beta|z) = \sum_{k=0}^{\infty} \frac{(\alpha_1)_k \cdots (\alpha_n)_k}{(\beta_1)_k \cdots (\beta_n)_k} z^k$$

is called the Clausen–Thomae hypergeometric series. The α_j are called numerator parameters and the β_j denominator parameters. Since $\beta_n = 1$ we get $(\beta_n)_k = k!$ and for $n = 2$ one recovers the Euler–Gauss hypergeometric series.

The hypergeometric series converges on the unit disc $\mathbb{D}$ and is a solution of the hypergeometric equation

$$[z(\theta + \alpha_1) \cdots (\theta + \alpha_n) - (\theta + \beta_1 - 1) \cdots (\theta + \beta_n - 1)]f = 0$$

with $\theta = zd/dz$ as before. This equation has regular singular points at $0, 1, \infty$ with local exponents given by the Riemann scheme

0	1	∞
$1 - \beta_j$	$0, 1, \cdots, n-2, \gamma$	α_j

by a direct computation with $\gamma = -1 + \sum_1^n (\beta_j - \alpha_j)$. This follows from the next theorem (after a substitution $z \mapsto z - 1$), which goes back to Pochhammer.

Theorem 3.1. *If we have given the linear differential equation*

$$(\partial^n + a_1 \partial^{n-1} + \cdots + a_{n-1}\partial + a_n)f = 0$$

with the functions $z \mapsto za_j(z)$ holomorphic on the unit disc $\mathbb{D}$ for all j and $\lim_{z\to 0} za_1 \notin -\mathbb{N}$ *then there exist $n-1$ linearly independent holomorphic solutions f on $\mathbb{D}$ with $\partial^{j-1} f(0) = f_{0j}$ freely prescribed for $j = 1, \cdots, n-1$.*

Proof. Let $A(z)$ be the matrix valued holomorphic function

$$\begin{pmatrix} 0 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 \\ a_n & a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_2 & a_1 \end{pmatrix}$$

and rewrite the scalar differential equation in matrix form $(\partial + A)F = 0$ as in Theorem 1.1. This vector solution is of the form $F = (f_1, \cdots, f_n)^t$ with $f_{j+1} = \partial f_j$ and $f = f_1$ a solution of the original scalar differential equation of order n. Multiplication by z gives the matrix form $(\theta + B)F = 0$ with $B(z) = zA(z)$ holomorphic on $\mathbb{D}$. If $B(z) = \sum B_k z^k$ then

$$\mathrm{Ker}(B_0) = \{F_0 = (f_{01}, \cdots, f_{0n})^t \in \mathbb{C}^n; \sum b_{0(n+1-j)} f_{0j} = 0\}$$

with $b_{0j} = \lim_{z\to 0} z a_j$. Since $b_{01} \neq 0$ the components f_{0j} of $F_0 \in \mathrm{Ker}(B_0)$ can be freely prescribed for $j = 1, \cdots, n-1$. Since $\det(k + 1 + B_0) = (k+1)^{n-1}(k + 1 + b_{01}) \neq 0$ for all $k \in \mathbb{N}$ by our assumption the recurrence relations of Proposition 1.26 can be solved for each $F_0 \in \mathrm{Ker}(B_0)$. The conclusion is that the original scalar equation of order n has holomorphic solutions on $\mathbb{D}$ with $\partial^{j-1} f(0) = f_{0j}$ freely prescribed for $j = 1, \cdots, n-1$. □

The Clausen–Thomae hypergeometric equation is the unique differential equation of order n with this special property at $z = 1$ and the above Riemann scheme of local exponents. It is free of accesory parameters, and as such a rigid differential equation. The proof is left to the reader as Exercise 3.2. The Euler integral formula has a natural generalization from the Euler–Gauss to the Clausen–Thomae hypergeometric function.

Theorem 3.2. *If $\Re(\beta_i) > \Re(\alpha_i) > 0$ for $i = 1, \cdots, n-1$ then the Clausen–Thomae hypergeometric function $F(\alpha;\beta|z)$ is given by*

$$\prod_{i=1}^{n-1} \frac{\Gamma(\beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i - \alpha_i)} \int_0^1 \cdots \int_0^1 \frac{\prod_1^{n-1} t_i^{\alpha_i - 1}(1-t_i)^{\beta_i - \alpha_i - 1}}{(1 - z t_1 \cdots t_{n-1})^{\alpha_n}} dt_1 \cdots dt_{n-1}$$

which for $n = 2$ boils down to the Euler integral formula. Note that the Euler integral breaks the permutation symmetry in the numerator parameters $\alpha_1, \cdots, \alpha_n$ and singles out α_n.

Proof. Substitute the binomial series

$$(1 - zt_1 \cdots t_{n-1})^{-\alpha_n} = \sum_{k=0}^{\infty} \frac{\Gamma(\alpha_n + k)}{\Gamma(\alpha_n)k!}(t_1 \cdots t_{n-1})^k z^k$$

in the integral formule and use the Euler Beta integral

$$\int_0^1 t_i^{\alpha_i+k-1}(1-t_i)^{\beta_i-\alpha_i-1} dt_i = \Gamma(\alpha_i + k)\Gamma(\beta_i - \alpha_i)/\Gamma(\beta_i + k)$$

to conclude that

$$F(\alpha, \beta; z) = \sum_{k=0}^{\infty} \frac{\Gamma(\alpha_n + k)}{\Gamma(\alpha_n)k!} \prod_{i=1}^{n-1} \frac{\Gamma(\alpha_i + k)\Gamma(\beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i + k)} z^k$$

and the result follows. □

The Kummer basis of local solutions around $z = \infty$ with exponents α_i all distinct modulo $\mathbb{Z}$ is given by

$$(-z)^{-\alpha_i} F((\alpha_i + 1)\underline{1} - \beta; (\alpha_i + 1)\underline{1} - \alpha | 1/z)$$

with $\underline{1} = (1, \cdots, 1) \in \mathbb{C}^n$. Note that one of the denominator parameters (although not the last but the i^{th} one) equals 1 as should. Let us check that all these functions have the Riemann scheme

0	1	∞
$\underline{1} - \beta$	$0, 1, \cdots, n-2, \gamma$	α

of the original $F(\alpha; \beta | z)$.

The Riemann scheme of $F((\alpha_i + 1)\underline{1} - \beta; (\alpha_i + 1)\underline{1} - \alpha | z)$ is equal to

0	1	∞
$-\alpha_i \underline{1} + \alpha$	$0, 1, \cdots, n-2, \gamma$	$(\alpha_i + 1)\underline{1} - \beta$

with the parameter

$$\gamma = -1 + \sum_{j=1}^{n} ((\alpha_i + 1) - \alpha_j - (\alpha_i + 1) + \beta_j) = -1 + \sum_{j=1}^{n} (\beta_j - \alpha_j)$$

the same as for the original hypergeometric function $F(\alpha; \beta | z)$. Since z gets replaced by $1/z$ one has to interchange the exponents at $z = 0$ and $z = \infty$ and so the Riemann scheme of $F((\alpha_i + 1)\underline{1} - \beta; (\alpha_i + 1)\underline{1} - \alpha | 1/z)$ becomes

0	1	∞
$(\alpha_i+1)\underline{1}-\beta$	$0,1,\cdots,n-2,\gamma$	$-\alpha_i\underline{1}+\alpha$

Finally the term $(-z)^{-\alpha_i}$ amounts to adding $-\alpha_i\underline{1}$ at $z=0$ and adding $\alpha_i\underline{1}$ at $z=\infty$ to the exponents. The final outcome is the Riemann scheme

0	1	∞
$\underline{1}-\beta$	$0,1,\cdots,n-2,\gamma$	α

of the original $F(\alpha;\beta|z)$.

Likewise the Kummer continuation formula also can be generalized.

Theorem 3.3. *Analytic continuation from* 0 *to* $-\infty$ *along the negative real axis yields the identity*

$$F(\alpha;\beta|z)=\sum_{i=1}^{n}\frac{\prod_{j=1}^{n-1}\Gamma(\beta_j)\prod_{j\neq i}\Gamma(\alpha_j-\alpha_i)}{\prod_{j=1}^{n-1}\Gamma(\beta_j-\alpha_i)\prod_{j\neq i}\Gamma(\alpha_j)}$$

$$\times(-z)^{-\alpha_i}F((\alpha_i+1)\underline{1}-\beta;(\alpha_i+1)\underline{1}-\alpha|1/z)$$

under the parameter restrictions $\beta_j\notin-\mathbb{N}$ *and* $(\alpha_i-\alpha_j)\notin\mathbb{Z}$ *for* $i\neq j$.

Proof. The coefficient of $(-z)^{-\alpha_i}F((\alpha_i+1)\underline{1}-\beta;(\alpha_i+1)\underline{1}-\alpha|1/z)$ in case $i=n$ can be derived from the Euler integral formula just like in the case $n=2$. Indeed for $\Re(\alpha_i-\alpha_n)>0$ for $i=1,\cdots,n-1$ this coefficient becomes

$$\lim_{z\to-\infty}(-z)^{\alpha_n}F(\alpha;\beta|z)=\lim_{z\to-\infty}\prod_{i=1}^{n-1}\frac{\Gamma(\beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i-\alpha_i)}$$

$$\int_0^1\cdots\int_0^1\frac{\prod_1^{n-1}t_i^{\alpha_i-1}(1-t_i)^{\beta_i-\alpha_i-1}}{(t_1\cdots t_{n-1}-1/z)^{\alpha_n}}dt_1\cdots dt_{n-1}=$$

$$\prod_1^{n-1}\frac{\Gamma(\beta_j)}{\Gamma(\alpha_j)\Gamma(\beta_j-\alpha_j)}\prod_1^{n-1}\int_0^1 t_j^{\alpha_j-\alpha_n-1}(1-t_j)^{\beta_j-\alpha_j-1}dt_j$$

which under the assumption $\Re(\alpha_j-\alpha_n),\Re(\beta_j-\alpha_j)>0$ for $j=1,\cdots,n-1$ simplifies to

$$\prod_1^{n-1}\frac{\Gamma(\beta_j)\Gamma(\alpha_j-\alpha_n)}{\Gamma(\alpha_j)\Gamma(\beta_j-\alpha_n)}$$

by the Euler Beta integral. The theorem follows by analytic continuation in the parameters and the permutation symmetry among the numerator parameters. ☐

So, many of the results for the Euler–Gauss hypergeometric function have a natural generalization to the Clausen–Thomae hypergeometric function. There is however one exception, namely the Gauss summation formula

$$_2F_1(\alpha, \beta; \gamma|1) = \frac{\Gamma(\gamma)\Gamma(\gamma - \alpha - \beta)}{\Gamma(\gamma - \alpha)\Gamma(\gamma - \beta)}$$

in case $\gamma \notin -\mathbb{N}$ and $\Re(\gamma - \alpha - \beta) > 0$. An explicit evaluation of the Clausen–Thomae hypergeometric function at $z = 1$ as a product of Γ factors does not seem possible in general for $n \geq 3$. Only for special restrictions on the parameters such an explicit summation is known, but the general structure behind these special restrictions remains mysterious [3]. Just to be impressed we mention some of these, for example Saalschütz theorem says

Theorem 3.4. *If α, β or γ is a nonpositive integer (and so the hypergeometric series terminates) and $\alpha + \beta + \gamma + 1 = \delta + \varepsilon$ then*

$$_3F_2(\alpha, \beta, \gamma; \delta, \varepsilon|1) = \frac{\Gamma(\delta)\Gamma(1 + \alpha - \varepsilon)\Gamma(1 + \beta - \varepsilon)\Gamma(1 + \gamma - \varepsilon)}{\Gamma(1 - \varepsilon)\Gamma(\delta - \alpha)\Gamma(\delta - \beta)\Gamma(\delta - \gamma)}$$

In case $\gamma = -n \in -\mathbb{N}$ (and $\delta \mapsto \gamma$ and $\varepsilon \mapsto 1 + \alpha + \beta - \gamma - n$) this theorem can be rewritten in the equivalent form

Theorem 3.5. *For $\gamma, 1 + \alpha + \beta - n - \gamma \notin -\mathbb{N}$ we have*

$$_3F_2(\alpha, \beta, -n; \gamma, 1 + \alpha + \beta - n - \gamma|1) = \frac{(\gamma - \alpha)_n(\gamma - \beta)_n}{(\gamma)_n(\gamma - \alpha - \beta)_n}$$

Proof. Comparison of the coefficients of z^n in the Kummer relation

$$(1 - z)^{\alpha+\beta-\gamma}{}_2F_1(\alpha, \beta; \gamma|z) = {}_2F_1(\gamma - \alpha, \gamma - \beta, \gamma|z)$$

(derived in Exercise 2.2) yields the relation

$$\sum_{k=0}^{n} \frac{(\alpha)_k(\beta)_k}{(\gamma)_k \; k!} \frac{(\gamma - \alpha - \beta)_{n-k}}{(n - k)!} = \frac{(\gamma - \alpha)_n(\gamma - \beta)_n}{(\gamma)_n \; n!}$$

and hence also

$$\sum_{k=0}^{n} \frac{(\alpha)_k(\beta)_k}{(\gamma)_k k!} \frac{(\gamma - \alpha - \beta)_n(-n)_k}{(1 + \alpha + \beta - \gamma - n)_k n!} = \frac{(\gamma - \alpha)_n(\gamma - \beta)_n}{(\gamma)_n \; n!}$$

which in turn implies

$$\sum_{k=0}^{n} \frac{(\alpha)_k(\beta)_k(-n)_k}{(\gamma)_k(1 + \alpha + \beta - \gamma - n)_k \; k!} = \frac{(\gamma - \alpha)_n(\gamma - \beta)_n}{(\gamma)_n(\gamma - \alpha - \beta)_n}$$

and proves Saalschütz's formula. □

Dixon's theorem gives another case for which the ${}_3F_2$ at unit argument can be summed.

Theorem 3.6. *For* $(1+\alpha-\beta), (1+\alpha-\gamma) \notin -\mathbb{N}$ *and* $\Re(1+\alpha-2\beta-2\gamma) > 0$ *we have*

$$
{}_3F_2(\alpha,\beta,\gamma;1+\alpha-\beta,1+\alpha-\gamma|1) = \\
\frac{\Gamma(1+\frac{1}{2}\alpha)\Gamma(1+\alpha-\beta)\Gamma(1+\alpha-\gamma)\Gamma(1+\frac{1}{2}\alpha-\beta-\gamma)}{\Gamma(1+\alpha)\Gamma(1+\frac{1}{2}\alpha-\beta)\Gamma(1+\frac{1}{2}\alpha-\gamma)\Gamma(1+\alpha-\beta-\gamma)}
$$

Proof. The Gauss summation formula gives

$$
{}_2F_1(\beta+n,\gamma+n;1+\alpha+2n|1) = \frac{\Gamma(1+\alpha+2n)\Gamma(1+\alpha-\beta-\gamma)}{\Gamma(1+\alpha-\beta+n)\Gamma(1+\alpha-\gamma+n)}
$$

and therefore

$$
\frac{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)\ {}_3F_2(\alpha,\beta,\gamma;1+\alpha-\beta,1+\alpha-\gamma|1)}{\Gamma(1+\alpha-\beta)\Gamma(1+\alpha-\gamma)}
$$

$$
= \sum_{n=0}^{\infty} \frac{\Gamma(\alpha+n)\Gamma(\beta+n)\Gamma(\gamma+n)}{n!\Gamma(1+\alpha-\beta+n)\Gamma(1+\alpha-\gamma+n)}
$$

$$
= \sum_{n=0}^{\infty} \frac{\Gamma(\alpha+n)\Gamma(\beta+n)\Gamma(\gamma+n)\ {}_2F_1(\beta+n,\gamma+n;1+a+2n|1)}{n!\Gamma(1+\alpha+2n)\Gamma(1+\alpha-\beta-\gamma)}
$$

$$
= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{\Gamma(\alpha+n)\Gamma(\beta+n+m)\Gamma(\gamma+n+m)}{n!m!\Gamma(1+\alpha+2n+m)\Gamma(1+\alpha-\beta-\gamma)}
$$

$$
= \sum_{k=0}^{\infty}\sum_{n=0}^{k} \frac{\Gamma(\alpha+n)\Gamma(\beta+k)\Gamma(\gamma+k)}{n!(k-n)!\Gamma(1+\alpha+n+k)\Gamma(1+\alpha-\beta-\gamma)}
$$

$$
= \sum_{k=0}^{\infty} \frac{\Gamma(\alpha)\Gamma(\beta+k)\Gamma(\gamma+k)\ {}_2F_1(\alpha,-k;1+\alpha+k|-1)}{k!\Gamma(1+\alpha+k)\Gamma(1+\alpha-\beta-\gamma)}
$$

$$
= \sum_{k=0}^{\infty} \frac{\Gamma(\alpha)\Gamma(\beta+k)\Gamma(\gamma+k)\Gamma(1+\frac{1}{2}\alpha)}{k!\Gamma(1+\alpha)\Gamma(1+\alpha-\beta-\gamma)\Gamma(1+\frac{1}{2}\alpha+k)}
$$

$$
= \frac{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)\ {}_2F_1(\beta,\gamma;1+\frac{1}{2}\alpha|1)}{\Gamma(1+\alpha)\Gamma(1+\alpha-\beta-\gamma)}
$$

$$
= \frac{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)\Gamma(1+\frac{1}{2}\alpha)\Gamma(1+\frac{1}{2}\alpha-\beta-\gamma)}{\Gamma(1+\alpha)\Gamma(1+\alpha-\beta-\gamma)\Gamma(1+\frac{1}{2}\alpha-\beta)\Gamma(1+\frac{1}{2}\alpha-\gamma)}
$$

using the definition of the terminating series

$$ {}_2F_1(\alpha,-k;1+\alpha+k|z)=\sum_{n=0}^{k}\frac{\Gamma(\alpha+n)\Gamma(1+\alpha+k)k!}{\Gamma(\alpha)\Gamma(1+\alpha+k+n)n!(k-n)!}(-z)^n $$

and the Kummer summation formula (see Exercise 2.3)

$$ {}_2F_1(\alpha,-k;1+\alpha+k|-1)=\frac{\Gamma(1+\alpha+k)\Gamma(1+\frac{1}{2}\alpha)}{\Gamma(1+\alpha)\Gamma(1+\frac{1}{2}\alpha+k)} $$

and again the Gauss summation formula

$$ {}_2F_1(\beta,\gamma;1+\tfrac{1}{2}\alpha|1)=\frac{\Gamma(1+\frac{1}{2}\alpha)\Gamma(1+\frac{1}{2}\alpha-\beta-\gamma)}{\Gamma(1+\frac{1}{2}\alpha-\beta)\Gamma(1+\frac{1}{2}\alpha-\gamma)} $$

and checking the cancellations in numerator and denominator. □

The little book by Bailey on generalized hypergeometric series [3] is full of similar formulas all the way up to ${}_7F_6$ (Dougall's theorem). Ramanujan rediscovered many of these hypergeometric identities in his Notebooks. In recent times Doron Zeilberger devised a computer algorithm for proving and finding new identities of this kind. That was a good idea because these proofs are just unpleasant verifications.

But what is the point? Are these formulas useful to anybody? Probably only to few people. They give explicit evaluations of so called "periods", which are integrals of algebraic functions over regions, both defined over the field $\mathbb{Q}$. Periods are in general transcendental numbers, and that in special cases they have closed formulas as product of Γ-factors is truly remarkable. For further reading see the paper by Kontsevich and Zagier [25].

The following quadratic transformation formula is due to Whipple [44].

Theorem 3.7. *For $1+\alpha-\beta$ and $1+\alpha-\gamma$ not negative integers and $|z|$ sufficently small we have the quadratic transformation formula*

$$ {}_3F_2(\alpha,\beta,\gamma;1+\alpha-\beta,1+\alpha-\gamma|z)= $$
$$ (1-z)^{-\alpha}\,{}_3F_2(\tfrac{1}{2}\alpha,\tfrac{1}{2}+\tfrac{1}{2}\alpha,1+\alpha-\beta-\gamma;1+\alpha-\beta,1+\alpha-\gamma|\frac{-4z}{(1-z)^2}) $$

Proof. The degree two map $z\mapsto w=-4z/(1-z)^2$ sends $z=-1,0,1,\infty$ to $w=1,0,\infty,0$ respectively and has two ramification points $w=1,\infty$. The Riemann scheme of local exponents of the hypergeometric series on the right hand side becomes in the variable w

$w=0$	$w=1$	$w=\infty$
0	0	$\alpha/2$
$\beta-\alpha$	1	$1/2+\alpha/2$
$\gamma-\alpha$	$1/2$	$1+\alpha-\beta-\gamma$

and therefore in the variable z takes the form

$z=-1$	$z=0$	$z=1$	$z=\infty$
0	0	α	0
1	$\beta-\alpha$	$1+\alpha$	$\beta-\alpha$
2	$\gamma-\alpha$	$2+2\alpha-2\beta-2\gamma$	$\gamma-\alpha$

and taking the factor $(1-z)^{-\alpha}$ into account it becomes

$z=-1$	$z=0$	$z=1$	$z=\infty$
0	0	0	α
1	$\beta-\alpha$	1	β
2	$\gamma-\alpha$	$2+\alpha-2\beta-2\gamma$	γ

The point $z=-1$ is just a regular point, and so the first column can be deleted and the remaining part of the diagram is the Riemann scheme of the hypergeometric function ${}_3F_2(\alpha,\beta,\gamma;1+\alpha-\beta,1+\alpha-\gamma|z)$. The formula follows since both sides are the solutions around $z=0$ with local exponent 0 and value 1 at $z=0$. □

In the final chapter of his thesis Levelt gave a conceptual proof of Dixon's theorem using the above quadratic transformation formula [26], and which is outlined in Exercise 3.5.

3.2 The monodromy according to Levelt

Let $V(\alpha,\beta)$ be the local solution space at the base point $\frac{1}{2}$ and consider the monodromy representation

$$M(\alpha,\beta):\Pi\to\mathrm{GL}(V(\alpha,\beta))$$

with Π the fundamental group of $\mathbb{M}=\mathbb{P}-\{0,1,\infty\}$ with generators g_0,g_1,g_∞ and relation $g_\infty g_1 g_0=1$ as before. The monodromy group is the image of Π under the monodromy representation. It is generated by the elements

$$h_0=M(\alpha,\beta)(g_0)\ ,\ h_1=M(\alpha,\beta)(g_1)\ ,\ h_\infty=M(\alpha,\beta)(g_\infty)$$

satisfying the relation $h_\infty h_1 h_0=1$. The local exponents at $z=0$ and $z=\infty$ in the Riemann scheme imply that

$$\det(t-h_\infty)=(t-a_1)\cdots(t-a_n)\ ,\ \det(t-h_0^{-1})=(t-b_1)\cdots(t-b_n)$$

with

$$a_j = \exp(2\pi i \alpha_j)\ ,\ b_j = \exp(2\pi i \beta_j)$$

while the linear map

$$(h_1 - \mathrm{Id}) \in \mathrm{End}(V(\alpha, \beta))$$

has rank at most one by Theorem 3.1.

Theorem 3.8. *Let $n \geq 2$ and $H < \mathrm{GL}(\mathbb{C}^n)$ be a subgroup generated by two matrices A, B such that $\mathrm{rk}(A - B) \leq 1$. Then H acts irreducibly on $\mathbb{C}^n$ if and only if A and B have disjoint sets of eigenvalues.*

Proof. Suppose H acts reducibly on $\mathbb{C}^n$. Let V_1 be a nontrivial invariant subspace of $\mathbb{C}^n$ and let V_2 be $\mathbb{C}^n/V_1$. Since $\mathrm{rk}(A - B) \leq 1$ it follows that A and B coincide either on V_1 or on V_2. Hence A and B have a common eigenvalue.

Suppose A and B have a common eigenvalue λ. If $W = \mathrm{Ker}(A - B)$ then $\dim W \geq (n - 1)$ by assumption. If $\dim W = n$ then $A = B$ and H acts reducibly on $\mathbb{C}^n$. Therefore we may assume that $\dim W = (n - 1)$. If A has an eigenvector in W then it must also be an eigenvector for B, since A and B coincide on W. This common eigenvector generates an invariant subspace in $\mathbb{C}^n$ of dimension one. Hence H acts reducibly on $\mathbb{C}^n$. Therefore we may assume that neither A nor B have an eigenvector in W.

We claim that $V = (A - \lambda)\mathbb{C}^n$ is an invariant subspace for H. Clearly $AV = A(A - \lambda)\mathbb{C}^n = (A - \lambda)A\mathbb{C}^n = (A - \lambda)\mathbb{C}^n = V$ and so V is invariant under A. Since $\mathrm{Ker}(A - \lambda)$ is nontrivial and has trivial intersection with the codimension one subspace W the dimension of $\mathrm{Ker}(A - \lambda)$ is one. Hence the dimension of V is $n - 1$ and so $V = (A - \lambda)W$. Since A and B coincide on W we get $V = (B - \lambda)W$ and by a similar argument as for A we get $V = (B - \lambda)\mathbb{C}^n$ and V is invariant under B. Hence V is a nontrivial invariant subspace for H and the representation of H on $\mathbb{C}^n$ becomes reducible. □

The next algebraic characterization of the monodromy group of the Clausen–Thomae hypergeometric equation is due to Levelt [26].

Theorem 3.9. *Suppose that $a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathbb{C}^\times$ with $a_i \neq b_j$ for all $1 \leq i, j \leq n$. Up to simultaneous conjugation in $\mathrm{GL}(n, \mathbb{C})$ there exist unique elements $A, B \in \mathrm{GL}(n, \mathbb{C})$ with*

$$\det(t - A) = \prod_{j=1}^{n} (t - a_j)\ ,\ \det(t - B) = \prod_{j=1}^{n} (t - b_j)$$

while the matrix $A - B$ has rank one.

Proof. First we shall prove the existence of A and B. We have to find matrices A, B in $\mathrm{GL}(n, \mathbb{C})$ with

$$\prod_{j=1}^{n}(t-a_j) = t^n + A_1 t^{n-1} + \cdots + A_n$$
$$\prod_{j=1}^{n}(t-b_j) = t^n + B_1 t^{n-1} + \cdots + B_n$$

as their characteristic polynomials. If we take

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -A_n \\ 1 & 0 & \cdots & 0 & -A_{n-1} \\ 0 & 1 & \cdots & 0 & -A_{n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -A_1 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 0 & \cdots & 0 & -B_n \\ 1 & 0 & \cdots & 0 & -B_{n-1} \\ 0 & 1 & \cdots & 0 & -B_{n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -B_1 \end{pmatrix}$$

then an easy computation shows

$$\det(t-A) = t^n + A_1 t^{n-1} + \cdots + A_n \ , \ \det(t-B) = t^n + B_1 t^{n-1} + \cdots + B_n$$

and $\mathrm{rk}(A-B) = 1$ holds trivially. This proves the existence of A and B.

In order to prove the uniqueness of $A, B \in \mathrm{GL}(n, \mathbb{C})$ up to a simultaneous conjugation let $W = \mathrm{Ker}(A-B) \subset \mathbb{C}^n$. By assumption W has codimension one in $\mathbb{C}^n$. Hence

$$V = W \cap A^{-1}W \cap \cdots \cap A^{-(n-2)}W$$

has dimension at least one. For $v \in V$ a nonzero vector the elements $A^i v$ for $i = 0, 1, \cdots, n-2$ all lie in W, which in turn implies that $A^i v = B^i v$ for $i = 0, 1, \cdots, n-1$. By the Cayley–Hamilton theorem the linear span of $A^i v = B^i v$ for $i = 0, 1, \cdots, n-1$ is invariant under the group generated by A and B. Since $a_i \neq b_j$ the action of this group on $\mathbb{C}^n$ is irreducible by the previous theorem, which in turn implies that $A^i v = B^i v$ for $i = 0, 1, \cdots, n-1$ is a basis of $\mathbb{C}^n$. Relative to this basis the matrices of A and B have the above form. □

Under the irreducibility condition $a_i \neq b_j$ the monodromy group of the Clausen–Thomae hypergeometric equation is obtained by

$$h_\infty = A \ , \ h_0 = B^{-1} \ , \ h_1 = A^{-1}B$$

and we will denote this monodromy group by $H(a,b)$. Indeed the linear map $h_1 - \mathrm{Id} = A^{-1}(B-A)$ has rank one. A linear transformation $h \in \mathrm{GL}(\mathbb{C}^n)$ is called a (complex) reflection if $h - \mathrm{Id}$ has rank one. The distinguished property of the Clausen–Thomae hypergeometric equation is that under the irreducibility condition $a_i \neq b_j$ the monodromy $h_1 = M(\alpha,\beta)(g_1)$ around the point 1 is a reflection. This makes the Clausen–Thomae hypergeometric equation a rigid equation, in the sense that it is characterized among all Fuchsian equation of order n with regular singular points $\{0,1,\infty\}$ by its local exponents at the three singular points. For rigid Fuchsian equations the monodromy group should be determined in linear algebra terms by the characteristic polynomials of the monodromy operators around the various regular singular points. For other examples of rigid equations we refer to work by Simpson [36]. A description in algebraic geometric terms of all rigid equations is due to Katz [23].

Corollary 3.10. *Under the irreducibility condition $a_i \neq b_j$ the monodromy group $H(a,b)$ of the Clausen–Thomae hypergeometric equation is defined in a suitable basis by matrices with entries in the ring $\mathbb{Z}[A_i, B_j, 1/A_n, 1/B_n]$.*

This is clear from the proof of the above theorem since $\det A = \pm A_n$ and $\det B = \pm B_n$. The rigidity of the Clausen–Thomae hypergeometric equation enables one to derive certain results by just looking at the Riemann schemes. In the next example the proof of Clausen's formula gives an illustration of this idea.

Example 3.11. *Clausen's formula says that*

$$ {}_2F_1(\alpha,\beta,\alpha+\beta+1/2;z)^2 = {}_3F_2(2\alpha,2\beta,\alpha+\beta;2\alpha+2\beta,\alpha+\beta+1/2|z) $$

with the ${}_2F_1$ on the left hand side a second order Euler–Gauss hypergeometric function and the ${}_3F_2$ on the right hand side a third order Clausen–Thomae hypergeometric function. Clausen's formula can be proved by comparison of the two Riemann schemes. The Riemann scheme for the ${}_2F_1$ is given by

0	1	∞
0	0	α
$1/2-(\alpha+\beta)$	$1/2$	β

while the Riemann scheme for the ${}_3F_2$ equals

0	1	∞
0	0	2α
$1/2-(\alpha+\beta)$	1	2β
$1-2(\alpha+\beta)$	$1/2$	$\alpha+\beta$

Observe that the latter Riemann scheme is just the second symmetric square of the former Riemann scheme. Moreover near the point $z = 1$ there is a two dimensional subspace of holomorphic solutions, corresponding to the local exponents $0, 1$. This proves Clausen's formula, and this particular third order Clausen–Thomae hypergeometric equation is just the second symmetric square of this particular second order Euler–Gauss hypergeometric equation.

Using Example 2.6 it follows from Clausen's formula that the hypergeometric function

$$_3F_2((p+6)/6p, (p-6)/6p, 1/6; 1/3, 2/3|z)$$

is an algebraic function for $p = 3, 4, 5$. The monodromy groups are subgroups of the symmetry groups of the Platonic solids tetrahedron, octahedron and icosahedron respectively.

Such algebraic $_3F_2$ hypergeometric functions were the starting point for subsequent work by Beukers and Heckman [4] leading to a full classification of algebraic $_nF_{n-1}$ hypergeometric functions. The method will be explained in later sections.

Example 3.12. *The symmetric group S_{n+1} on $n+1$ letters acts on $\mathbb{C}^{n+1}$ by permutations of the coordinates. This action is reducible, but the restriction to the invariant linear subspace V of vectors with zero sum of coordinates is an irreducible representation, called the reflection representation of S_{n+1}. The nearest neighbour transpositions $s_i = (i\, i+1)$ for $i = 1, \cdots, n$ generate the symmetric group S_{n+1}. It is easy to see that the symmetric group S_{n+1} is also generated by the elements*

$$h_\infty = A = s_1 \cdots s_n \ , \ h_0^{-1} = B = s_1 \cdots s_{n-1} \ , \ h_1 = A^{-1}B = s_n$$

considered as elements of $\mathrm{GL}(V)$ and the relation $h_\infty h_1 h_0 = \mathrm{Id}$ is trivial. Moreover $A = (1\ 2\ \cdots n+1)$ and $B = (1\ 2 \cdots n)$ implies

$$\det(t - A) = t^n + t^{n-1} + \cdots + t + 1 \ , \ \det(t - B) = t^n - 1$$

and so the symmetric group S_{n+1} acting on V by the reflection representation is an example of a hypergeometric group.

3.3 The criterion of Beukers–Heckman

Throughout this section we have given $a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathbb{C}^\times$ for which the irreducibility condition

$$a_i \neq b_j$$

of the previous section holds. Let $H(a,b) < \mathrm{GL}(n,\mathbb{C})$ be the corresponding hypergeometric group acting irreducibly on $\mathbb{C}^n$. In this section we will discuss a criterion for finiteness of the hypergeometric group $H(a,b)$ due to Beukers and Heckman [4]. Independently similar results were obtained around the same time by Kontsevich, but they remained unpublished after he learned about our preprint.

Theorem 3.13. *If both sets $a = \{a_1, \cdots, a_n\}$ and $b = \{b_1, \cdots, b_n\}$ are invariant under the substitution $z \mapsto \overline{1/z}$ then there exists a nondegenerate Hermitian form of $\mathbb{C}^n$ which is invariant under the hypergeometric group $H(a,b)$.*

Proof. Let $A, B \in \mathrm{GL}_n(\mathbb{C})$ with $\mathrm{rk}(A - B) = 1$ be the generators of the hypergeometric group $H(a,b)$ as in Theorem 3.9. Let us denote $X^\dagger = \overline{X}^t$ for $X \in \mathrm{Mat}_n(\mathbb{C})$. Since $\mathrm{rk}(A^\dagger - B^\dagger) = 1$ it is clear that $A^\dagger, B^\dagger$ generate a hypergeometric group with parameter sets

$$\overline{a} = \{\overline{a}_1, \cdots \overline{a}_n\} \ , \ \overline{b} = \{\overline{b}_1, \cdots, \overline{b}_n\}$$

and likewise since $\mathrm{rk}(A^{-1} - B^{-1}) = 1$ it follows that A^{-1}, B^{-1} generate a hypergeometric group with

$$1/a = \{1/a_1, \cdots, 1/a_n\} \ , \ 1/b = \{1/b_1, \cdots, 1/b_n\}$$

as parameter sets.

If $\overline{a} = 1/a$ and $\overline{b} = 1/b$ we conclude from Levelt's theorem the existence of a matrix $s \in \mathrm{GL}_n(\mathbb{C})$ with

$$A^\dagger = sA^{-1}s^{-1} \quad , \quad B^\dagger = sB^{-1}s^{-1}$$

and so $s = A^\dagger s A = B^\dagger s B$. If $\langle \cdot, \cdot \rangle$ denotes the standard Hermitian inner product on $\mathbb{C}^n$ then the sesquilinear forms

$$(u,v) \mapsto \langle su, v \rangle \ , \ (u,v) \mapsto \langle u, sv \rangle$$

on $\mathbb{C}^n$ are nondegenerate and invariant under $H(a,b)$. Hence

$$(u,v) \mapsto (\langle su, v \rangle + \langle u, sv \rangle)/2 \ , \ (u,v) \mapsto (\langle su, v \rangle - \langle u, sv \rangle)/2i$$

are both invariant Hermitian forms on $\mathbb{C}^n$, and certainly one of them is nondegenerate. □

Theorem 3.14. *Suppose that $\{a_1, \cdots, a_n\}$ and $\{b_1, \cdots, b_n\}$ have all modulus one. Suppose the exponents α_i and β_j are contained in $(0, 1]$ and numbered by increasing argument. Let $m_j = \#\{i; \alpha_i < \beta_j\}$ for $j = 1, \cdots, n$. Then the signature (p, q) of the invariant Hermitian form for the hypergeometric group $H(a, b)$ is given by*

$$|p - q| = |\sum_{j=1}^{n}(-1)^{j+m_j}| \, .$$

In particular the invariant Hermitian form is definite if and only if the two sets

$$\{a_1, \cdots, a_n\} \, , \, \{b_1, \cdots, b_n\}$$

interlace on the unit circle.

A quick proof that the interlacing property of the two eigenvalue sets of A and B implies definiteness of the invariant Hermitian form goes as follows. First observe that the signature of the invariant Hermitian form does not change as long as a_i and b_j vary continuously over the unit circle while $a_i \neq b_j$ throughout the variation. Moreover any pair of interlacing eigenvalue value sets on the unit circle can be continuously deformed into any other such pair.

Hence it is sufficient to have just one example for each n of two eigenvalue sets, interlacing on the unit circle, for which the invariant Hermitian form is definite. But from Example 3.12 we know that the invariant Hermitian form is definite for the case

$$\det(t - A) = t^n + t^{n-1} + \cdots + t + 1 \, , \, \det(t - B) = t^n - 1$$

of the reflection representation of the symmetric group S_{n+1}. Clearly these two eigenvalue sets $\sqrt[n+1]{1} - \{1\}$ and $\sqrt[n]{1}$ interlace on the unit circle.

Proof. Throughout this proof we have as usual $A = h_\infty$, $B = h_0^{-1}$ and so $h_1 = h_\infty^{-1} h_0^{-1} = A^{-1}B$. We know that $h_1 - 1 = A^{-1}(B - A)$ has rank one, which by definition means that h_1 is a (complex) reflection with $\det(h_1) = c$ the special eigenvalue of h_1. Let $\langle \cdot, \cdot \rangle$ be the invariant Hermitian form on $\mathbb{C}^n$. Irreducibility of $H(a, b)$ implies that it is nondegenerate. If $c \neq 1$ then the unitary reflection h_1 is given by the formula

$$h_1(v) = v + (c - 1)\frac{\langle v, r \rangle}{\langle r, r \rangle} r$$

for some nonzero $r \in \mathbb{C}^n$. Indeed $h_1(v) = v$ if $\langle v, r\rangle = 0$ and $h_1(r) = cr$, and so r is just an eigenvector of h_1 with eigenvalue c. Our assumption $c \neq 1$ implies that $\langle r, r\rangle \neq 0$ as should in order that the formula makes sense.

The formula for h_1 can be rewritten in the form

$$(B - A)v = (c-1)\frac{\langle v, r\rangle}{\langle r, r\rangle} Ar$$

and using that for a rank one linear map D of the form $D(v) = \langle v, w\rangle u$ with $u, v, w \in \mathbb{C}^n$ the determinant of $1 + D$ is given by $\det(1 + D) = 1 + \langle u, w\rangle$ we get

$$\det((B-t)(A-t)^{-1}) = \det(1+(B-A)(A-t)^{-1}) = 1+(c-1)\frac{\langle (A-t)^{-1}Ar, r\rangle}{\langle r, r\rangle}$$

which in turn implies under the assumption that all a_i are distinct

$$\prod_{i=1}^{n} \frac{b_i - t}{a_i - t} = 1 + (c-1)\sum_{i=1}^{n} \frac{a_i}{a_i - t}\,\frac{\langle r_i, r\rangle}{\langle r, r\rangle}$$

with $r = \sum r_i$ and $Ar_i = a_i r_i$ the eigenvalue decomposition of r for A. Since $r = \sum r_i$ is an orthogonal decomposition we find by taking residues at $t = a_j$ that

$$a_j(c-1)\frac{\langle r_j, r_j\rangle}{\langle r, r\rangle} = (b_j - a_j)\prod_{i\neq j}\frac{b_i - a_j}{a_i - a_j}$$

and since $c = b_1 \cdots b_n a_1^{-1} \cdots a_n^{-1}$ we get after division by $a_j c^{1/2}$ the result

$$\frac{\langle r_j, r_j\rangle}{\langle r, r\rangle} = \frac{(b_j^{1/2}a_j^{-1/2} - b_j^{-1/2}a_j^{1/2})}{c^{1/2} - c^{-1/2}}\prod_{i\neq j}\frac{b_i^{1/2}a_j^{-1/2} - b_i^{-1/2}a_j^{1/2}}{a_i^{1/2}a_j^{-1/2} - a_i^{-1/2}a_j a_j^{1/2}}$$

which becomes

$$\frac{\langle r_j, r_j\rangle}{\langle r, r\rangle} = \frac{\sin\pi(\beta_j - \alpha_j)}{\sin\pi\sum_i(\beta_i - \alpha_i)}\prod_{i\neq j}\frac{\sin\pi(\beta_i - \alpha_j)}{\sin\pi(\alpha_i - \alpha_j)}$$

with $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_n \leq 1$ and $0 < \beta_1 < \beta_2 < \cdots < \beta_n \leq 1$ by assumption. Our assertion follows simply by determination of the sign of the latter products. $\square$

The next theorem gives an arithmetic criterion for the hypergeometric group $H(a, b)$ being finite.

Theorem 3.15. *Suppose the parameters $a_1, \cdots, a_n, b_1, \cdots, b_n$ are roots of unity, and say*

$$\mathbb{Z}[a_1, \cdots, a_n, b_1, \cdots, b_n] = \mathbb{Z}[\exp(2\pi i/h)]$$

for some $h \in \mathbb{N}$. Then the group $H(a, b)$ is finite if and only if for each $k \in (\mathbb{Z}/h\mathbb{Z})^\times$ the two sets

$$a^k = \{a_1^k, \cdots, a_n^k\} \ , \ b^k = \{b_1^k, \cdots, b_n^k\}$$

interlace on the unit circle.

Proof. For $\zeta_h = \exp(2\pi i/h)$ the ring of integers $\mathbb{Z}[\zeta_h]$ of the cyclotomic field $\mathbb{Q}(\zeta_h)$ is a free $\mathbb{Z}$-module of rank $m = \varphi(n)$ with basis ζ_h^k for $k \in (\mathbb{Z}/h\mathbb{Z})^\times$. For $k \in (\mathbb{Z}/h\mathbb{Z})^\times$ and $\zeta^h = 1$ the Galois automorphism

$$\sigma_k(\zeta) = \zeta^k$$

identifies the Galois group with $(\mathbb{Z}/h\mathbb{Z})^\times$.

By Corollary 3.10 we have $H(a, b) < \mathrm{GL}(n, \mathbb{Z}[\zeta_h])$. Multiplication by an algebraic integer $\zeta \in \mathbb{Z}[\zeta_h]$ in the basis ζ_h^k has a square matrix of size m with rational integral coefficients, whose eigenvalues are the Galois conjugates $\sigma_k(\zeta)$. In this way the diagonal embedding

$$\prod_{k\in(\mathbb{Z}/h\mathbb{Z})^\times} \sigma_k : H(a, b) \to \prod_{k\in(\mathbb{Z}/h\mathbb{Z})^\times} H(a^k, b^k)$$

realizes $H(a, b)$ as a subgroup of $\mathrm{GL}(mn, \mathbb{Z})$. Since a subgroup of $\mathrm{GL}(N, \mathbb{Z})$ is finite if and only if it leaves invariant a positive definite Hermitian form the theorem follows. □

The theorem gives a very quick way of deciding if a given hypergeometric group $H(a, b)$ is finite, or equivalently if a given hypergeometric function ${}_nF_{n-1}(\alpha; \beta|z)$ is an algebraic function of z. What is not clear at all is that in fact one can derive from it (using the classification of finite complex reflection group by Shephard and Todd [35]) a complete classification of finite hypergeometric groups [4].

Remark 3.16. *By the same method one can prove that the hypergeometric group $H(a, b) < \mathrm{GL}(n, \mathbb{Z}[\zeta_h])$ is a discrete subgroup of $\mathrm{GL}(n, \mathbb{C})$ if the sets a^k and b^k interlace on the unit circle for all $k \in (\mathbb{Z}/h\mathbb{Z})^\times$ with $1 < k < h/2$. For a discussion of the geometric representation of algebraic integers in algebraic number theory we refer to [38].*

Example 3.17. *The image under* $s \mapsto \exp(2\pi i s)$ *of the two sets*

$$\{\alpha_j\} = \{1/30, 7/30, 11/30, 13/30, 17/30, 19/30, 23/30, 29/30\}$$

$$\{\beta_j\} = \{6/30, 10/30, 12/30, 15/30, 18/30, 20/30, 24/30, 30/30\}$$
$$= \{1/5, 1/3, 2/5, 1/2, 3/5, 2/3, 4/5, 1\}$$

interlace on the unit circle. Note that the characteristic polynomials

$$\prod_{j=1}^{8}(t - a_j) = \Phi_{30}(t) \ , \ \prod_{j=1}^{8}(t - b_j) = \Phi_1(t)\Phi_2(t)\Phi_3(t)\Phi_5(t)$$

are defined over $\mathbb{Z}$. *Here* $\Phi_m(t)$ *is the* m^{th} *cyclotomic polynomial of degree* $\varphi(m)$. *Hence both sets* $\{a_j\}$ *and* $\{b_j\}$ *are stable under raising to the power* $k \in (\mathbb{Z}/30\mathbb{Z})^\times$. *Therefore the group* $H(a, b)$ *is finite (with order* 696.729.600*) and so the Clausen–Thomae hypergeometric function with these parameters is an algebraic function.*

Let me now explain the origin of this fancy example, which after all is not too strange to write down for someone who is familiar with the theory of finite Coxeter groups

3.4 Intermezzo on Coxeter groups

Suppose $M = (m_{ij})_{1 \le i,j \le n}$ is a Coxeter matrix which means that $m_{ii} = 1$ for all i and $m_{ij} = m_{ji} \in \mathbb{N}_{\ge 2}$ for all $i \ne j$.

Definition 3.18. *The Coxeter group* $W = W(M)$ *associated to the Coxeter matrix* M *is given by the presentation*

$$W = \langle s_i; i = 1, \cdots, n \rangle / \{(s_i s_j)^{m_{ij}} = 1\}$$

so in particular $s_i^2 = 1$, *hence* s_i *is an involution.*

Definition 3.19. *The Coxeter diagram associated to the Coxeter matrix* M *is a marked graph, with nodes indexed by* $i = 1, \cdots, n$. *The* i^{th} *and the* j^{th} *node are connected if* $m_{ij} \ge 3$, *and the edge is marked* m_{ij} *if* $m_{ij} \ge 4$. *So an unmarked edge between the* i^{th} *and* j^{th} *node means* $m_{ij} = 3$, *while no edge between the* i^{th} *and* j^{th} *node means* $m_{ij} = 2$.

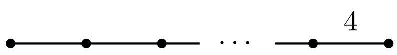
A Coxeter diagram is called crystallographic if $m_{ij} \in \{1, 2, 3, 4, 6\}$ for all i, j. Finite Coxeter groups corresponding to crystallographic Coxeter diagrams are also called Weyl groups. Both the symmetric group S_{n+1} and the hyperoctahedral group $C_2^n \rtimes S_n$ are Weyl groups.

Example 3.20. *The symmetric group* $W = S_{n+1}$ *has a Coxeter presentation with generators* $s_1 = (12), \cdots, s_n = (n\ n+1)$ *the nearest neighbour transpositions. The Coxeter matrix is given by*

$$m_{ij} = \begin{cases} 1 & \text{if } i = j \\ 2 & \text{if } |i-j| \geq 2 \\ 3 & \text{if } |i-j| = 1 \end{cases}$$

So the Coxeter diagram with nodes numbered from left to right is of the form

Example 3.21. *The hyperoctahedral group* $W = C_2^n \rtimes S_n$ *has a Coxeter presentation with generators* $s_1 = (12), \cdots, s_{n-1} = (n-1\ n) \in S_n$ *and* $s_n = (1, \cdots, 1, -1) \in C_2^n$*. The Coxeter diagram with nodes numbered from left to right is of the form*



Consider a Euclidean vector space V with basis $e_1, \cdots, e_n$ and with inner product given by the Gram matrix

$$\langle e_i, e_j \rangle = -2\cos(\pi/m_{ij})$$

for all i, j. Define the orthogonal reflection

$$s_i : V \to V\ ,\ s_i(v) = v - \langle v, e_i \rangle e_i$$

with mirror the orthogonal complement of e_i. It is easy to check that this assignment extends to a homomorphism $W \to \mathrm{O}(V)$. This is called the reflection representation of the Coxeter group W. A fundamental result in the theory of Coxeter groups is the theorem of Tits.

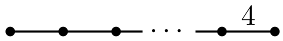
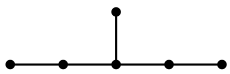
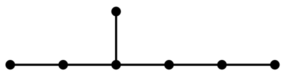
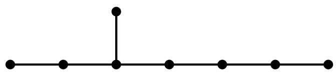
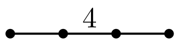
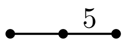
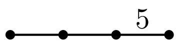
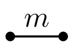
Theorem 3.22. *The reflection representation of a Coxeter group is faithful.*

The inner product on V is positive definite if and only if the Coxeter group W is finite. From now on we assume that the Coxeter group W is a finite group.

It turns out that for finite Coxeter groups the Coxeter diagram has no loops. The Coxeter element is the product of the involutive generators taken in some order. One can show that all Coxeter element are conjugated in W. Suppose in addition that the Coxeter diagram is connected. The order of a Coxeter element is called the Coxeter number, usually denoted h. The eigenvalues of a Coxeter element in the reflection representation are $\exp(2\pi i m_j/h)$ with

$$1 = m_1 \leq m_2 \leq \cdots \leq m_n = (h-1)$$

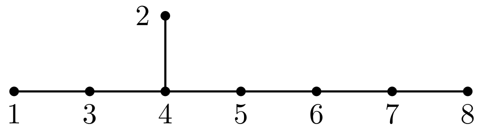
the sequence of exponents.

Name	Coxeter diagram	h	Exponents
A_n		$n+1$	$1, 2, \cdots, n$
B_n		$2n$	$1, 3, 5, \cdots, 2n-1$
D_n		$2n-2$	$1, 3, \cdots, 2n-3, n-1$
E_6		12	$1, 4, 5, 7, 8, 11$
E_7		18	$1, 5, 7, 9, 11, 13, 17$
E_8		30	$1, 7, 11, 13, 17, 19, 23, 29$
F_4		12	$1, 5, 7, 11$
H_3		10	$1, 5, 9$
H_4		30	$1, 11, 19, 29$
$\mathrm{I}_2(m)$	 $m \geq 5$	m	$1, m-1$

Above is a list of the finite Coxeter groups, which are irreducible in the sense that the Coxeter diagram is connected, or equivalently for which the reflection representation is irreducible. This classification can be found in various text books [6], [21]. In the first column we have the Cartan symbol, with the subindex n for the number of nodes of the Coxeter diagram. In the second column we have the Coxeter diagram, in the third column the Coxeter number, and in the last column the exponents. This classification of finite Coxeter groups is one of the most fundamental classifications in mathematics. For example it plays a crucial role in the classification of the simple algebraic groups.

With this basic knowledge of Coxeter groups in mind it is easy to see that the finite hypergeometric group of Example 3.17 is contained (and in fact equal to) the Coxeter group $W(\mathrm{E}_8)$.

Example 3.23. *Consider the Coxeter diagram of type* E_8 *with the nodes traditionally numbered by*



In Example 3.17 the monodromy group $H(a,b)$ *is the subgroup of the Coxeter group* $W(\mathrm{E}_8)$ *generated by*

$$h_\infty = s_2s_1s_3s_5s_6s_7s_8s_4 \ , \ h_1 = s_4 \ , \ h_0 = s_8s_7s_6s_5s_3s_1s_2$$

for which the topological relation $h_\infty h_1 h_0 = 1$ *indeed is true. So* h_∞ *is a Coxeter element of type* E_8 *and* h_0 *is a Coxeter element of type* $\mathrm{A}_1+\mathrm{A}_2+\mathrm{A}_4$. *The element* h_1 *is indeed a reflection. It can be shown that* $H(a,b) = W(\mathrm{E}_8)$, *which is a group of order almost* 700 *million. Note that* $W(\mathrm{E}_8) \cong \mathrm{O}_8^+(2)$ *and* $\mathrm{O}_8^+(2) = 2.G.2$ *with* $G \cong \mathrm{PSO}_8^+(2)$ *a simple group of order* 174.182.400.

Example 3.24. *Recall Clausen's formula*

$${}_2F_1(\alpha,\beta,\alpha+\beta+1/2;z)^2 = {}_3F_2(2\alpha,2\beta,\alpha+\beta;2\alpha+2\beta,\alpha+\beta+1/2|z)$$

as discussed in Example 3.11, and look at the particular example

$${}_2F_1(1/4,-1/12,2/3;z)^2 = {}_3F_2(1/2,-1/6,1/6;1/3,2/3|z)$$

with $\alpha = 1/4, \beta = -1/12$ *and* $\gamma = \alpha+\beta+1/2 = 2/3$. *The function on the left hand side is algebraic with projective monodromy group the tetrahedral group* A_4. *Indeed the exponent differences are* $\frac{1}{2},\frac{1}{3},\frac{1}{3}$ *at the points* $1,0,\infty$ *respectively. So the right hand side is again an algebraic function. The monodromy group of the latter is* $C_2 \times A_4 \cong C_2^3 \rtimes A_3$ *(of index* 2 *in* $W(\mathrm{B}_3) = C_2^3 \rtimes S_3$*) in its three dimensional reflection representation. Indeed the eigenvalues match for*

$$h_\infty = -(234) \ , \ h_1 = -(12)(34) \ , \ h_0 = (123)$$

and $h_\infty h_1 h_0 = 1$ *as should.*

3.5 Lorentzian Hypergeometric Groups

Let us suppose that the parameters α, β of the hypergeometric equation satisfy the regularity condition

$$0 < \alpha_1 < \cdots < \alpha_n \leq 1 \ , \ 0 < \beta_1 < \cdots < \beta_n \leq 1$$

and the usual irreducibility condition $\alpha_i \neq \beta_j$ for all $i, j = 1, \cdots, n$. Let us denote $a_j = \exp(2\pi i \alpha_j), b_j = \exp(2\pi i \beta_j)$ and let $H(a, b) < \mathrm{GL}_n(\mathbb{C})$ be the hypergeometric group of Levelt's theorem with generators A, B satisfying

$$\det(t - A) = \prod(t - a_j) \ , \ \det(t - b) = \prod(t - b_j)$$

and $\mathrm{rk}(A - B) = 1$. Let $\langle \cdot, \cdot \rangle$ be the up to a real scalar unique Hermitian form on $\mathbb{C}^n$ invariant under $H(a, b)$ given by Theorem 3.14. If in addition we assume $c = b_1 \cdots b_n a_1^{-1} \cdots a_n^{-1} \neq 1$ then $C = A^{-1}B$ is a unitary complex reflection of the form

$$C(v) = v + (c - 1)\frac{\langle v, r \rangle}{\langle r, r \rangle} r$$

for some nonzero $r \in \mathbb{C}^n$. If $r = \sum r_j$ with $Ar_j = a_j r_j$ then we have shown in the proof of Theorem 3.14 that

$$\frac{\langle r_j, r_j \rangle}{\langle r, r \rangle} = \frac{\sin \pi(\beta_j - \alpha_j)}{\sin \pi \sum_i (\beta_i - \alpha_i)} \prod_{i \neq j} \frac{\sin \pi(\beta_i - \alpha_j)}{\sin \pi(\alpha_i - \alpha_j)}$$

which in turn implies that the signature (p, q) of $\langle \cdot, \cdot \rangle$ is given by

$$|p - q| = |\sum_{j=1}^{n} (-1)^{j + m_j}|$$

with $m_j = \#\{i; \alpha_i < \beta_j\}$ for $j = 1, \cdots, n$. In particular the Hermitian form is definite if and only if the sets a and b interlace on the unit circle.

Definition 3.25. *Let $n \geq 3$. Let us assume that $\alpha_1 < \beta_1$ which we always may by possibly interchanging α and β. The two sets $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$ almost interlace on the unit circle, if there exists one index k with*

$$\cdots < \alpha_{k-1} < \beta_{k-1} < \beta_k < \alpha_k < \alpha_{k+1} < \beta_{k+1} < \cdots$$

while at all other places the sequences a and b interlace on the unit circle, or if there exists one index k with

$$\cdots < \beta_{k-1} < \alpha_k < \alpha_{k+1} < \beta_k < \beta_{k+1} < \alpha_{k+2} < \cdots$$

while at all other places the sequences a and b interlace on the unit circle, or if there exist two indices $k < l$ with

$$\cdots < \beta_{k-1} < \alpha_k < \alpha_{k+1} < \alpha_{k+2} < \beta_k < \cdots$$
$$\cdots < \alpha_{l+1} < \beta_{l-1} < \beta_l < \beta_{l+1} < \alpha_{l+2} < \cdots$$

while at all other places the sequences a and b interlace on the unit circle, or if there exist two indices $k < l$ with

$$\cdots < \alpha_k < \beta_k < \beta_{k+1} < \beta_{k+2} < \alpha_{k+1} < \cdots$$
$$\cdots < \beta_{l+1} < \alpha_l < \alpha_{l+1} < \alpha_{l+2} < \beta_{l+2} < \cdots$$

while at all other places the sequences a and b interlace on the unit circle.

In the first two cases we have $m_j = j$ for $j \neq k$ while $m_k = k \mp 1$, and hence $|p-q| = n-2$. In the third case we have $m_j = j$ for $j < k$, $m_j = j+2$ for $k \leq j < l$, $m_l = l+1$ and $m_j = j$ for $j \geq l$ and therefore $|p-q| = n-2$. The same conclusion holds in the fourth case. Hence if a and b almost interlace on the unit circle the Hermitian form has Lorentzian signature. Conversely, it can be shown that Lorentzian signature only happens if a and b almost interlace on the unit circle.

Theorem 3.26. *Let $n \geq 3$. Suppose the two sets a and b almost interlace on the unit circle, and suppose that $\overline{a} = a$ and $\overline{b} = b$, which together mean that the hypergeometric group $H(a,b)$ is a subgroup of the real Lorentz group. Then the dimension $n = 2m+1$ is odd, and after possible interchange of the two parameter sets a and b we have*

$$0 < \beta_1 < \alpha_1 < \cdots < \alpha_m < \alpha_{m+1} = \tfrac{1}{2} < \alpha_{m+2} < \cdots < \beta_{n-1} < \beta_n = 1$$

and so there is interlacing on the unit circle except around $a_{m+1} = -1$ with a_m and a_{m+2} as nearest neighbors from $a \sqcup b$, and around $b_n = 1$ with b_1 and b_{n-1} as nearest neighbors from $a \sqcup b$. Moreover we have $c = \det C = -1$ and hence

$$C(v) = v - 2\frac{\langle v, r\rangle}{\langle r, r\rangle} r\ .$$

If the signature of the Hermitian form is taken $(n-1,1)$ then $\langle r, r\rangle < 0$. Hence for the natural action of $H(a,b)$ on hyperbolic space

$$H^{2m} = \{[v] \in \mathbb{P}^{2m}(\mathbb{R}); \langle v, v\rangle < 0\}$$

of dimension $2m$ (see Exercise 2.10 for the definition of Lobachevsky space) the element C acts as an involution with fixed point $[r]$ in H^{2m} (rather than a reflection with mirror $r^{\perp}$ in case $\langle r, r\rangle$ would have been a positive number).

Proof. Under the reality restriction $\overline{a} = a$ and $\overline{b} = b$ only the last two of the four cases of Definition 3.25 can occur, and in that case the clustering of such triples from a and from b should be symmetric under complex conjugation, and therefore the clustering takes place around ± 1. We may assume after a possible interchange of a and b that $b_n = 1$ (in accordance with the standard convention $\beta_n = 1$) and so $a_{m+1} = -1$, and the three nearest points from $a \sqcup b$ to 1 are $\{b_1, b_{n-1}, b_n = 1\}$ and to -1 are $\{a_m, a_{m+1} = -1, a_{m+2}\}$. Such a conjugation symmetric configuration of points on the unit circle can indeed only occur if $n = 2m + 1$ is odd, and clearly $c = -1$. Hence C is given by the desired formula for some $r \in \mathbb{R}^n$ with $\langle r, r\rangle \neq 0$.

It remains to show that under the assumption that $\langle \cdot, \cdot \rangle$ has signature $(n-1, 1)$ we have $\langle r, r\rangle < 0$. Using the formula

$$\frac{\langle r_j, r_j\rangle}{\langle r, r\rangle} = \frac{\sin \pi(\beta_j - \alpha_j)}{\sin \pi \sum_i (\beta_i - \alpha_i)} \prod_{i \neq j} \frac{\sin \pi(\beta_i - \alpha_j)}{\sin \pi(\alpha_i - \alpha_j)}$$

it follows by straightforward inspection that $\langle r_j, r_j\rangle / \langle r, r\rangle < 0$ for all j except for $j = m$. Hence $\langle r, r\rangle < 0$ if the signature is taken $(n-1, 1)$. The last sentence is an immediate consequence of this. $\square$

Remark 3.27. *The above result holds and was proved under the regularity assumption that $\alpha_i \neq \alpha_j, \beta_i \neq \beta_j$ for all $i \neq j$. However the hypergeometric group $H(a, b)$ is well defined and acts irreducibly just under the condition $a_i \neq b_j$ for all i, j. Hence by a deformation argument the theorem equally holds if either $a_m = a_{m+1} = a_{m+2} = -1$, either $b_1 = b_{n-1} = b_n = 1$, or both of these conditions hold.*

The above theorem and remark were obtained by Fuchs, Meiri and Sarnak [16]. Using this result and the classification of finite hypergeometric groups they were able to give a classification of the integral Lorentzian hypergeometric groups, for which the parameters a, b of the hypergeometric group $H(a, b)$ are almost interlacing on the unit circle and $\prod(t - a_i)$ and $\prod(t - b_i)$ have integral coefficients, or equivalently relative to a suitable basis $H(a, b)$ is contained in $\mathrm{O}(L)$ with L an integral Lorentzian lattice (a free Abelian group of finite rank with an integral symmetric bilinear $\langle \cdot, \cdot \rangle$ form or Lorentz signature). If $H(a, b)$ has finite index in $\mathrm{O}(L)$ then $H(a, b)$ is called an arithmetic group, otherwise $H(a, b)$ is called a thin group. The question whether an integral Lorentzian hypergeometric group is arithmetic or thin was the central theme of the work by Fuchs, Meiri and Sarnak with the following two results.

Theorem 3.28. *Any integral Lorentzian hypergeometric group $H(a,b)$ in rank $n=3$ is arithmetic.*

Conjecture 3.29. *Any integral Lorentzian hypergeometric group $H(a,b)$ in sufficiently high rank n is thin.*

Fuchs, Meiri and Sarnak developed a technique which enabled them to verify that for about half the cases in their classification table of Lorentzian hypergeometric groups $H(a,b)$ the conjecture holds. In this way they found no example of an arithmetic integral Lorentzian hypergeometric group in rank $n \geq 5$, and they mention that the conjecture might even hold for $n \geq 5$. In the rest of this section we shall discuss some of their work.

We say that two hypergeometric groups $H(a,b)$ and $H(c,d)$ differ by a scalar shift if $c = za, d = zb$ for some $z \in \mathbb{C}^\times$. In case all these parameters lie on the unit circle this means that the corresponding real parameters $\alpha, \beta, \gamma, \delta$ satisfy $\gamma \equiv \alpha + \zeta\underline{1}, \delta \equiv \beta + \zeta\underline{1}$ modulo $\mathbb{Z}^n$ for some $\zeta \in \mathbb{R}$. Let us first discuss the case of a real Lorentzian hypergeometric group in dimension $n=3$ and prove the above theorem.

Proof. Consider the ${}_2F_1$ hypergeometric function with real parameters and Riemann scheme given by

$z=0$	$z=1$	$z=\infty$
$0, 1-\gamma$	$0, \gamma-\alpha-\beta$	α, β

with monodromy group $H((a,b),(1,c))$ in the usual notation $a = \exp(2\pi i\alpha)$, $b = \exp(2\pi i\beta)$, $c = \exp(2\pi i\gamma)$. After multiplication by $z^{(\gamma-1)/2}$ and taking $\gamma = \alpha+\beta+1/2$, $\delta = \alpha-\beta+1/2$ we get the modified scalar shifted Riemann scheme

$z=0$	$z=1$	$z=\infty$
$(\gamma-1)/2, (1-\gamma)/2$	$0, 1/2$	$\delta/2, 1/2-\delta/2$

If $|\gamma-1|+|\delta-1/2| < 1/2$ and we put $c^{1/2} = \exp(\pi i(\gamma-1))$, $d^{1/2} = \exp(\pi i\delta)$ then the two subsets $\underline{c}^{1/2} = \{c^{1/2}, c^{-1/2}\}$ and $\underline{d}^{1/2} = \{d^{1/2}, -d^{-1/2}\}$ of the unit circle do not interlace. Hence the hypergeometric group $H(\underline{d}^{1/2}, \underline{c}^{1/2})$ leaves invariant an indefinite Hermitian form $\langle\cdot,\cdot\rangle$ on $\mathbb{C}^2$. The conclusion is that $H(\underline{d}^{1/2}, \underline{c}^{1/2})$ is subgroup of

$$\mathrm{SU}^{\pm}_{1,1}(\mathbb{C}) = \{g \in \mathrm{GL}_2(\mathbb{C}); \langle gu, gv\rangle = \langle u,v\rangle \; \forall\, u,v \in \mathbb{C}^2, \det g = \pm 1\}$$

which is called the complex Lorentz group in dimension 2.

The isotropic vectors for $\langle\cdot,\cdot\rangle$ on the associated projective line $\mathbb{P}$ form an invariant circle. A connected component of its complement is a hyperbolic disc $\mathbb{D}$, and $H(\underline{d}^{1/2}, \underline{c}^{1/2})$ acts projectively on $\mathbb{D}$ as index two subgroup of the group generated by the antiholomorphic reflections in the sides of a hyperbolic triangle with angles $\pi|\gamma - 1|, \pi|\delta - 1/2|, \pi/2$ in accordance with the Schwarz–Klein theory.

The second symmetric square of the above modified Riemann scheme is equal to

$z = 0$	$z = 1$	∞
$0, \gamma - 1, 1 - \gamma$	$0, 1, 1/2$	$\delta, 1/2, 1 - \delta$

which is the Riemann scheme of a ${}_3F_2$ hypergeometric function. Its hypergeometric group $H(\underline{d}, \underline{c})$ has parameters $\underline{c} = (c, 1, 1/c)$ and $\underline{d} = (d, -1, 1/d)$. Since $|\gamma - 1| + |\delta - 1/2| < 1/2$ the parameters $\underline{c}$ and $\underline{d}$ almost interlace on the unit circle. Hence $H(\underline{d}, \underline{c})$ is contained in the real Lorentz group $\mathrm{O}(\mathbb{R}^{2,1})$ in dimension 3. Conversely, any real Lorentzian hypergeometric group in dimension $n = 3$ is of this form.

The second symmetric square induces the so called spin homomorphism

$$\mathrm{SU}^{\pm}_{1,1}(\mathbb{C}) \longrightarrow \mathrm{O}(\mathbb{R}^{2,1})$$

with both kernel and cokernel of order 2. Under this spin homomorphism the hypergeometric group $H(\underline{d}^{1/2}, \underline{c}^{-1/2})$ maps onto the hypergeometric group $H(\underline{d}, \underline{c})$. Therefore $H(\underline{d}^{1/2}, \underline{c}^{-1/2})$ is a discrete cocompact (or cofinite volume) subgroup of $\mathrm{SU}^{\pm}_{1,1}(\mathbb{C})$ if and only if the hypergeometric group $H(\underline{d}, \underline{c})$ is a discrete cocompact (or cofinite volume) subgroup of $\mathrm{O}(\mathbb{R}^{2,1})$.

Let us now assume that $\gamma - 1 = 1/k, \delta - 1/2 = 1/l$ with $k, l \in \mathbb{N} \sqcup \{\infty\}$ and $1/k + 1/l < 1/2$. Then $H(\underline{d}^{1/2}, \underline{c}^{-1/2})$ is a discrete cofinite volume subgroup of $\mathrm{SU}^{\pm}_{1,1}(\mathbb{C})$ by the Schwarz–Klein theory, and hence $H(\underline{d}, \underline{c})$ is a discrete cofinite volume subgroup of $\mathrm{O}(\mathbb{R}^{2,1})$. More precisely, cofinite volume can be sharpened to cocompact if $k, l \neq \infty$. If in addition $H(\underline{d}, \underline{c})$ is a subgroup of an integral Lorentz group $\mathrm{O}(L)$ with L an integral Lorentzian lattice rank 3, then the index of $H(\underline{d}, \underline{c})$ in $\mathrm{O}(L)$ is given by

$$[\mathrm{O}(L) : H(\underline{d}, \underline{c})] = \mathrm{vol}\{\mathrm{O}(\mathbb{R}^{2,1})/H(\underline{d}, \underline{c})\} : \mathrm{vol}\{\mathrm{O}(\mathbb{R}^{2,1})/\mathrm{O}(L)\}$$

and so is finite. This proves Theorem 3.28. □

Let us keep the notation of the above proof. In these cases the group $H(\underline{d}, \underline{c})$ is defined over the ring $\mathbb{Z}[c+1/c, d+1/d]$ and so $H(\underline{d}, \underline{c})$ is an integral Lorentzian group if both $c + 1/c = 2\cos(2\pi/k)$ and $d + 1/d = 2\cos(2\pi/l)$ are integers, or equivalently if $k, l \in \{3, 4, 6, \infty\}$. The possible pairs (k, l) are easily classified with outcome

k	3	4	6	∞
l	∞	$6,\infty$	$4,6,\infty$	$3,4,6,\infty$

in accordance with the tabulation

$$\alpha = (1/2 - 1/l, 1/2, 1/2 + 1/l)\ ,\ \beta = (1/k, 1 - 1/k, 1)$$

in [16]. In these cases the hypergeometric group $H(\underline{d}, \underline{c})$ is a subgroup of $\mathrm{O}(L)$ with L an integral Lorentzian lattice. In case $k, l \neq \infty$ this lattice has no nonzero isotropic vectors, because cocompactness of $H(\underline{d}, \underline{c})$ in $\mathrm{SU}^{\pm}_{1,1}(\mathbb{C})$ is equivalent to absence of cusps in the quotient space $\mathbb{D}/H(\underline{d}^{1/2}, \underline{c}^{1/2})$. This is a phenomenon for small rank because of Meyer's theorem [34].

Theorem 3.30. *Any indefinite integral lattice of rank $n \geq 5$ has nonzero isotropic vectors.*

Remark 3.31. *For $\gamma - 1 = 1/k, \delta - 1/2 = 1/l$ with $k, l \in \mathbb{N} \sqcup \{\infty\}$ and $1/k + 1/l < 1/2$ the real Lorentzian group $H(\underline{d}, \underline{c})$ is commensurable with the Schwarz triangle group with parameters $(2, k, l)$. This group is definied over the ring of real algebraic integers $\mathbb{O} = \mathbb{Z}[2\cos(2\pi/k), 2\cos(2\pi/l)]$, and therefore the hypergeometric group $H(\underline{d}, \underline{c})$ is an arithmetic group over $\mathbb{O}$ if $\underline{d}^{\sigma}$ and $\underline{c}^{\sigma}$ interlace on the unit circle for all $\sigma \neq 1$ in the Galois group of $\mathbb{O}$ over $\mathbb{Z}$. These arithmetic triples $(2, k, l)$ with $k \leq l$ have been classified by Takeuchi, who found 37 cocompact and 4 cofinite volume cases [39].*

The classification of integral Lorentzian hypergeometric groups in rank $n \geq 5$ reduces to the classification of finite hypergeometric groups in rank n, $n-1$ or $n-2$. The proof is trivial using Theorem 3.26.

Theorem 3.32. *Let $H(a, b)$ be an integral Lorentzian hypergeometric group in rank $n = 2m + 1$ with $-1 \in a$ and $1 \in b$. If both $-1 \in a$ and $1 \in b$ have multiplicity one then the sets*

$$a' = a \sqcup \{1\} - \{-1\}\ ,\ b' = b \sqcup \{-1\} - \{1\}$$

of cardinality n interlace on the unit circle. If $-1 \in a$ has multiplicity three and $1 \in b$ has multiplicity one then the sets

$$a' = a \sqcup \{1\} - \{-1, -1\}\ ,\ b' = b - \{1\}$$

(so $-1 \in a'$ gets multiplicity one) of cardinality $n - 1 = 2m$ interlace on the unit circle. If both $-1 \in a$ and $1 \in b$ have multiplicity three then the sets

$$a' = a - \{-1, -1\}\ ,\ b' = b - \{1, 1\}$$

(so both $-1 \in a'$ and $1 \in b'$ get multiplicity one) of cardinality $n-2$ interlace on the unit circle.

If both $-1 \in a$ and $1 \in b$ have multiplicity three then either $n = 3$ and we recover the case $k = l = \infty$ in our classification for $n = 3$, or $n = 5$ with $\pm i \in a$ and $\omega, \omega^2 \in b$, or $n = 5$ with $-\omega, -\omega^2 \in a$ and $\pm i \in b$.

Example 3.33. *Recall from Exercise 3.12 that the hypergeometric function ${}_8F_7(\alpha'; \beta'|z)$ with parameters*

$$\alpha' = (1/30, 7/30, 11/30, 13/30, 17/30, 19/30, 23/30, 29/30)$$

and β' given by

β'
$(1/12, 1/4, 5/12, 1/2, 7/12, 3/4, 11/12, 1)$
$(1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8, 1)$
$(1/7, 2/7, 3/7, 1/2, 4/7, 5/7, 6/7, 1)$
$(1/5, 1/3, 2/5, 1/2, 3/5, 2/3, 4/5, 1)$
$(1/5, 1/4, 2/5, 1/2, 3/5, 3/4, 4/5, 1)$
$(1/8, 1/3, 3/8, 1/2, 5/8, 2/3, 7/8, 1)$
$(1/12, 1/3, 5/12, 1/2, 7/12, 2/3, 11/12, 1)$
$(1/18, 5/18, 7/18, 1/2, 11/18, 13/18, 17/18, 1)$

is algebraic with monodromy group $W(\mathrm{E}_8)$.

Then it is clear that the hypergeometric group $H(a, b)$ with parameters

$$\alpha = (1/30, 7/30, 11/30, 13/30, 17/30, 19/30, 23/30, 29/30, 1)$$

(add 1 to α') and β given by

β
$(1/12, 1/4, 5/12, 1/2, 1/2, 1/2, 7/12, 3/4, 11/12)$
$(1/8, 1/4, 3/8, 1/2, 1/2, 1/2, 5/8, 3/4, 7/8)$
$(1/7, 2/7, 3/7, 1/2, 1/2, 1/2, 4/7, 5/7, 6/7)$
$(1/5, 1/3, 2/5, 1/2, 1/2, 1/2, 3/5, 2/3, 4/5)$
$(1/5, 1/4, 2/5, 1/2, 1/2, 1/2, 3/5, 3/4, 4/5)$
$(1/8, 1/3, 3/8, 1/2, 1/2, 1/2, 5/8, 2/3, 7/8)$
$(1/12, 1/3, 5/12, 1/2, 1/2, 1/2, 7/12, 2/3, 11/12)$
$(1/18, 5/18, 7/18, 1/2, 1/2, 1/2, 11/18, 13/18, 17/18)$

(β is obtained from β' by deleting 1 and giving 1/2 multiplicity three) is an integral Lorentzian group. Indeed a, b are almost interlacing with no a_i on the unit circle between $b_4 = b_5 = b_6 = -1$ and likewise no b_i between $a_8 = \zeta_{30}^{-1}, a_9 = 1, a_1 = \zeta_{30}$.

Example 3.34. *Suppose that* $1 \le j \le n$ *with* $\gcd(j, n+1) = 1$. *Recall from Exercise 3.9 that the hypergeometric function* ${}_nF_{n-1}(\alpha', \beta'|z)$ *with parameters*

$$\alpha' = (\tfrac{1}{n+1}, \cdots, \tfrac{n}{n+1}) \ , \ \beta' = (\tfrac{1}{j}, \cdots, \tfrac{j-1}{j}, \tfrac{1}{n+1-j}, \cdots, \tfrac{n-j}{n+1-j}, 1)$$

is algebraic with monodromy group the symmetric group S_{n+1}.

If $n = 2m+1$ *then the parameters*

$$\alpha = (\tfrac{1}{n+1}, \cdots, \tfrac{m}{n+1}, \tfrac{m+2}{n+1}, \cdots \tfrac{n}{n+1}, 1) \ , \ \beta = (\tfrac{1}{j}, \cdots, \tfrac{j-1}{j}, \tfrac{1}{n+1-j}, \cdots, \tfrac{n-j}{n+1-j}, \tfrac{1}{2})$$

are almost interlacing. If $n = 2m$ *then the parameters*

$$\alpha = (\tfrac{1}{n+1}, \cdots, \tfrac{n}{n+1}, 1) \ , \ \beta = (\tfrac{1}{j}, \cdots, \tfrac{j-1}{j}, \tfrac{1}{n+1-j}, \cdots, \tfrac{n-j}{n+1-j}, \tfrac{1}{2}, \tfrac{1}{2})$$

are almost interlacing. Note that in β *the parameter* $1/2$ *occurs with multiplicity three.*

We now outline the method of Fuchs, Meiri and Sarnak for proving that an integral Lorentzian hypergeometric group $H(a, b)$ for rank $n \ge 5$ is thin. Their method uses computer computations, which we have not bothered to check.

First recall the usual notation. The hypergeometric group $H(a, b)$ has generators $A = h_\infty, B = h_0^{-1}, C = h_1$ with $AC = B$. Moreover we have $\det(t - A) = \prod(t - a_i)$, $\det(t - B) = \prod(t - b_i)$ and $\det C = c = b_1 \cdots b_n / a_1 \cdots a_n = -1$ in the integral Lorentzian case. The transformation C is an involution of the form

$$v \mapsto C(v) = v - 2\frac{\langle v, r\rangle}{\langle r, r\rangle} r$$

for some vector $r \in \mathbb{C}^n$ of negative norm $\langle r, r\rangle < 0$.

Let us assume that A has finite order, which in turn implies that all its eigenvalues a_i are distinct. In the proof of Theorem 3.14 we have derived the signature formula

$$\frac{\langle r_j, r_j\rangle}{\langle r, r\rangle} = \frac{\sin\pi(\beta_j - \alpha_j)}{\sin\pi\sum_i(\beta_i - \alpha_i)} \prod_{i \ne j} \frac{\sin\pi(\beta_i - \alpha_j)}{\sin\pi(\alpha_i - \alpha_j)}$$

with $r = \sum r_i, Ar_i = a_i r_i$ the eigenvalue decomposition of r for A by taking the residue of the identity

$$\prod_{i=1}^{n} \frac{t - b_i}{t - a_i} = 1 + (1 - c)\sum_{i=1}^{n} \frac{a_i}{t - a_i} \frac{\langle r_i, r\rangle}{\langle r, r\rangle}$$

at $t = a_j$. Note that the Laurent expansion at ∞ of the function on the left side has integral coefficients.

If instead we first multiply this identity by t^{k-1} and then take the residue at $t = a_j$ we get

$$\mathrm{Res}_{t=a_j}\{t^{k-1}\prod_{i=1}^{n}\frac{t-b_i}{t-a_i}\} = (1-c)\frac{\langle A^k r_j, r\rangle}{\langle r, r\rangle}$$

which in turn implies that the numbers

$$(c-1)\frac{\langle A^k r, r\rangle}{\langle r, r\rangle} = \mathrm{Res}_{t=\infty}\{t^{k-1}\prod_{i=1}^{n}\frac{t-b_i}{t-a_i}\}$$

are integers for all $k \geq 1$. Recall that $c = -1$.

With the normalization $\langle r, r\rangle = -2$ and under the assumption $\langle A^k r, r\rangle$ being odd for at least one $k \in \mathbb{Z}$ we see that the integral span $L = \sum A^k r$ is an even Lorentzian lattice. In turn we can calculate (by computer and case by case) the invariant factors of the finite Abelian group L^*/L with $L^* = \{v \in \mathbb{Q} \otimes L; \langle v, l\rangle \in \mathbb{Z}\ \forall\ l \in L\}$ the dual rational Lorentzian lattice. The question we like to understand is whether $H(a, b) < \mathrm{O}(L)$ has finite or infinite index, or equivalently whether $H(a, b)$ is arithmetic on thin.

Lemma 3.35. *For $\langle v, v\rangle = -2$ we denote by $u \mapsto i_v(u) = u + \langle u, v\rangle v$ the orthogonal involution of $\mathbb{R} \otimes L$. Then the involutions i_v for $v = A^k r$ and some $k \in \mathbb{Z}$ generate a finite index normal subgroup $N(a, b)$ of $H(a, b)$.*

Proof. The group $H(a, b)$ is generated by the elements A and C, and so the elements $A^k C A^{-k}$ for $k \in \mathbb{Z}$ generate a normal subgroup $N(a, b)$ with index a divisor of the (finite by assumption) order of A. □

Lemma 3.36. *For $\langle x, x\rangle = 2$ we denote by $u \mapsto s_x(u) = u - \langle u, x\rangle x$ the orthogonal reflection of $\mathbb{R} \otimes L$. If $\langle v, v\rangle = -2, \langle w, w\rangle = -2, \langle v, w\rangle = -3$ then*

$$i_v i_w = s_x s_y$$

with $x = v - w, y = v - 2w$ and $\langle x, x\rangle = \langle y, y\rangle = 2, \langle x, y\rangle = 3$.

Proof. This is a straightforward calculation, since

$$i_v i_w(u) = i_v(u) + \langle u, w\rangle i_v(w) = u + \langle u, v - 3w\rangle v + \langle u, w\rangle w$$

while

$$s_x s_y(u) = s_x(u) - \langle u, y\rangle s_x(y) = u - \langle u, x - 3y\rangle x - \langle u, y\rangle y$$

which are easily checked to coincide. □

Lemma 3.37. *Let $W_2(L)$ be the normal subgroup of $\mathrm{O}(L)$ generated by the reflections s_x for any norm 2 vector $x \in L$. Define an equivalence relation $\sim$ on the set $S = \{A^k r; k \in \mathbb{Z}\}$ generated by $v \sim w$ if $\langle v, w\rangle = -3$. If the set S is a single equivalence class then the involution subgroup $N(a,b)$ is commensurable with a subgroup of the reflection group $W_2(L)$ of the of the even Lorentzian lattice $L = \sum A^k r$.*

Proof. The subgroup $N_+(a,b)$ of products of an even number of the generating involutions has index two in the full involution subgroup $N(a,b)$ of $H(a,b)$. Let $v, w \in S$. Because S is a single equivalence class there is is a sequence $v_1 = v, v_2, \cdots, v_{n-1}, v_n = w \in S$ with $\langle v_i, v_{i+1}\rangle = -3$. Then we get

$$i_v i_w = (i_{v_1} i_{v_2})(i_{v_2} i_{v_3}) \cdots (i_{v_{n-1}} i_{v_n}) \in W(L)$$

by the previous lemma. □

Suppose for the moment that L is an integral Lorentzian lattice with scalar product $\langle \cdot, \cdot \rangle$ of signature $(n-1, 1)$. A vector $x \in L$ of positive norm is called a root if $2x/\langle x, x\rangle \in L^*$. We denote by $R_k(L)$ the set of all roots in L of norm k and by $R(L) = \sqcup R_k(L)$ the full root system of L. Norm 1 and norm 2 vectors are always roots, but for particular lattices higher norm vectors can be root as well. We denote by

$$s_x(v) = v - 2\frac{\langle v, x\rangle}{\langle x, x\rangle} x$$

the orthogonal reflection with mirror the Lorentzian hyperplane perpendicular to the root $x \in R(L)$. Let $W(L) < \mathrm{O}(L)$ be the subgroup generated by the reflections in all roots of L, and let $W_k(L) < W(L)$ be the subgroup generated by the reflections in norm k roots.

Definition 3.38. *The lattice L is called reflective (respectively k-reflective) if the subgroup $W(L) < \mathrm{O}(L)$ (respectively $W_k(L) < \mathrm{O}(L)$) has finite index.*

Vinberg has deviced an algorithm to decide whether a given Lorentzian lattice is reflective. The idea is simple. Reflection groups have a canonical fundamental domain for the action on the associated hyperbolic space

$$H(L) = \{v \in \mathbb{R} \otimes L; \langle v, v\rangle < 0\}/\mathbb{R}^\times$$

of dimension $(n-1)$. The main theorem of Coxeter group theory says that the closure C of a connected component of the complement of all mirrors in $H(L)$ is a fundamental domain in the strong sense, that is each orbit of $W(L)$ on $H(L)$ intersects the fundamental chamber C is a single point. The reflection group $W(L)$ permutes the set of all chambers in a simply transitive way, and so $H(L)$ has a tessellation by congruent copies of the fundamental chamber C. Think of the Circle Limit IV picture of M.C. Escher. The Vinberg algorithm starts by choosing a "controlling vector" $c \in C$, and determining the "walls" of C in increasing order of hyperbolic distance to c. The Vinberg algorithm might terminate, in which case C has finitely many walls, but it is also possible that C has infinitely many walls. In the latter case C has infinite hyperbolic volume, and $W(L)$ is thin. If the Vinberg algorithm terminates, then one subsequently has to decide whether C has finite or infinite hyperbolic volume, which in turn is equivalent whether $W(L)$ is arithmetic or thin.

The first example that Vinberg worked out (together with Kaplinskaja to deal with the substantial calculations for $n = 18, 19$) was the standard odd unimodular Lorentzian lattice $\mathbb{Z}^{n,1}$ [43].

Theorem 3.39. *The lattice $\mathbb{Z}^{n,1}$ is reflective if and only if $n \leq 19$.*

There are many variations of this theorem. For example, the next result is due to Everitt, Ratcliffe and Tschantz [14] with a quick proof in [20].

Theorem 3.40. *The lattice $\mathbb{Z}^{n,1}$ is 1-reflective if and only if $n \leq 8$.*

The largest rank 22 of a reflective Lorentzian lattice L is an example due to Borcherds [5]. He took for L the even index 2 sublattice of $\mathbb{Z}^{21,1}$. Esselmann has shown that Borcherds' example is optimal [13].

Theorem 3.41. *All reflective Lorentzian lattices have rank at most 22. Moreover, the example found by Borcherds is the unique such lattice in rank 22 and all others have rank at most 20, with the highest rank example of Vinberg and Kaplinskaja showing that the bound 20 is also sharp.*

For our purpose we need the following classfication theorem of 2-reflective even Lorentzian lattices by Nikulin [27].

Theorem 3.42. *Let U be the even unimodular Lorentzian lattice of rank 2. Let K be an even integral Euclidean lattice, and $L = U \oplus K$ the corresponding even integral Lorentzian lattice. If in addition L is 2-reflective then either L is 2-elementary in the sense that $L^*/L \cong (\mathbb{Z}/2\mathbb{Z})^m$ for some $m \in \mathbb{N}$, or K is one of the following root lattices*

rk	K
2	A_2
3	$\mathrm{A}_1 \oplus \mathrm{A}_2, \mathrm{A}_3$
4	$\mathrm{A}_1^2 \oplus \mathrm{A}_2, \mathrm{A}_1 \oplus \mathrm{A}_3, \mathrm{A}_2^2, \mathrm{A}_4$
5	$\mathrm{A}_1^2 \oplus \mathrm{A}_3, \mathrm{A}_1 \oplus \mathrm{A}_2^2, \mathrm{A}_1 \oplus \mathrm{A}_4, \mathrm{A}_2 \oplus \mathrm{A}_3, \mathrm{A}_5, \mathrm{D}_5$
6	$\mathrm{A}_1 \oplus \mathrm{A}_5, \mathrm{A}_1 \oplus \mathrm{D}_5, \mathrm{A}_2^3, \mathrm{A}_2 \oplus \mathrm{A}_4, \mathrm{A}_2 \oplus \mathrm{D}_4, \mathrm{A}_3^2, \mathrm{E}_6$
7	$\mathrm{A}_1 \oplus \mathrm{E}_6, \mathrm{A}_2 \oplus \mathrm{D}_5, \mathrm{A}_3 \oplus \mathrm{D}_4, \mathrm{A}_7, \mathrm{D}_7$
8	$\mathrm{A}_2 \oplus \mathrm{E}_6$
10	$\mathrm{A}_2 \oplus \mathrm{E}_8$
11	$\mathrm{A}_3 \oplus \mathrm{E}_8$

Besides these the only other 2*-reflective even Lorentzian lattices are*

$$\langle -2^k \rangle \oplus \mathrm{D}_4, \langle -6 \rangle \oplus \mathrm{A}_2^2, U(4) \oplus \mathrm{A}_1^3, U(4) \oplus \mathrm{D}_4$$

for $k = 2, 3, 4$. *Here* $U(m)$ *is the rank two lattice* $\mathbb{Z}^{1,1}$ *with scalar product multiplied by* $m \in \mathbb{N}$*, and* $\langle m \rangle$ *is the rank one lattice* $\mathbb{Z}$ *with scalar product multiplied by* $m \in \mathbb{Z}$*. Recall that the invariant factors are* $n+1$ *for* A_n*,* $2, 2$ *or* 4 *for* D_n *if* n *is even or odd respectively, and* $9 - n$ *for* E_n *if* $n = 6, 7, 8$*.*

The moral of this theorem is that there is a rather short list of integral Lorentzian lattices, which are even, 2-reflective but not 2-elementary. By direct inspection these lattices are determined by their rank together with the invariant factors.

We can now explain the approach of Fuchs, Meiri and Sarnak towards their Conjecture 3.29. For an integral Lorentzian hypergeometric group $H(a,b)$ of rank $n \geq 5$ consider the Lorentzian lattice L spanned by the norm 2 vectors $A^k r$ for $k \in \mathbb{Z}$. Assume L is integral and even, which as explained before we can check by computer in plenty of examples. Determine the invariant factors of L^*/L, which again should be done by computer. Suppose that the set $\{A^k r; k \in \mathbb{Z}\}$ is a single equivalence class for the equivalence relation generated by $u \sim v$ if $\langle u, v \rangle = -3$. By the above lemmata the hypergeometric group $H(a,b)$ is commensurable with its normal 2-reflection subgroup $N(a,b)$, which in turn is commensurable with a subgroup of the 2-reflection group $W_2(L)$ of L. But by simple inspection of the invariant factors of L^*/L it so happens that in all examples L does not occur on the above list in the theorem of Nikulin. Hence $W_2(L)$ is thin, and therefore by commensurability a fortiori $H(a,b)$ is thin.

At this point it is good to have the following philosophical remark of the French mathematician René Thom about the nature of mathematics in mind. Thom distinguishes "rich structures" versus "poor structures" in mathematics. Simple groups form a rich structure, general groups a poor structure. Regular convex polytopes form a rich structure, general convex polytopes a poor structure. Finite reflection subgroups of an orthogonal group form a rich structure, while general finite subgroups of an orthogonal group form a poor structure. Now as a general rule of the thumb rich structures become simpler and more rare if the parameters (such as dimension or cardinality) increase, while poor structures become more complicated and more abundant if the parameters increase.

Apparently integral Lorentzian arithmetic hypergeometric groups in rank n are a rich structure, while the general integral Lorentzian hypergeometric groups form a poor structure. The conjecture of Fuchs, Meiri and Sarnak about the absence of integral Lorentzian arithmetic hypergeometric groups for large rank is in accordance with the above expressed philosophy. Probably in general arithmetic groups are a rich structure and thin groups a poor structure. The lesson to be learned of this section is that monodromy groups of period maps from algebraic geometry are often thin. Arithmetic monodromy groups should be the exception and thin monodromy groups the rule.

3.6 Prime Number Theorem after Tchebycheff

In this section we discuss the proof by Tchebycheff of a weak version of the Prime Number Theorem. His proof is very elegant. See also page 622 of the interview from 2005 with Selberg [2]. It was pointed out by Rodriguez-Villegas [32] that a crucial step in this proof of Tchebycheff is the same interlacing property that we encountered in Example 3.17.

Let $\pi(x) = \#\{p; p \le x\}$ denote the standard prime counting function. Introduce the numbers

$$A = \log \frac{2^{\frac{1}{2}} 3^{\frac{1}{3}} 5^{\frac{1}{5}}}{30^{\frac{1}{30}}} = 0.92129022 \cdots \ , \ B = 6A/5 = 1.105550428 \cdots$$

which enter in the argument below. In 1852 Tchebycheff proved in an elementary way the following result towards the Prime Number Theorem [40].

Theorem 3.43. *We have*

$$\frac{Ax}{\log x}(1 + o(x)) < \pi(x) < \frac{Bx}{\log x}(1 + o(x))$$

Introduce the following three prime counting functions for $x > 0$

$$\pi(x) = \sum_{p \le x} 1 \ , \ \theta(x) = \sum_{p \le x} \log p \ , \ \psi(x) = \sum_{p^m \le x} \log p$$

with p always denoting a prime number, and $m = 1, 2, 3, \cdots$ denoting a positive integer. It is obvious that

$$\psi(x) = \sum_{p \le x} \left[\frac{\log x}{\log p} \right] \log p$$

with $[\log x / \log p]$ the largest integer m with $p^m \le x$. In turn

$$\psi(x) = \theta(x) + \theta(x^{\frac{1}{2}}) + \theta(x^{\frac{1}{3}}) + \theta(x^{\frac{1}{4}}) + \cdots$$

is clear as well.

Theorem 3.44. *We have* $\psi(x) = \theta(x) + O(x^{\frac{1}{2}} \log^2 x)$.

Proof. Clearly $\theta(x^{\frac{1}{m}}) = 0$ if $x < 2^m$ or equivalently $\log x / \log 2 < m$. For $m \ge 2$ we have

$$\theta(x^{\frac{1}{m}}) \le (x^{\frac{1}{m}} \log x)/m \le x^{\frac{1}{2}} \log x$$

using $\theta(x) \le x \log x$, which in turn implies that

$$\sum_{m \ge 2} \theta(x^{\frac{1}{m}}) \le x^{\frac{1}{2}} \log x \cdot \frac{\log x}{\log 2} < 2x^{\frac{1}{2}} \log^2 x$$

using $2 \log 2 > 1$. Hence we have

$$\theta(x) \le \psi(x) \le \theta(x) + O(x^{\frac{1}{2}} \log^2 x)$$

which proves the theorem. □

The Prime Number Theorem is usually stated in the form

$$\pi(x) \approx \frac{x}{\log x}$$

but can be reformulated as

$$\psi(x) \approx x$$

and the proof of Tchebycheff will focus on the latter formulation.

Theorem 3.45. *We have* $T(x) \stackrel{\text{def}}{=} \sum_{k \ge 1} \psi(x/k) = \log([x]!)$.

Proof. For $n = [x]$ a natural number the numbers

$$1, 2, \cdots, n$$

include just $[n/p] = [x/p]$ multiples of p, and $[n/p^2] = [x/p^2]$ multiples of p^2, and so on. Hence

$$n! = \prod_p p^{k_p}\ ,\ k_p = \sum_{m \geq 1} [x/p^m]$$

which can be rewritten as

$$\log(n!) = \sum_p k_p \log p = \sum_{p,m} [x/p^m] \log p\ .$$

Observe that

$$[x/p^m] = l \geq 1 \Leftrightarrow x/kp^m \geq 1 \text{ exactly for } k = 1, 2, \cdots, l$$

and therefore (with the sum in the middle term over those triples p, m, k with $p^m \leq x/k$)

$$\log(n!) = \sum_{p,m,k} \log p = \sum_{k \geq 1} \psi(x/k)$$

which proves the theorem. □

The problem is to turn the good asymptotic understanding of $T(x)$ by Stirling's formula into asymptotic understanding of $\psi(x)$. For this purpose Tchebycheff made the following crucial step. If we introduce the function

$$F(x) = T(x) + T(x/30) - T(x/2) - T(x/3) - T(x/5)$$

and use

$$T(x) = \sum_{k \geq 1} \psi(x/k)$$

then we can rewrite

$$F(x) = \sum_{k \geq 1} A_k \psi(x/k)$$

with

$$A_k = \begin{cases} +1 & \text{if } k \text{ is not divisible by } 2, 3, 5 \\ 0 & \text{if } k \text{ is divisible by exactly one number from } 2, 3, 5 \\ -1 & \text{if } k \text{ is divisible by at least two numbers from } 2, 3, 5 \end{cases}$$

For example if k is divisble by 2 but not by $3, 5$ then the term $\psi(x/k)$ enters in $T(x)$ and in $-T(x/2)$, but does not enter in $T(x/30) - T(x/3) - T(x/5)$. Hence $A_k = 0$ in that case. A direct verification shows that

$$A_k = \begin{cases} +1 & \text{if } k \equiv 1, 7, 11, 13, 17, 19, 23, 29 \mod 30 \\ -1 & \text{if } k \equiv 6, 10, 12, 15, 18, 20, 24, 30 \mod 30 \\ 0 & \text{if else} \end{cases}$$

Observe that the two sequences of natural numbers

$$\{k; A_k = +1\} \qquad \{k; A_k = -1\}$$

interlace. It is the same interlacing property that we have seen in Example 3.17.

Corollary 3.46. *We can write*

$$F(x) = \psi(x) - \psi(x/6) + \psi(x/7) - \psi(x/10) + \psi(x/11) - \psi(x/12) + \cdots$$

with alternating plus and minus signs, which in turn implies the key inequality

$$\psi(x) - \psi(x/6) < F(x) < \psi(x)$$

because $\psi(x) = \sum_p [\log x / \log p] \log p$ *is monotonic increasing in* x.

Recall Stirling's formula

$$n! = \sqrt{2\pi n} \exp(n \log n - n + \theta/12n)$$

for some $0 < \theta < 1$.

Corollary 3.47. *Using* $T(x) = \log([x]!)$ *and Stirling's formula we have the inequalities*

$$\frac{1}{2}\log(2\pi) + x\log x - x - \frac{1}{2}\log x < T(x)$$

$$T(x) < \frac{1}{2}\log(2\pi) + x\log x - x + \frac{1}{2}\log x + 1/12$$

as lower and upper bound for $T(x)$.

Corollary 3.48. *Using* $F(x) = T(x) + T(x/30) - T(x/2) - T(x/3) - T(x/5)$ *we have the inequalities*

$$F(x) < Ax + \frac{5}{2}\log x - \frac{1}{2}\log(1800\pi) + 2/12 < Ax + \frac{5}{2}\log x$$

$$F(x) > Ax - \frac{5}{2}\log x + \frac{1}{2}\log(450/\pi) - 3/12 > Ax - \frac{5}{2}\log x$$

with $A = \frac{1}{2}\log 2 + \frac{1}{3}\log 3 + \frac{1}{5}\log 5 - \frac{1}{30}\log 30 = 0.92129022\cdots$.

Using the key inequality of Corollary 3.46

$$\psi(x) - \psi(x/6) < F(x) < \psi(x)$$

we get

$$Ax - \frac{5}{2}\log x < \psi(x)\ ,\ \psi(x) - \psi(x/6) < Ax + \frac{5}{2}\log x$$

and the second inequality can be iterated. Indeed

$$\psi(x) < Ax + \frac{5}{2}\log x + \psi(x/6)$$

$$< Ax(1 + 1/6) + \frac{5}{2}(2\log x - \log 6) + \psi(x/6^2)$$

$$< Ax(1 + 1/6 + 1/6^2) + \frac{5}{2}(3\log x - (1 + 2)\log 6) + \psi(x/6^3)$$

$$< Ax(1+1/6+\cdots+1/6^m)+\frac{5}{2}((m+1)\log x-\frac{1}{2}m(m+1)\log 6)+\psi(x/6^{m+1})$$

$$< \frac{6}{5}Ax + O(\log^2 x)$$

since

$$\psi(x/6^{m+1}) = 0 \Leftrightarrow x/6^{m+1} < 2 \Leftrightarrow (m+1) > \frac{\log(x/2)}{\log 6}$$

This ends our discussion of the proof of the following theorem of Tchebycheff.

Theorem 3.49. *We have*

$$Ax + O(\log x) < \psi(x) < Bx + O(\log^2 x)$$

with $A = 0.92129022\cdots$ *and* $B = 6A/5 = 1.105550428\cdots$ *.*

Equivalently we arrive at

$$\frac{Ax}{\log x}(1 + o(x)) < \pi(x) < \frac{Bx}{\log x}(1 + o(x))$$

and so

$$\pi(x) \asymp \frac{x}{\log x}$$

which is Tchebycheff's weak version of the Prime Number Theorem.

Remark 3.50. *The proof of Tchebycheff has two main ideas. The first step is to work with the prime counting function $\psi(x)$ instead of the usual function $\pi(x)$, and to consider the function $T(x) = \sum \psi(x/k) = \log([x]!)$.*

The second step is to turn good asymptotic understanding for $T(x)$ from Stirling's formula into good asymptotic understanding for $\psi(x)$. A first try might be to consider

$$F(x) = T(x) - 2T(x/2) = \psi(x) - \psi(x/2) + \psi(x/3) - \psi(x/4) + \cdots$$

which in turn implies that

$$F(x) < \psi(x) < F(x) + \psi(x/2) \ .$$

Now the same method of proof works in a simpler way leading to

$$Ax + O(\log x) < \psi(x) < Bx + O(\log x)$$

with $A = \log 2 = 0.693\cdots$ and $B = 2\log 2 = 1.386\cdots$. Having established this special case first it might be not unreasonable to try

$$F(x) = T(x) + T(2x/N) - T(x/p) - T(x/q) - T(x/r)$$

with $p \geq q \geq r \geq 2$ and $1/p + 1/q + 1/r = 1 + 2/N$. Of course, the dihedral case $(p, q, r, N) = (m, 2, 2, 2m)$ gives back the previous case. There are just a few other possibilities

p	q	r	N
m	2	2	$2m$
p	3	2	$12p/(6-p)$

with $p = 3, 4, 5$ and $N = 12, 24, 60$. These numbers are also familiar from the classification of the Platonic solids (tetrahedron, octahedron and icosahedron).

For $q = 3, r = 2$ the coefficients A_k for $p = 3$ are given by

$$A_k = \begin{cases} +1 & \text{if } k \equiv 1, 5 \text{ mod } 6 \\ -1 & \text{if } k \equiv 3, 6 \text{ mod } 6 \\ 0 & \text{if else} \end{cases}$$

and for $p = 4$ become

$$A_k = \begin{cases} +1 & \text{if } k \equiv 1, 5, 7, 11 \text{ mod } 12 \\ -1 & \text{if } k \equiv 4, 6, 8, 12 \text{ mod } 12 \\ 0 & \text{if else} \end{cases}$$

and so the same interlacing property holds for all these cases.

If $q = 3, r = 2$ *and* $p = 3, 4, 5$ *and so* $N = 12p/(6-p) = 12, 24, 60$ *respectively one gets*

$$A = \tfrac{1}{2}\log 2 + \tfrac{1}{3}\log 3 + \tfrac{1}{p}\log p - \tfrac{2}{N}\log(N/2)$$

which amounts to $A = 0.780\cdots$ *for* $p = 3$, $A = 0.852\cdots$ *for* $p = 4$ *and* $A = 0.921\cdots$ *for* $p = 5$. *Likewise* $B = 3A/2 = 1.171\cdots$ *for* $p = 3$, $B = 4A/3 = 1.136\cdots$ *for* $p = 4$ *and* $B = 6A/5 = 1.105\cdots$ *for* $p = 5$. *All in all, the method gives the sharpest bounds for the icosahedron with* $(p, q, r, N) = (5, 3, 2, 60)$, *and this is the case discussed by Tchebycheff.*

3.7 Exercises

Exercise 3.1. *Verify using Theorem 3.1 that the Riemann scheme of the hypergeometric function* ${}_nF_{n-1}(\alpha; \beta|z)$ *is of the form*

0	1	∞
$\underline{1} - \beta$	$0, 1, \cdots, n-2, \gamma$	α

with $\gamma = -1 + \sum_1^n (\beta_j - \alpha_j)$.

Exercise 3.2. *Show that the Clausen–Thomae hypergeometric equation is characterized among the Fuchsian equations with regular singular points at* $z = 0, 1, \infty$ *by its Riemann scheme, together with the fact the point* $z = 1$ *is a "special regular singular" point in the sense that all coefficients* a_j *of the linear differential equation*

$$(\partial^n + a_1\partial^{n-1} + \cdots + a_n)f = 0$$

have at most simple poles at $z = 1$.

Exercise 3.3. *Show that for a linear differential equation*

$$(\partial^n + a_1\partial^{n-1} + \cdots + a_n)f = 0$$

with a regular singular point at $z = 0$ *the property that there exist local holomorphic solutions* $f(z)$ *around* $z = 0$ *with* $\partial^{j-1}f(0)$ *freely prescribed for* $j = 1, \cdots, n-1$ *implies that* $z = 0$ *is a special regular singular point, in the sense that* za_j *are holomorphic around* $z = 0$ *for* $j = 1, \cdots, n$.

Exercise 3.4. *Check that for* β_i *all distinct modulo* $\mathbb{Z}$ *the functions*

$$(-z)^{1-\beta_i} {}_nF_{n-1}(\alpha + (1-\beta_i)\underline{1}; \beta + (1-\beta_i)\underline{1}|z)$$

defined on domain $\mathbb{C} - [0, \infty)$ *form a Kummer basis around* $z = 0$ *for the hypergeometric equation with parameters* α, β. *Conclude from Theorem 3.3 that the monodromy representation for the Clausen–Thomae hypergeometric equation is in principle explicitly computable as function of the parameters.*

Exercise 3.5. *After multiplication of Whipple's quadratic transformation formula by* $z^{\alpha/2}$ *the limit for* $z \uparrow 1$ *on the left hand side, which is given by Dixon's formula, can be evaluated by taking the limit for* $w \to -\infty$ *on the right hand side, which in turn can be evaluated using Kummer's continuation formula. Check the details of this proof of Dixon's formula, using the duplication formula* $\Gamma(x/2)\Gamma((x+1)/2) = 2^{1-x}\sqrt{\pi}\,\Gamma(x)$ *for the* Γ*-function.*

Exercise 3.6. *Show that the linear span of all products* gh *with the functions* g, h *solutions of a Fuchsian equation* $(\partial^n + a_1\partial^{n-1} + \cdots + a_n)f = 0$ *of order* n *are solutions of a Fuchsian equation of order* $n(n+1)/2$, *which is called the second symmetric power of the original equation.*

Exercise 3.7. *A matrix* $A \in \mathrm{Mat}_n(\mathbb{C})$ *is called regular if its commutant* $\{X \in \mathrm{Mat}_n(\mathbb{C}); AX = XA\}$ *has dimension* n. *Show using Jordan normal form that* A *is regular if and only if for each eigenvalue* a *of* A *the eigenspace* $\ker(A - a)$ *has dimension one. Show that the matrix*

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -A_n \\ 1 & 0 & \cdots & 0 & -A_{n-1} \\ 0 & 1 & \cdots & 0 & -A_{n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -A_1 \end{pmatrix}$$

is a regular with eigenvalues $a_1, \cdots, a_n$ *given by* $\det(t - A) = \prod(t - a_i)$ *and so such matrices* A *form a slice of dimension* n *for the conjugation orbits of regular matrices.*

Exercise 3.8. *Suppose* $H < \mathrm{GL}_n(\mathbb{C})$ *is an irreducible subgroup and let* $b : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ *be a nonzero bilinear form that is is invariant under* H.

- *Show that* b *is nondegenerate.*
- *Show that* b *is unique up to a nonzero scalar.*

- *For* $u, v \in \mathbb{C}^n$ *let us denote*

$$g(u,v) = [b(u, v + b(v,u))]/2 \quad , \quad \omega(u,v) = [b(u,v) - b(v,u)]/2$$

the corresponding symmetrized and skew symmetrized bilinear forms. Show that one of these forms is zero while the other is nondegenerate.

- *Suppose there exists an element* h *in* H *with* $\mathrm{rk}(h-1) = 1$ *and so one has* $h(v) = v + f(v)w$ *for all* $v \in \mathbb{C}^n$ *and some nonzero vectors* $f \in \mathrm{Hom}(\mathbb{C}^n, \mathbb{C})$ *and* $w \in \mathbb{C}^n$*. Show that* $\det(h) = 1 + f(w)$*. Show that*

$$f(u)b(w,v) + f(v)b(u,w) + f(u)f(v)b(w,w) = 0$$

for all $u, v \in \mathbb{C}^n$*, and hence*

$$f(w)b(w,v) + f(v)(1 + f(w))b(w,w) = 0$$

for all $v \in \mathbb{C}^n$*. Show that* $f(w) \neq 0$ *implies that* $b = g$ *is symmetric (so* H *is a subgroup of the orthogonal group), while* $f(w) = 0$ *implies that* $b = \omega$ *is skew symmetric (so* H *is a subgroup of the symplectic group). Finally remark that in case* $b = g$ *we have* $\det(h) = -1$*, while* $\det(h) = 1$ *in case* $b = \omega$*.*

Exercise 3.9. *In the notation of Example 3.12 show that for* $j = 1, \cdots, n$ *with* $\gcd(j, n+1) = 1$ *the symmetric group* S_{n+1} *is generated by the two permutations* $s_1 \cdots s_{j-1} s_{j+1} \cdots s_n$ *and* s_j*, and conclude that the hypergeometric function*

$${}_nF_{n-1}(\tfrac{1}{n+1}, \cdots, \tfrac{n}{n+1}; \tfrac{1}{j}, \cdots, \tfrac{j-1}{j}, \tfrac{1}{n+1-j}, \cdots, \tfrac{n-j}{n+1-j}|z)$$

is algebraic with monodromy group isomorphic to S_{n+1}*. Hint: The special case* $j = 1$ *is easy, and the general case can be reduced to it.*

Exercise 3.10. *Show that the hypergeomtric group* $H(a,b)$ *with parameters* $a = \sqrt[2n]{1} - \sqrt[n]{1}$ *and* $b = \sqrt[n]{1}$ *is isomorphic (as subgroup of* $W(\mathrm{B}_n)$*) to* $C_2^n \rtimes C_n$ *with* C_n *the cyclic group of order* n*. Note that the case* $n = 3$ *was already discussed in Example 3.24.*

Exercise 3.11. *Show that the hypergeometric function* ${}_3F_2(\alpha; \beta|z)$ *with parameters* α *equal to* $(1/10, 1/2, 9/10)$ *and* β *given by*

β
$(1/5, 4/5)$
$(1/3, 2/3)$

is algebraic with monodromy group $W(\mathrm{H}_3)$.

Likewise show that the hypergeometric function ${}_4F_3(\alpha;\beta|z)$ *with parameters* α *equal to* $(1/30, 11/30, 19/30, 29/30)$ *and* β *given by*

β
$(1/10, 1/2, 9/10)$
$(1/5, 1/2, 4/5)$
$(1/3, 1/2, 2/3)$
$(1/4, 1/2, 3/4)$

is algebraic with monodromy group $W(\mathrm{H}_4)$. *Compare these tables (and the ones of the previous and next exercises) with the tables of algebraic hypergeometric functions in [4].*

Exercise 3.12. *Show that the hypergeometric function* ${}_6F_5(\alpha;\beta|z)$ *with parameters* α *equal to* $(1/12, 1/3, 5/12, 7/12, 2/3, 11/12)$ *and* β *given by*

β
$(1/8, 3/8, 1/2, 5/8, 7/8)$
$(1/5, 2/5, 1/2, 3/5, 4/5)$

is algebraic with monodromy group $W(\mathrm{E}_6)$.

Likewise show that the hypergeometric function ${}_7F_6(\alpha;\beta|z)$ *with parameters* α *equal to* $(1/18, 5/18, 7/18, 1/2, 11/18, 13/18, 17/18)$ *and* β *given by*

β
$(1/7, 2/7, 3/7, 4/7, 5/7, 6/7)$
$(1/5, 1/3, 2/5, 3/5, 2/3, 4/5)$
$(1/12, 1/3, 5/12, 7/12, 2/3, 11/12)$

is algebraic with monodromy group $W(\mathrm{E}_7)$.

Finally show that the hypergeometric function ${}_8F_7(\alpha;\beta|z)$ *with parameters* α *equal to* $(1/30, 7/30, 11/30, 13/30, 17/30, 19/30, 23/30, 29/30)$ *and* β *given by*

β
$(1/12, 1/4, 5/12, 1/2, 7/12, 3/4, 11/12)$
$(1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8)$
$(1/7, 2/7, 3/7, 1/2, 4/7, 5/7, 6/7)$
$(1/5, 1/3, 2/5, 1/2, 3/5, 2/3, 4/5)$
$(1/5, 1/4, 2/5, 1/2, 3/5, 3/4, 4/5)$
$(1/8, 1/3, 3/8, 1/2, 5/8, 2/3, 7/8)$
$(1/12, 1/3, 5/12, 1/2, 7/12, 2/3, 11/12)$
$(1/18, 5/18, 7/18, 1/2, 11/18, 13/18, 17/18)$

is algebraic with monodromy group $W(\mathrm{E}_8)$.

Exercise 3.13. *A glance at the table of parameters for algebraic hypergeometric functions in [4] shows that for example for $W(\mathrm{E}_8)$ there are more cases than the ones found in the previous exercise, notably for α having fractions with denominators 20 and 24. Together with 30 these are exactly the regular degrees in the sense of Springer [37]. Springer studies regular elements of such orders as generalization of Coxeter elements. An open question is whether these interlacing parameter sets can be understood using such "Springer elements" as a generalization of the results of the previous exercise for Coxeter elements?*

Exercise 3.14. *The monodromy group of Example 2.6 is a hypergeometric group with parameters*

$$\{\zeta_{12}\zeta_{2p}, \zeta_{12}/\zeta_{2p}\}\ ,\ \{\zeta_3^2, 1\}$$

and $\zeta_k = \exp(2\pi i/k)$. Dividing both generators A and B by ζ_3 gives a hypergeometic group $H(a,b)$ with parameters

$$a = (-\zeta_4\zeta_{2p}, -\zeta_4/\zeta_{2p})\ ,\ b = (\zeta_3, \zeta_3^2)$$

with $p \neq 6$ to assure irreducibility. Show that for $p \geq 7$ all Galois conjugates of $H(a,b)$ different from the identity and complex conjugation have parameters that interlace on the unit circle if and only if

$$p = 7, 8, 9, 10, 11, 12, 14, 16, 18, 24, 30, \infty\ .$$

This list was found by Fricke and Klein, and extended by Takeuchi to a complete list of arithmetic triangle groups [15], [39].

Exercise 3.15. *Consider $\mathbb{C}^2$ with a definite Hermitian form $\langle\cdot,\cdot\rangle$. For $X \in \mathrm{End}(\mathbb{C}^2)$ the adjoint $X^\dagger \in \mathrm{End}(\mathbb{C}^2)$ is defined by $\langle Xu, v\rangle = \langle u, X^\dagger v\rangle$ for all $u, v \in \mathbb{C}^2$. Show that the group*

$$\mathrm{SU}_2^\pm(\mathbb{C}) = \{A \in \mathrm{GL}(\mathbb{C}^2); A^\dagger A = 1, \det(A) = \pm 1\}$$

acts on the real vector space of traceless Hermitian operators

$$\mathrm{Herm}_2(\mathbb{C}) = \{X \in \mathrm{End}(\mathbb{C}^2); X^\dagger = X, \mathrm{tr}(X) = 0\}$$

by conjugation, leaving the trace form $(X,Y) \mapsto \mathrm{tr}(XY^\dagger)$ invariant. Show that the induced group homomorphism

$$\mathrm{SU}_2^\pm(\mathbb{C}) \twoheadrightarrow \mathrm{O}_3(\mathbb{R})\ ,$$

called the spin homomorphism, is surjective with kernel of order 2*. In the next exercise we treat the Lorentzian analogue of this well known Euclidean situation.*

Exercise 3.16. *Consider* $\mathbb{C}^2$ *with an indefinite Hermitian form* $\langle \cdot, \cdot \rangle$. *For* $X \in \mathrm{End}(\mathbb{C}^2)$ *the adjoint* $X^* \in \mathrm{End}(\mathbb{C}^2)$ *is defined by* $\langle Xu, v \rangle = \langle u, X^* v \rangle$ *for all* $u, v \in \mathbb{C}^2$. *Show that* $X \mapsto X^*$ *is an antilinear involution on* $\mathrm{End}(\mathbb{C}^2)$. *Show that the Hermitian form* $\langle X, Y \rangle = \mathrm{tr}(XY^*)$ *on* $\mathrm{End}(\mathbb{C}^2)$ *turns the space*

$$\mathrm{Herm}_{1,1}(\mathbb{C}) = \{X \in \mathrm{End}(\mathbb{C}^2); X^* = X, \mathrm{tr}(X) = 0\}$$

of traceless self adjoint operators into a real Lorentz space $\mathbb{R}^{2,1}$ *of signature* $(2,1)$. *Show that the group*

$$\mathrm{SU}^{\pm}_{1,1}(\mathbb{C}) = \{A \in \mathrm{GL}(\mathbb{C}^2); A^* A = 1, \det(A) = \pm 1\}$$

acts on $\mathrm{Herm}_{1,1}(\mathbb{C})$ *by conjugation. Show that the associated spin homomorphism* $\mathrm{SU}^{\pm}_{1,1}(\mathbb{C}) \to \mathrm{O}(\mathbb{R}^{2,1})$ *has both kernel and cokernel of order* 2.

References

[1] D.V. Anosov and A.A. Bolibruch, The Riemann–Hilbert problem, Aspects of Mathematics 22, Vieweg, Braunschweig, 1994.

[2] N.A. Baas and Ch.F. Skau, The lord of the numbers, Atle Selberg. On his life and mathematics, Bulletin AMS 45 (2008), 617-649.

[3] W.N. Bailey, Generalized Hypergeometric Series, Cambidge University Press, Cambridge, 1935.

[4] F. Beukers and G.J. Heckman, Monodromy for the hypergeometric function ${}_nF_{n-1}$, Invent. Math. 95 (1989), 325-354.

[5] R. Borcherds, Automorphism Groups of Lorentzian Lattices, Journal of Algebra 111 (1987), 133-153.

[6] N. Bourbaki, Éléments de Mathematiques, Groupes et Algèbres de Lie, Chapitres 4,5 et 6, Masson, Paris, 1981.

[7] T. Clausen, Über die Fälle, wenn die Reihe von die Form $y = 1 + \cdots$ ein Quadrat von der form $z = 1 + \cdots$ hat, J. Reine Angew. Math. 3 (1828), 89-91.

[8] H.S.M. Coxeter, Finite groups generated by unitary reflections, Abh. Math. Sem. Hamburg 31 (1967), 125-135.

[9] W.J.M. Dekkers, Moduli spaces for pairs of 2×2 matrices and for certain connections on $\mathbb{P}^1(\mathbb{C})$, PhD, Radboud University Nijmegen, 1976.

[10] Pierre Deligne, Équatiosn Différentielles à Points Singuliers Réguliers, Lecture Notes in Mathematics 163, Springer-Verlag, Berlin, 1970.

[11] A.C. Dixon, Summation of a certain series, Proc. London Math. Soc. 35:1 (1902), 284-291.

[12] F. Dyson, Birds and Frogs, Notices of the AMS 56:2 (2009), 212-213.

[13] F. Esselmann, Über die maximale Dimension von Lorentz-Gittern mit coendlicher Spiegelungsgruppe, J. Number Theory 61 (1996), 103-144.

[14] Brent Everitt, John Ratcliffe and Steven T. Tschantz, Right-angled Coxeter polytopes, hyperbolic 6-manifolds and a problem of Siegel, Preprint 2010.

[15] K.E.R. Fricke and F. Klein, Vorlesungen über die Theorie der automorphen Funktionen I, Teubner, Leipzig, 1897.

[16] Elena Fuchs, Chen Meiri and Peter Sarnak, Hyperbolic monodromy groups for the hypergeometric equation and its Cartan involutions, arXiv:math.GR/13050729, 2013.

[17] Harish-Chandra, Spherical Functions on a Semisimple Lie Group I, Amer. J. Math 80 (1958), 241-310.

[18] Harish-Chandra, Spherical Functions on a Semisimple Lie Group II, Amer. J. Math 80 (1958), 553-613.

[19] Gert Heckman, Coxeter groups, Informal Lecture Notes, Fall 2013.

[20] Gert Heckman and Sander Rieken, An Odd Presentation for $W(\mathrm{E}_6)$, Preprint 2014.

[21] J.E. Humphreys, Reflection Groups and Coxeter Groups, Cambridge Studies in Advanced Mathematics 29, Cambridge University Press, Cambridge, 1990.

[22] C. Jordan, Cours d' Analyse, 1887.

[23] N. Katz, Rigid Local Systems, Annals of Mathematics Studies 139, Princeton University Press, Princeton, 1995.

[24] F. Klein, Vorlesungen über die hypergeometrische Funktion, Springer Verlag, Berlin, 1933.

[25] M. Kontsevich and D. Zagier, Periods, Mathematics Unlimited – 2001 and Beyond, B. Engquist and W. Schmid editors, Springer-Verlag, Berlin (2001), 771-808.

[26] A.H.M. Levelt, Hypergeometric Functions, PhD, University of Amsterdam, 1961.

[27] V.V. Nikulin, On factor groups of the automorphism groups of hyperbolic forms modulo subgroups generated by 2-reflections, Soviet Math. Dokl. 20 (1979), 1156-1158.

[28] E. Papperitz, Über verwandte $S-$Funktionen, Mathematische Annalen 25 (1885), 212-221.

[29] L. Pochhammer, Zur Theorie der Euler'schen Integrale, Mathematische Annalen 35 (1890), 495-526.

[30] A. Pressley, Elementary Differential Geometry, Springer Undergraduate Mathematical Series, 2008.

[31] G.F.B. Riemann, Beiträge zur Theorie der durch die Gaussische hypergeometrische Reihe $F(\alpha, \beta, \gamma; x)$ darstellbaren Funktionen (1857), Gesammelte mathematische Werke, Teubner, Leipzig, 1892.

[32] F. Rodriguez-Villegas, Integral ratios of factorials and algebraic hypergeometric functions, Notes from a lecture at Oberwolfach, 2004, http://www.ma.utexas.edu/users/villegas/publications.html

[33] H.A. Schwarz, Über diejenigen Fälle in welchen die Gaussische hypergeometrische Reihe einer algebraische Funktion ihres vierten Elementes darstellt, Crelle J. 75 (1873), 292-335.

[34] Jean-Pierre Serre, Cours d'arithmétique, Presses Universitaires de France, 1970.

[35] G.C. Shephard and J.A. Todd, Finite unitary reflection groups, Canadian J. Math. 6 (1954), 274-304.

[36] C.T. Simpson, Product of Matrices, Canadian Mathematical Society Conference Proceeding, Volume 12 (1992), 157-185.

[37] T.A. Springer, Regular elements of finite reflection groups, Invent. Math. 25 (1974), 159-198.

[38] I. Stewart and D.O. Tall, Algebraic Number Theory and Fermat's Last Theorem, Third Edition, Taylor and Francis, UK, 2001.

[39] K. Takeuchi, Arithmetic triangle groups, J. Math. Soc. Japan 29 (1977), 91-106.

[40] P.L. Tchebycheff, Mémoire sur les nombres premiers, J. Math. Pures et Appliquées XVII (1852), 366-390.

[41] J. Thomae, Über die höheren hypergeometrische Reihen, Math. Ann. 2 (1870), 427-444.

[42] E.C. Titchmarsh, Eigenfunction Expansions, Clarendon Press, Oxford, 1962.

[43] È.B. Vinberg, Hyperbolic reflection groups, Russian Math. Surv. 40 (1980), 31-75.

[44] F.J.W. Whipple, Some transformations of generalized hypergeometric series, Proc. London Math. Soc. 26:2 (1927), 257-272.

[45] E.T. Whittaker and G.N. Watson, A Course in Modern Analysis, Cambridge University Press, Cambridge, 1963.