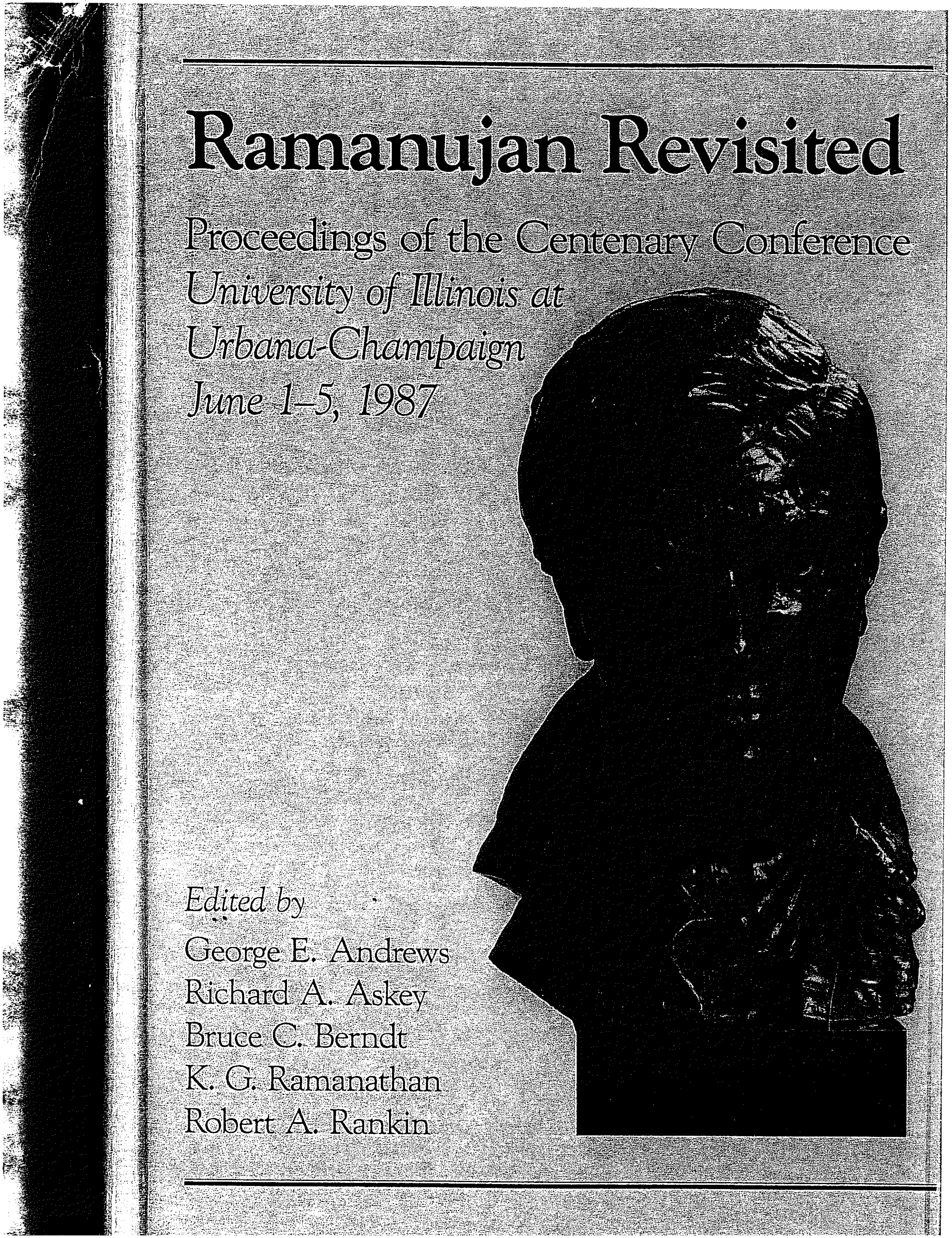

Approximations and complex multiplication according to Ramanujan

by

D.V. Chudnovsky*), G.V. Chudnovsky*)

Department of Mathematics
Columbia University
New York, NY 10027

*)This work was supported in part by the N.S.F., U.S. Air Force and O.C.R.E.A.E. program.

Introduction

This talk revolves around two focuses: complex multiplications (for elliptic curves and Abelian varieties) in connection with algebraic period relations, and (diophantine) approximations to numbers related to these periods. Our starting point is Ramanujan's works [1], [2] on approximations to π via the theory of modular and hypergeometric functions. We describe in chapter 1 Ramanujan's original quadratic period--quasiperiod relations for elliptic curves with complex multiplication and their applications to representations of fractions of π and other logarithms in terms of rapidly convergent nearly integral (hypergeometric) series. These representations serve as a basis of our investigation of diophantine approximations to π and other related numbers. In Chapter 2 we look at period relations for arbitrary CM-varieties following Shimura and Deligne. Our main interest lies with modular (Shimura) curves arising from arithmetic Fuchsian groups acting on H. From these we choose arithmetic triangular groups, where period relations can be expressed in the form of hypergeometric function identities. Particular attention is devoted to two (commensurable) triangle groups, (0,3;2,6,6) and (0,3;2,4,6), arising from the quaternion algebra over $\mathbb{Q}$ with the discriminant $D = 2\cdot 3$. We also touch upon the algebraic

independence problem for periods and quasiperiods of general CM-varieties and particularly CM-curves associated with the triangle groups (hypergeometric curves as we call them). The diophantine approximation problem for numbers connected with periods, particularly for multiples of π, is analyzed using Padé approximations to power series representing these numbers. We give a brief review of Padé approximations, their effective construction, and problems of analytic and arithmetic (p-adic) convergence of Padé approximants. Padé approximations to nearly integral power series (G-functions) are used in connection with Ramanujan-like representations of $1/\pi$ and other similar period constants. We discuss measures of irrationality for algebraic multiples of π and related numbers that follow from Padé approximation methods.

The problem of uniformization by nonarithmetic subgroups is discussed in connection with the Whittaker conjecture [11] on an explicit expression for accessory parameters in the (Schottkey-type) uniformization of hyperelliptic Riemann surfaces of genus $g \geq 2$ by Fuchsian groups. On the basis of numerical computations of monodromy groups of linear differential equations, we concluded that the conjecture [11] is generically incorrect. Moreover, it appears that accessory parameters in the uniformization problem of Riemann surfaces defined over $\bar{\mathbb{Q}}$ are nonalgebraic with the exception of uniformization by arithmetic subgroups and of cases when the differential equations are reduced to hypergeometric ones (the monodromy group is connected to one of the triangle groups). We briefly describe numerical and theoretical results on the transcendence of elements of the monodromy groups of linear differential equations over $\bar{\mathbb{Q}}(x)$.

We conclude the paper with a discussion of numerical approximations to solutions of algebraic and differential equations. We present generalizations of our previous results [12] on the complexity of approximations to solutions of linear differential equations. We describe a new, "bit-burst" method of evaluation of solutions of linear differential equations everywhere on their Riemann surface. In the worst case, an evaluation with n bits of precision at an n-bit point requires $O(M_{bit}(n)\log^3 n)$ boolean (bit) operations, where $M_{bit}(n)$ is the boolean complexity of an n-bit multiplication. For functions satisfying additional arithmetic conditions (e.g. E-functions and G-functions) the factor $\log^3 n$ could be further decreased to $\log^2 n$ or, even, $\log n$. We also describe the natural parallelizations of the presented algorithms that are well suited for practical implementation.

We want to thank the organizers of the Ramanujan conference and, particularly, B. Berndt and R. Askey. We want to thank the IBM Computer Algebra Group for their help and access to the SCRATCHPAD System. We thank SYMBOLICS Corporation for the use of their workstation 3645.

1. Complex multiplications and Ramanujan's period relations

Most of the material in this talk evolves around mathematics closely associated with one of the earliest Ramanujan papers "Modular equations and approximations to π" [1] published in 1914, which according to Hardy [2] "is mainly Indian work, done before he came to England." In that or another way the same kind of mathematics appears in later Ramanujan research, including his notebooks. Hardy interpreted many of Ramanujan's results and identities as connected mainly with "complex multiplication", and Ramanujan's interest in resolving modular equations in explicit radicals was later picked up in Watson's outstanding series [3] of "Singular Moduli" papers. Singular moduli themselves, and general modular equations relating automorphic functions with respect to congruence subgroup of a full modular group are traditional subjects of late $\overline{\text{XIX}}$ century mathematics, whose importance is clearly realized in modern number theory and algebraic geometry, particularly in diophantine geometry in connection with arithmetic theory of elliptic curves and rational points on them. Also modular equations turned up as a convenient tool of fast operational complexity algorithms of computation of π and of values of elementary functions (see corresponding chapters in Borweins' book [5]).

Instead of complex multiplication as merely a subject of "singular moduli" of elliptic functions we will touch upon the complex multiplication in a slightly more general context: from the point of view of nontrivial endomorphisms of certain classes of Jacobians of particular algebraic curves. (We are not going to discuss a variety of complex "complex multiplication" subjects on L-functions and Abelian varieties, though we'll have to borrow particular consequences of vast theories developed in general by Shimura, Deligne and others.)

The choice of curves is clearly determined by Ramanujan's interest: these are curves with 4 critical points, whose Abelian periods are expressed via hypergeometric integrals (for simplicity one can call those curves hypergeometric ones), see [10].

Transforming one of the 4 critical points into infinity and normalizing, one recovers the Gauss hypergeometric function integrals representative of these periods. We display these well-known expressions:

$$F(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \cdot \int_0^1 x^{b-1}(1-x)^{c-b-1}(1-xz)^{-a}dx,$$

$(c > b > 0)$, with the expansion near $z = 0$:

$$F(a,b;c;z) = 1 + \frac{a\cdot b}{c\cdot 1} \cdot z + \frac{a(a+1)\cdot b(b+1)}{c\cdot(c+1)\cdot 1\cdot 2} \cdot z^2 + \dots .$$

Our primary interest in complex multiplication is not arithmetical but transcendental; rather than to study the number-theoretical functions associated with complex multiplication invariants, we want to know the basic facts about the values of these invariants: are these numbers transcendental (over $\mathbb{Z}$)? If algebraic relations do exist, over what fields of definition do they exist? These basic questions of transcendence, algebraic independence and linear independence are the subject of diophantine approximations. With these questions come their qualitative counterparts: when numbers are irrational (transcendental), how well are they approximated by rationals? Can these best approximations be determined effectively and/or explicitly? (Usually one asks in this context: can one determine the continued fraction expansion of the number?)

The class of numbers we are interested in is generated over $\bar{\mathbb{Q}}$ by periods and quasiperiods of Abelian varieties, i.e. by integrals

$$\int_\gamma \omega \quad \text{and} \quad \int_\gamma \eta$$

for differentials ω and η of the first and the second kind, respectively, from $H^1_{DR}(A)$, and $\gamma \in H_1(A,\mathbb{Z})$, for an Abelian variety A defined over $\bar{\mathbb{Q}}$.

In particular, when A is a Jacobian $J(\Gamma)$ of a non-singular curve Γ over $\bar{\mathbb{Q}}$, we are looking at periods and quasiperiods forming the full Riemann matrix of Γ-total of $2g \times 2g$ elements, where g denotes the genus of Γ.

In this context, complex multiplications, understood as nontrivial endomorphisms of A, are usually expressed as non-trivial algebraic relations among the elements of Riemann's original $2g \times g$ pure period matrix π of A. These "purely period relations" are well known, and are mainly algebraic in nature, and in one-dimensional case ($g = 1$), give pairs of periods whose ratio is a "singular module". An interesting thing discovered by Ramanujan in this classical (even in his time) field was the existence of new quasiperiods relations. In the Weierstrass-like notation, commonly accepted nowdays, period and quasiperiod relations in the elliptic curve case can be described as follows.

One starts with an elliptic curve over $\bar{\mathbb{Q}}$ with a Weierstrass equation $y^2 = 4x^3 - g_2x - g_3$ ($g_2, g_3 \in \bar{\mathbb{Q}}$), having the fundamental periods ω_1, ω_2 (with $\mathrm{Im}\frac{\omega_2}{\omega_1} > 0$) and the corresponding quasi-periods η_1, η_2. The only relation between ω_i and η_j that always holds is the Legendre identity:

$$\eta_1\omega_2 - \eta_2\omega_1 = 2\pi i.$$

Thus π belongs to the field generated by periods and quasiperiods over $\bar{\mathbb{Q}}$. In the complex multiplication case $\tau = \frac{\omega_2}{\omega_1}$ is an imaginary quadratic number.

Whenever $\tau \in \mathbb{Q}(\sqrt{-d})$ for $d > 0$, invariants g_2 and g_3 can be chosen from the Hilbert class field of $\mathbb{Q}(\sqrt{-d})$, and this field is the minimal extension with this property.

A priori complex multiplication means only a single relation between ω

It seems that until Ramanujan's paper nobody explicitly stated the existence of the second relation between periods and quasiperiods. This relation is the following one:

Whenever τ is a quadratic number: $A\tau^2 + B\tau + C = 0$, the four numbers: $\omega_1, \omega_2, \eta_1, \eta_2$ are linearly dependent over $\bar{\mathbb{Q}}$ only on two of them.

Explicitly:

$$\begin{aligned} &\omega_2 = \tau\omega_1, \\ &A\tau\eta_2 - C\eta_1 + \alpha\omega_1 = 0 \end{aligned} \tag{1.1}$$

for $\alpha \in \bar{\mathbb{Q}}$ ($\alpha \in \mathbb{Q}(\tau, g_2, g_3)$).

The relations (1.1) are not entirely original; Legendre's investigation of the lemmiscate case provides with (1.1) in two cases, where τ is equivalent to i or to ρ under $SL_2(\mathbb{Z})$; moreover, these cases were clearly known to Euler, who evaluated the appropriate complete elliptic integrals. However, those two particular cases are "wrong ones": in these cases the importance of the relation (1.1) is lost because $\alpha = 0$, and it is hard to understand the need for its appearance. In a few other special singular moduli cases, (1.1) appears in the classical literature, see [4].

Of course, Ramanujan did not use the Weierstrass equations and preferred the Legendre ones, where one can see the hypergeometric functions instantly.

In order to pass to Legendre notations [5], one puts function

$$k^2 = \frac{e_2 - e_3}{e_1 - e_3}$$

(an automorphic function $k^2 = k^2(\tau)$ with respect to the principal congruence subgroup $\Gamma(2)$ of $\Gamma(1) = SL_2(\mathbb{Z})$ in the variable τ in $H = \{\tilde{z}: \operatorname{Im} \tilde{z} > 0\}$). Then the periods and quasiperiods are expressed through the complete elliptic integrals of the first and second kind:

$$K(k) = \frac{\pi}{2}\cdot F(\tfrac{1}{2},\tfrac{1}{2};1;k^2),$$

$$E(k) = \frac{\pi}{2}\cdot F(-\tfrac{1}{2},\tfrac{1}{2};1;k^2).$$

We also look at $K(k')$, $E(k')$ for $k'^2 = 1 - k^2$; then ω_1, η_1 are expressed in terms of $K = K(k)$, $E = E(k)$, while ω_2, η_2 are expressed in terms of iK' and iE':

$$\omega_1 = \frac{K}{\sqrt{e_1-e_3}}, \quad \omega_2 = \frac{iK'}{\sqrt{e_1-e_3}};$$

$$\eta_1 = \sqrt{e_1-e_3}\cdot E - e_1\omega_1, \quad \eta_2 = -\sqrt{e_1-e_3}\; iE' - e_3\omega_2.$$

Invariant α in (1.1)--a nontrivial part of the Ramanujan quasiperiod relation--is easily recognized as one of the simplest values of "nonholomorphic Eisenstein series". Weil's treatise [6] creates a clear impression that this quantity and its algebraicity had been known to Kronecker (or even Eisenstein). It seems to us that though similar and more general functions were carefully examined, this particular connection had been reconstructed by Weil, and cannot be easily separated

from his own work on period relations. The "nonholomorphic" Eisenstein series are too important to be ignored.

The usual Eisenstein series associated with the lattice $\mathcal{L}$ of periods: $\mathcal{L} = \mathbb{Z}\omega_1 \oplus \mathbb{Z}\omega_2$ are

$$G_k(\omega_1,\omega_2) = \sum_{\substack{\omega\in\mathcal{L}\\ \omega\neq 0}} \omega^{-k} \quad \text{for } k = 4,6,\ldots .$$

The corresponding normalized inhomogeneous series $E_k(\tau)$ are defined as

$$G_k(\omega_1,\omega_2) = \left(\frac{2\pi i}{\omega_2}\right)^k\cdot\frac{-B_k}{k!}\cdot E_k(\tau),$$

or

$$E_k(\tau) = 1 - \frac{2k}{B_k}\cdot\sum_{n=1}^{\infty} \sigma_{k-1}(n)\cdot q^n$$

for $\sigma_{k-1}(n) = \Sigma_{d|n} d^{k-1}$, and $q = e^{2\pi i\tau}$.

These q-series were subject of a variety of Ramanujan's studies [1-2] with his preferred notations a P,Q,R for $E_k(\tau)$ with $k = 2,4,6$, respectively.

For $k = 2$ the proper definition of $G_k(\omega_1,\omega_2)$ is a nonholomorphic one arising from

$$H(s,z) = \sum_{\omega\in\mathcal{L}} (\bar{z}+\bar{\omega})\,|z+\omega|^{-2s}$$

as:

$$G_2(\omega_1,\omega_2) = \lim_{s\to\infty^+} \sum_{\omega\in\mathcal{L}}{}^{*} \omega^{-2}\cdot|\omega|^{-2s}.$$

In the $E_k(\tau)$ notations, the quasiperiod relation is expressed by means of the function

$$s_2(\tau) \overset{\text{def}}{=} \frac{E_4(\tau)}{E_6(\tau)}\cdot\left(E_2(\tau) - \frac{3}{\pi \operatorname{Im}(\tau)}\right), \tag{1.2}$$

which is invariant under the action of $\Gamma(1)$ but nonholomorphic. It is this object that was studied by Ramanujan in connection with α in (1.1). Ramanujan actually proved in [1] that this function admits algebraic values whenever τ is imaginary quadratic. Moreover, Ramanujan [1] presented a variety of algebraic expressions for this function, differentiating modular equations.

His work, or Weil's, shows that the function in (1.2) has values from the Hilbert class field $\mathbb{Q}(\tau,j(\tau))$ of $\mathbb{Q}(\tau)$ for quadratic τ. The relation of (1.2) to α is simple: for $\beta = s_2(\tau)$ from (1.2), $\alpha = (B + 2A\tau)\beta\cdot g_3/g_2$.

Amazingly, Mordell in his notes [1] on Ramanujan paper missed the true importance of (1.1) or (1.2), stating merely "... Ramanujan's method of obtaining purely algebraical approximations appears to be new." These relations were rediscovered in the 70's (among the rediscovers was Siegel [7]), see particularly [9], and stimulated search for multidimensional generalizations of the period relations promoted by Weil [8]. We'll talk about generalizations of elliptic period relations later, but meanwhile let us look on relations (1.1) once more. One can combine (1.1) with the Legendre relation to arrive to a phenomenally looking "quadratic relation" derived by Ramanujan, that expresses π in terms of squares of ω_1 and η_1 only (no ω_2 and η_2!) All this is interesting, as an algebraic identity, but Ramanujan transforms these quadratic relations into rapidly convergent generalized hypergeometric representation of simple algebraic multiples of $1/\pi$. To do this he needed not only modular functions but also hypergeometric function identities.

We reproduce first Ramanujan's own favorite [1], which was used by Gosper in 1985 in his $17.5\cdot 10^6$ decimal digit computation of π:

$$\frac{9801}{2\sqrt{2}\pi} = \sum_{n=0}^{\infty} \{1103 + 26390n\} \frac{(4n)!}{n!^4\cdot(4\cdot 99)^{4n}} \qquad (44).$$

(Numeration here is temporarily borrowed from [1].)

The reason for this pretty representation of $1/\pi$ lies in the representation of $(K(k)/\pi)^2$ as a ${}_3F_2$-hypergeometric function. Apparently there are four classes of such representations all of which were determined by Ramanujan: these are four distinct classes of ${}_3F_2$-representation of $1/\pi$, all based on special cases of Clausen identity (and all presented by Ramanujan [1]):

$$F(a,b;a+b+\tfrac{1}{2};z)^2 = {}_3F_2\left(\begin{matrix}2a,a+b,2b\\ a+b+\frac{1}{2},2a+2b\end{matrix}\Big|\, z\right). \qquad (1.3)$$

Unfortunately, the Clausen identity is a unique one--no other nontrivial relation between parameters makes a product of hypergeometric functions a generalized hypergeometric function.

We display the basis for the Ramanujan's series representation for $1/\pi$. We'll discuss them later in connection with arithmetic triangle groups. Meanwhile, what is good in these identities for diophantine approximations? First of all, Ramanujan approximations to π are indeed remarkably fast numerical schemes of evaluation of π. Unlike some other numerical schemes these are series schemes, that can be accelerated into Padé approximation schemes. These Padé approximations schemes are better numerically, but more important, they are nontrivial arithmetically good rational approximations to algebraic multiples of π, that provide nontrivial measures of diophantine approximations.

(Deviating momentarily, we want to compare Padé approximations vs. power series approximations in numerical evaluation of functions and constants. Remarkably, asymptotically there is no significant difference between Boolean complexities of evaluation of Padé approximations to solutions of linear differential equations and of power series approximations within a given precision. Unfortunately, asymptotically there is no gain in the degree of approximations either; moreover there is a significant difference in storage requirements. Padé approximations require more storage. Even in cases when explicit Padé approximations are known, gains of using them can be visible only in about hundreds of digits of precision; not below or above. That is why unless special circumstances call for (like uniform approximations with a minimal storage in ordinary precision range), Padé approximations and continued fraction expansion techniques should not be used for numerical evaluation.

However, in diophantine approximations we have no choice. Only Padé approximations and a vast army of their generaliza-

tions are capable of approximating functions and constants and tell something of their arithmetic nature, of their irrationalities and transcendences, measures of approximation, etc.).

All Ramanujan's quadratic period relations (four types) can be deduced from one series by modular transformations, and we choose the series as the one associated with the modular invariant $J = J(\tau)$. In the place of Ramanujan's nonholomorphic function we take, as above in (1.2):

$$s_2(\tau) = \frac{E_4(\tau)}{E_6(\tau)}\left(E_2(\tau) - \frac{3}{\pi \operatorname{Im}(\tau)}\right).$$

Then the Clausen identity gives the following ${}_3F_2$-representation for an algebraic multiple of $1/\pi$:

$$\sum_{n=0}^{\infty} \left\{\frac{1}{6}(1-s_2(\tau))+n\right\}\cdot\frac{(6n)!}{(3n)!n!^3}\cdot\frac{1}{J(\tau)^n} = \frac{(-J(\tau))^{1/2}}{\pi}\cdot\frac{1}{(d(1728-J(\tau))^{1/2}}. \tag{1.4}$$

Here if $\tau = (1+\sqrt{-d})/2$. If $h(-d) = 1$ the second factor in the right hand side is a rational number. The largest one class discriminat $-d = -163$ gives the most rapidly convergent series (though coefficients are slightly strange):

$$\sum_{n=0}^{\infty} \{c_1+n\}\cdot\frac{(6n)!}{(3n)!n!^3}\,\frac{(-1)^n}{(640{,}320)^n} = \frac{(640{,}320)^{3/2}}{163\cdot 8\cdot 27\cdot 7\cdot 11\cdot 19\cdot 127}\cdot\frac{1}{\pi}. \tag{1.5}$$

Here

$$c_1 = \frac{13{,}591{,}409}{163\cdot 2\cdot 9\cdot 7\cdot 11\cdot 19\cdot 127}$$

(and, of course, $J(\frac{1+\sqrt{-163}}{2}) = -(640{,}320)^3$).

Ramanujan [1] provides instead of this a variety of other formulas connected mainly with the three other triangle groups commensurable with $\Gamma(1)$.

Four classes of ${}_3F_2$ representations of algebraic multiples of $1/\pi$ correspond to four ${}_3F_2$ hypergeometric functions (that are squares of ${}_2F_1$ representations of complete elliptic integrals via the Clausen identity). These are

$${}_3F_2\left({1/2\ ,\ 1/6\ ,\ 5/6 \atop 1\ ,\ 1}\Big| x\right) = \sum_{n=0}^{\infty} \frac{(6n)!}{(3n)!\,n!^3}\left(\frac{x}{12^3}\right)^n \tag{1.6}$$

$${}_3F_2\left({1/4\ ,\ 3/4\ ,\ 1/2 \atop 1\ ,\ 1}\Big| x\right) = \sum_{n=0}^{\infty} \frac{(4n)!}{n!^4}\left(\frac{x}{4^4}\right)^n \tag{1.7}$$

$${}_3F_2\left({1/2\ ,\ 1/2\ ,\ 1/2 \atop 1\ ,\ 1}\Big| x\right) = \sum_{n=0}^{\infty} \frac{(2n)!^3}{n!^6}\left(\frac{x}{2^6}\right)^n \tag{1.8}$$

$${}_3F_2\left({1/3\ ,\ 2/3\ ,\ 1/2 \atop 1\ ,\ 1}\Big| x\right) = \sum_{n=0}^{\infty} \frac{(3n)!}{n!^3}\cdot\frac{(2n)!}{n!^2}\left(\frac{x}{3^3\cdot 2^2}\right)^n. \tag{1.9}$$

The first function is the one arising in (1.4) with $x = 12^3/J(\tau)$. Other three functions correspond to modular transformations of $J(\tau)$. This means that appropriate $x = x(\tau)$ is a modular function of higher level (e.g. in (1.8), $x = 4k^2(1-k^2)$ for $k^2 = k^2(\tau)$), and series (1.7)-(1.9) for the same τ have slower convergence rates than the series in (1.6).

Representations similar to (1.5) can be derived for any of the series (1.6)-(1.9) for any singular moduli $\tau \in \mathbb{Q}(\sqrt{-d})$

and for any class number $h(-d)$, thus extending Ramanujan list [1] ad infinum. There is a simple recipe to generate these new identities, instead of elaborate procedure proposed in [1] (see also [5]) based on differentiating of modular equations. To derive these identities one needs the explicit expressions of $x_j = x(\tau_j)$ with the representatives τ_j in H of algebraically conjugate values of automorphic function $x = x(\tau)$ for $\tau \in \mathbb{Q}(\sqrt{-d})$ (say $\tau = \sqrt{-d}$ or $\tau = \frac{1+\sqrt{-d}}{2}$). E.g. for $x(\tau) = 12^3/J(\tau)$, τ_j: $j = 1,\ldots,h(-d)$ corresponds to the class number of $\mathbb{Q}(\sqrt{-d})$. The necessary values of $s_2(\tau_j)$ are easy to compute from q-series representation of $E_k(\tau)$, if to use the formula (1.2). These computations can be carried out in bounded precision, because, as we know, $s_2(\tau_j)$ lies in the Hilbert class field of $\mathbb{Q}(\sqrt{-d})$ and because, whenever $J(\tau_j)$ is algebraically conjugate to $J(\tau_i)$, numbers $s_2(\tau_j)$ and $s_2(\tau_i)$ are also algebraically conjugate. This allows us to express $s_2(\tau)$ in terms of $J(\tau)$ and $\sqrt{-d}$ explicitly using only finite precision approximations to all $s_2(\tau_j)$. This way one obtains rapidly convergent ${}_3F_2$ representations of algebraic multiples of π by nearly integral power series. When $h(-d) > 1$, these series contain nonrational numbers, making the series (1.5) the fastest convergent series representing a multiple of $1/\pi$, and having rational number entries only.

Even before Ramanujan's remarkable approximations to π, singular moduli evaluations were used to approximate multiples of π by logarithms of algebraic numbers (usually the values of modular invariants). One of the first series of such approximations belongs to Hermite [13]. Of course, by now it is reproduced in hundreds of papers and we have to give a customary example. One is looking here at the expansion of the modular invariant near infinity:

$$J(\tau) = q^{-1} + 744 + 196884q + 21493760q^2 + \ldots$$

for $q = e^{2\pi i\tau}$. The elementary theory of complex multiplication shows that for $\tau = (1+\sqrt{-d})/2$, $q^{-1} = e^{-\pi\sqrt{d}}$ is very close to an algebraic integer $J(\tau)$ of degree $h(-d)$. Usual examples (see description in [14]) involve the largest one class discriminant -163, $d = 163$, when:

$$e^{\pi\sqrt{163}} = -262537412640768743.9999999999992\ldots\ .$$

There is a variety of these and similar approximations of π by logarithms of other classical automorphic functions. One of the most popular, studied by Shanks et. al. [15], has a simple q-expansion:

For $f = f_1(\sqrt{-d})^{-24} = (k/4k'^2)^2$ at $\tau = \sqrt{-d}$,

$$\log f + \sqrt{d}\cdot\pi = 24\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}q^k(1-q^k)^{-1}.$$

It is also known that the right side can be expanded in powers of f:

$$\log f + \sqrt{d}\cdot\pi = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{a_n}{n}\cdot f^n.$$

Here a_n are integers. Shanks [15] looks at specialization of this formula for $-d$ with class groups of special structure for relatively large d. These approximations are not technically approximations to π, but rather to a linear form in π and in another logarithm. All of them are natural consequences of Schwarz theory and the representation of the function inverse to the automorphic one (say $J(\tau)$) as a ratio of two solutions of a hypergeometric equation. One such formula is

$$\pi i\cdot\tau = \ln(k^2) - \ln(16) + \frac{G(\frac{1}{2},\frac{1}{2};1;k^2)}{F(\frac{1}{2},\frac{1}{2};1;k^2)}, \qquad (1.10)$$

and another (our favorite) is Fricke's [16]

$$2\pi i\cdot\tau = \ln J + \frac{G(\frac{1}{12},\frac{5}{12};1;\frac{12^3}{J})}{F(\frac{1}{12},\frac{5}{12};1;\frac{12^3}{J})}. \qquad (1.11)$$

Here $G(a,b;c;x) = \Sigma_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n n!}\cdot\{\Sigma_{j=0}^{n-1}(\frac{1}{a+j} + \frac{1}{b+j} - \frac{2}{c+j})\}$ is the hypergeometric function (of the second kind) in the exceptional case, when there are logarithmic terms.

Perhaps the most popular approximations to linear forms in π and in another logarithm are associated with Stark-Baker solution to one-class and two-class problems (cf. [17]). Stark's approach [18] is based on Kronecker's limit formula, and in a way, similar to approximations given above, one represents

$$L(1,\chi)\cdot L(1,\chi\chi') - \frac{\pi^2}{6}\prod_{p|k}(1-\frac{1}{p^2})\sum_f\frac{\chi(a)}{a}$$

for $\chi(\cdot) = (\frac{k}{\cdot})$, $\chi'(\cdot) = (\frac{-d}{\cdot})$, $k > 0$ as a rapidly convergent

$q^{1/k}$-series (here $f = ax^2 + bxy + cy^2$ runs through a complete set of unequivalent quadratic forms with the discriminant $-d$). Using Dirichlet's class number formula one obtains an exceptional approximation (as above) to the linear form in two logarithms:

$$|h(-kd)\cdot \log \varepsilon_{\sqrt{k}} - q\pi\sqrt{d}| < e^{-\pi O(\sqrt{d}/k)}$$

for an arbitrary fundamental unit $\varepsilon_{\sqrt{k}}$ of $\mathbb{Q}(\sqrt{k})$ and for one-class discriminant $-d$. These remarkable linear forms were generalized by Stark to three-term linear forms in the class number two case.

(While these unusually good approximations can be used in the class number problem--approximations are so good that they are impossible for large d--these linear forms cannot be used for the analysis of arithmetic properties of the individual logarithms, like π, entering the linear form. Moreover, as the class number grows, the number of the terms in the linear form grows.)

Important developments initiated by Ramanujan in his truly algebraic approximations to $1/\pi$ can be extended to the analysis of linear forms in logarithms presented above. In fact, each term in these linear forms can be separately represented by a rapidly convergent series in $1/J$ with nearly integral coefficients.

For this one takes an automorphic function $\varphi(\tau)$ with respect to one of the congruence subgroups of $\Gamma(1)$ and expand

functions like $F(\frac{1}{12},\frac{5}{12};1;12^3/J)$, $G(\frac{1}{12},\frac{5}{12};1,;12^3/J)$ in powers of $\varphi(\tau)$ instead of $J(\tau)$. Whenever $\varphi(\tau)$ is automorphic with respect to a classical triangle group, we arrive to the corresponding usual hypergeometric functions.

Other logarithms, like π, can be represented as values of convergent series satisfying Fuchsian linear differential equations. This is particularly clear for $\log \varepsilon_{\sqrt{k}}$ of a fundamental unit $\varepsilon_{\sqrt{k}}$ of a real quadratic field $\mathbb{Q}(\sqrt{k})$. To represent this number as a convergent series (in, say, $1/J(\tau)$) one uses Kronecker's limit formula expressing this logarithm $\log \varepsilon_{\sqrt{k}}$ in terms of products of values of Dedekind's Δ-function ("Jugandtraum", see [6]). Such an expression of $\log \varepsilon_{\sqrt{k}}$ in terms of power series in $1/J(\tau)$ for $\tau = (1+\sqrt{-d})/2$ for any $d \equiv 3(8)$, depends, unfortunately, on k, because k is related to the level of the appropriate modular form $\varphi = \varphi_k(\tau)$.

For $k = 5$ Siegel [19] made an explicit computation that expresses $J(\tau)$ in terms of the resolvent $\varphi_5(\tau)$ of 5-th degree modular equation known from the classical theory of 5-th degree equations. His relations [19] were:

$$(\varphi-\varepsilon^3)((\varphi-\varepsilon)(\varphi^2+\varepsilon^{-1}\varphi+\varepsilon^{-2}))^3 + (\varphi/\sqrt{5})^5 J = 0$$

and

$$\varphi(\tau)(= \varphi_5(\tau)) = \varepsilon^{h(-5p)/2}$$

for $\tau = (1+\sqrt{-p})/2$ and $\varepsilon = \varepsilon_{\sqrt{5}}$. Here one has $p \equiv 3(5)$; if $p \equiv 2(5)$ and replaces ε by ε^{-1} in the expression of $J = J(\tau)$.

This, in combination with Ramanujan's approximation to π, allows one to express log ϵ (its multiple) as a convergent series in $1/J$ (or in $1/\wp$).

2. Period relations for Abelian varieties with complex multiplication.

The most general approach to period relations emerged in works of Shimura [20-21], Deligne [22-23] and others after extensive work in the 60's and 70's by Shimura, mainly on algebraic values of automorphic forms of $\mathbb{Q}$-rational reductive algebraic groups G (see references in [24]). In this work, with many important special examples treated by Siegel and Weil (see [8]), period relations are constructed for CM-points in bounded symmetric domains, and then values of automorphic forms at CM-points are expressed in terms of these periods. Among values of such automorphic forms are values of L-functions corresponding to Grossencharacters at integral points and values of various Eisenstein series. To describe period relations in invariant form we use the language of CM-types of Shimura-Taniyama [25].

A totally imaginary quadratic extension K of a totally real algebraic number field is called a CM-field. We look at $\mathbb{Z}$-module generated by formal sums of σ_i for embeddings σ_i of K into $\mathbb{C}$. An Abelian variety A defined over $\bar{\mathbb{Q}}$ is said to have a CM-type (K,Φ) if

$$\dim(A) = g, \quad [K:\mathbb{Q}] = 2g, \text{ there exist an injection } i: K \to \mathrm{End}(A) \otimes \mathbb{Q} \tag{2.1}$$

(i.e. A is a CM-variety), and

$$\Phi = \sum_{i=1}^{n} \sigma_i, \tag{2.2}$$

where σ_i runs through a set of representatives of pairs of complex conjugate embeddings of K into $\mathbb{C}$, and for $j = 1,..$ $j = 1,\ldots,n$ there exists a $\bar{\mathbb{Q}}$-rational holomorphic 1-form $\omega_j \neq 0$ on A such that

$$\omega_j \circ i(a) = a^{\sigma_j} \omega_j \text{ for all } a \in K$$

(with $i(a) \in \mathrm{End}(A)$).

Shimura introduces then "CM-periods" $p(\sigma_j,\Phi)$ depending on K, Φ, σ_j such that

$$\frac{\int_\gamma \omega_j}{\pi \cdot p(\sigma_j,\Phi)} \in \bar{\mathbb{Q}}$$

for all 1-cycles γ on A ($\gamma \in H_1(A,\mathbb{Z})$).

Shimura and Deligne [20-23] have shown how all periods and quasiperiods of a CM-variety A can be expressed algebraically through $p(\sigma,\Phi)$. For example, if σ is an embedding of K into $\mathbb{C}$, and $\Phi_1,\ldots,\Phi_m$ are CM-types of K, then $\prod_{i=1}^m p(\sigma,\Phi_i)^{c_i}$ for $c_i \in \mathbb{Z}$ depends, up to an algebraic (i.e. $\bar{\mathbb{Q}}$) factor, only on $\Sigma_{i=1}^m c_i\Phi_i$ and σ.

One of the results that determines the number of algebraically independent generators of period and quasiperiod matrix of A (the dimension of the Hodge group of A) is the following one [21,23]:

Let I_K^0 be the module of all formal linear combinations $\Sigma_\sigma c_\sigma \cdot \sigma$ for $c_\tau \in \mathbb{Q}$ such that $c_\sigma + c_{\sigma\rho}$ does not depend on σ (ρ is the complex conjugation). The rank, $r(K,\Phi)$ of a CM-type (K,Φ) is the dimension of the subspace of I_K^0 generated over $\mathbb{Q}$ by all Φ_γ for $\gamma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$.

Then for a CM-variety A with the CM-type (K,Φ), the number of algebraically independent elements among periods $\int_\gamma \omega$ and quasiperiods $\int_\gamma \eta$ is bounded by the rank $r(K,\Phi)$.

Here ω and η are, respectively, the differentials of the first and second kind on A, and γ is a 1-cycle.

We conjectured [26] that, in fact, the number of algebraically independent numbers among periods and quasiperiods is exactly the rank $r(K,\Phi)$. This conjecture is correct at least for $g = 1$. In the case $g = 1$ the rank is always 2, there are, in fact, two algebraically independent numbers among CM periods and quasiperiods; the fact that the rank is 2 can be considered to be a consequence of Ramanujan's quasiperiod relation.

The rank is bounded as follows:

$$2 + \log_2 g \leq r(K,\Phi) \leq g + 1$$

(with both inequalities achievable).

These period relations and the framework within which they are achieved: values of Eisenstein series and the theory of motives, are the most far reaching generalizations of "quadradic period relations" discovered by Ramanujan in the elliptic curve case. Additional information, also often sought by Ramanujan in the elliptic curve case, include the determination of period relations up to $\mathbb{Q}$-factors (i.e. resolution of algebraic equations in radicals), and the determination of periods in terms of products of Γ-functions in Abelian cases [22]. The Γ-product representations for periods are known in the case of elliptic curves with complex multiplication as Selberg-Chowla formulas [6]. Ramanujan [1] observed particular cases of these formulas. Important additions to Selberg-Chowla formulas include computation of algebraic factors in terms of values of p-adic Γ-functions, due to Gross and Deligne, see [27].

Most of this recently developed fascinating theory of period relations is of primary interest to algebraic geometry and escapes an easy translation into magic identities of Ramanujan type. It is therefore reasonable to restrict ourselves to the cases which Ramanujan might have himself stumbled upon by looking only on hypergeometric equations.

Instead of arithmetic quotients of general bounded symmetric domains we will look only at discrete arithmetic subgroups acting on the upper half-plane. These arithmetic subgroups were well studied by the late $\overline{\text{XIX}}$ century, particularly by Fricke and Klein. If to restrict ourselves to automorphic forms expressed in terms of hypergeometric integrals, we arrive to the celebrated class of Schwarz triangle groups denoted according to canonical notations as

$$(0.3;\ell_1,\ell_2,\ell_3) \tag{2.3}$$

where $2 \leq \ell_i \leq \infty$.

For the triangle group with the signature (2.3) to be Fuchsian, the condition of non- Euclideanity has to be satisfied:

$$\frac{1}{\ell_1} + \frac{1}{\ell_2} + \frac{1}{\ell_3} < 1. \tag{2.4}$$

(If (2.4) is not satisfied we arrive at the famous Schwarz's cases of finite polyhedral groups, if the inequality in (2.4) is reversed, or to an elliptic case uniformizing the curves of genus 0 and 1, if the equality is substituted in (2.4).)

The triangle groups with the signature (2.3) are explicitly expressed in terms of hypergeometric functions as follows. For $\lambda = 1/\ell_1$, $\mu = 1/\ell_2$, $\nu = 1/\ell_3$ we put $a = \frac{1}{2}(1-\lambda-\mu+\nu)$, $b = \frac{1}{2}(1-\lambda-\mu-\nu)$, $c = 1-\lambda$. Then the function

$$\tau = S(\lambda,\mu,\nu;z),$$

inverse to the function $\varphi(\tau) = z$, automorphic in H with respect to the triangle group $(0,3;\ell_1,\ell_2,\ell_3)$, is expressed as the ratio of two linearly independent solutions of the hypergeometric equation with the parameters a,b,c defined above. E.g. for $\lambda(= 1/\ell_1) \neq 0$, one can put

$$\tau = C\cdot z^{1/\ell_1}\cdot\frac{F(a+1-c,b+1-c;2-c;z)}{F(a,b;c;z)}.$$

Not all triangle groups are arithmetic; one easily recognizes those arithmetic triangle groups that correspond to (are commensurable with) a full modular group $\Gamma(1)$ or its principal congruence subgroup $\Gamma(2)$. However, the list of arithmetic triangle groups is rather large. The first sequence of arithmetic triangle subgroups had been described by Fricke and Klein in [28] in connection with ternary quadratic forms. Later in the turn of the $\overline{\underline{\text{XX}}}$ century an extensive investigation of triangle groups had been conducted by Hutchinson, Young and Morris, following earlier work of Hurwitz and Burkhardt. In these works detailed investigations of special classes of arithmetic triangle groups were carried out, see references R. Morris, [29]. This American contribution was much more than just automorphic function study, because period relations on Riemann surfaces of high genera with four critical points were described in detail; unfortunately this investigation seems to be largely ignored.

Recently a complete classification of all arithmetic triangle groups was presented by Takenchi [30]. For a precise definition of an arithmetic subgroup see Swinnerton-Dyer [31]. (Roughly speaking an arithmetic triangle group is one which is commensurable with the Fuchsian group arising from the order of a quaternion algebra over a totally real field.)

Computations of Takenchi [30] confirmed the findings of Fricke-Klein [28] (essentially, up to commensurability, the arithmetic triangle subgroups were already known.)

If one looks at the commensurability classes of arithmetic triangle subgroups, one of the classes corresponds to the

full modular group with its most important representative being

$$\begin{aligned}&(0,3;2,3,\infty),\\&(0,3;2,4,\infty),\\&(0,3;2,6,\infty),\\&(0,3;2,\infty,\infty).\end{aligned}\tag{2.5}$$

These are exactly four cases of hypergeometric functions studied by Ramanujan [1] in connection with period relations. Here one always has $\ell_1 = 2$ to have the Clausen identity. The squares of hypergeometric function corresponding to four classes in (2.5) are listed explicitly in (1.6)-(1.9) as ${}_3F_2$-hypergeometric functions. The four triangle subgroups listed in (2.5), two of which correspond simply to $\Gamma(1)$ and $\Gamma(2)$ and two others to Hecke groups G_q with $q = 4,6$ are well studied in classical and modern literature. The commensurability relations between groups can be easily translated into algebraic transformations of the corresponding hypergeometric functions, i.e. into modular relations of low degrees. Nothing more than the usual elliptic theory of modular functions occurs here, and one can always deduce all the identities from a single group, corresponding, say, to $(0,3;2,3,\infty)$--the case of the modular invariant $J(\tau)$--our favorite.

Other arithmetic triangle subgroups are less redundant. Among those of special interest to us are 17 classes [30] of commensurability of triangle subgroups corresponding to 12 totally real fields with class number 1 and quaternion algebras over them. In each of these cases there is a rich theory of arithmetic values of functions automorphic with respect to arithmetic triangle groups Γ acting on H with a compact H/Γ.

One of these arithmetic subgroups is a particularly distinguished Hurwitz group

$$(0,3;2,3,7).$$

This is a group with the minimal volume of H/Γ (among all Fuchsian groups of the first kind). Factors of this Hurwitz group are known to attain the maximal order $84(g-1)$ of automorphism group of a Riemann surface of genus $g > 1$.

The picture of the action of this triangle subgroup often illustrates papers, books and conference posters (see Figure 1).

Remarkably, very little is known about the arithmetic properties of values of automorphic functions corresponding to Hurwitz's and other arithmetic triangle compact groups, though Shimura in his papers (see review in [24]) built a complete theory of complex multiplication in most of these cases. He determined the Hilbert class fields and the action of Frobenius there (including the case of Hurwitz group).

A single commensurability class of triangle subgroups corresponds to quaternion algebras over $\mathbb{Q}$ with discriminant $D = 2\cdot 3$. This class contains

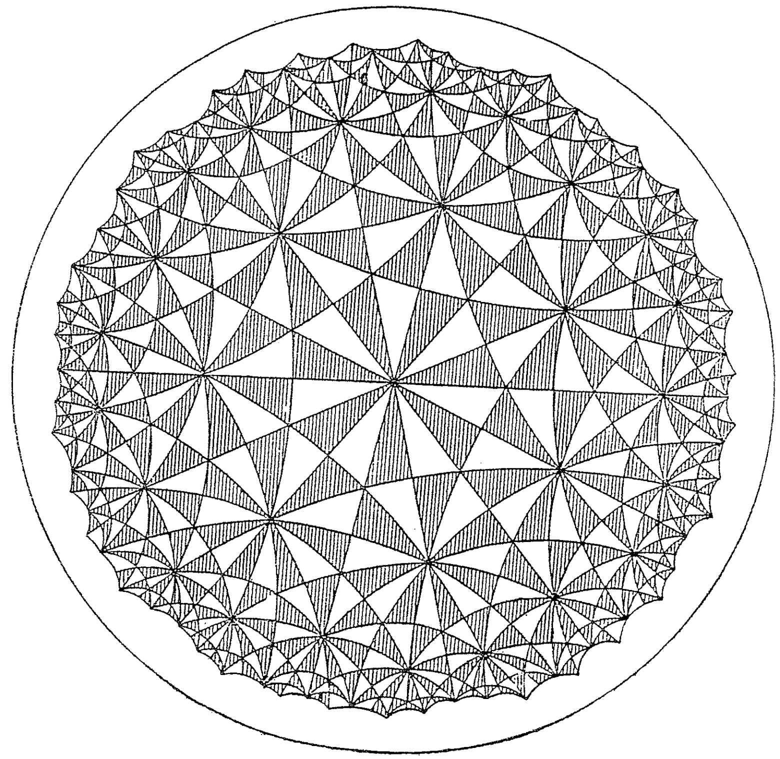

Figure 1

$$(0,3;2,4,6) \quad \text{and} \quad (0,3;2,6,6).$$

These cases are quite complex from the point of view of the underlying hypergeometric integrals and were partially studied by Hutchinson and Dalaker, see [32]. (These two cases are two cases of the general curve with four critical points (hypergeometric curves) for which the moduli of all integrals of the first kind, i.e. all periods of the differential of the first kind on this curve, are expressed linearly in terms of a single parameter. The problem of determination of all hypergeometric curves possessing this property had been studied by Morris [29], who proved that in addition to 4 cases commensurable with the full modular group there are only two additional cases: (0,3;2,4,6) and (0,3;2,6,6). These two cases lead to curves of genera 23 and 6, respectively (see below). This seems slightly surprising because the monodromy groups of the corresponding differential equations are rather small as subgroups of $SL_2(\mathcal{O})$ for a ring $\mathcal{O}$ of integers in $\mathbb{Q}(\sqrt{2},\sqrt{3})$. We are dealing here, however, with Riemann surfaces depending on a single parameter that admit a large group of automorphisms some of which are nontrivial. In a sense these particular Riemann surfaces represent a one-parameter generalization of the famous Klein curves

$$x^3y + y^3z + z^3x = 0$$

--a famous genus 3 curve with 168 automorphism group, and its less famous quintic counterpart

$$x^4y + y^4z + z^4x = 0$$

studied by Snyder and Scorza.

To see why it is unusual for a hypergeometric curve to have its period matrix to depend linearly only on a single parameter, let us look at the general expression of a curve

with 4 critical points:

$$y^q = (x-a_0)^{\alpha_0}(x-a_1)^{\alpha_1}(x-a_2)^{\alpha_2}(x-a_3)^{\alpha_3},$$

(when $a_0 = \infty$, one puts $\alpha_0 = 0$). Elements of Riemann's period matrix of this curve have the form of hypergeometric integrals

$$\int (x-a_1)^{\beta_1}(x-a_2)^{\beta_2}(x-a_3)^{\beta_3}dx,$$

for a variety of exponents $(\beta_1,\beta_2,\beta_3)$. Not all of these integrals can be linearly dependent; and these linear dependences are usually nontrivial automorphisms).

We look now at Jacobians of hypergeometric curves (curves with four critical points), i.e. at all periods of differentials of the first kind, and not only obvious ones, expressed by hypergeometric functions themselves.

Only the simplest case of $(0;3;\infty,\infty,\infty)$--i.e. of $K(k) = \frac{\pi}{2}F(\frac{1}{2},\frac{1}{2};1;k^2)$--clearly leads to elliptic function field. Already other triangle subgroups, commensurable with this one, give rise to nontrivial Riemann surfaces of hypergeometric type. E.g. the two Hecke triangle groups in (2.5) lead to Riemann surfaces of genra 2. These are, respectively,

$$y^3 = (x-a_0)(x-a_1)(x-a_2)^2(x-a_3)^2$$

for $(0,3;2,6,\infty)$, and

$$y^4 = (x-a_0)(x-a_1)^2(x-a_2)^2(x-a_3)^3$$

for $(0,3;2,4,\infty)$.

In both cases $g = 2$, and the Jacobian of each of the curves is isogenous to the product of two elliptic curves, which are isogenous to each other. Explicit modular equations express integrals of the first kind on these curves (and not only their periods) through classical elliptic integrals.

For two other triangle subgroups associated with quaternion algebra over $\mathbb{Q}$ with the discriminant $D = 2\cdot 3$, arising hypergeometric curves are even more complicated. For the triangle group $(0,3;2,6,6)$ we get a genus $g = 3$ equation

$$y^6 = (x-a_0)^2(x-a_1)^2(x-a_2)^3(x-a_3)^5, \tag{2.6}$$

and for $(0,3,2,4,6)$ we arrive at $g = 23$ curve

$$y^{24} = (x-a_0)(x-a_1)^{11}(x-a_2)^{17}(x-a_3)^{19}. \tag{2.7}$$

The first curve (2.6) has its Jacobian isogeneous to the product of an elliptic curve and a Jacobian of a hyperelliptic curve. The second curve is harder to describe. Its Jacobian possesses many nontrivial Jacobian factors (of genera 2 there are at least eight, and there is a factor of genera 5).

(All other (arithmetic) triangle subgroups lead to hypergeometric curves of high genera ($g > 2$), which have period matrices that cannot be expressed linearly in terms of a single parameter [29], [32]. This means that monodromy groups of more than a single hypergeometric equation are involved in the description of the monodromy group of a single hypergeometric curve. This means that the totality of numbers in the

period and quasiperiod matrix of the corresponding hypergeometric Riemann surfaces involves more than values of a single automorphic function. This makes the corresponding theory of complex multiplication very complicated.)

The cases of $(0,3;2,4,6)$ and $(0,3;2,6,6)$ (and other arithmetic triangle groups) immediately lend themselves to the generalization of the Ramanujan period relations. For each of these cases we can look at the function $z = \varphi(\tau)$ automorphic in H with respect to the corresponding arithmetic triangle group see [24]. We normalize this function like the classical modular invariant by its values in the vertices of the fundamental triangle. (Say we put, following [24]:

$$\varphi(e_2) = 1, \ \varphi(e_m) = 0, \ \varphi(e_n) = \infty$$

for vertices e_i ($m = n = 6$ or $m = 4$, $n = 6$) that are fixed points of elliptic elements γ_i of orders i in the triangle group: $\prod\gamma_i = -I_2$.)

The theory of complex multiplication in the quaternion case (Eichler-Shimura [24]) shows that for τ in H which is imaginary quadratic, the field $\mathbb{Q}(\tau,\varphi(\tau))$ is an explicit Abelian extension of $\mathbb{Q}(\tau)$.

For example, whenever $\mathbb{Q}(\tau)$ has the class number 1, the values of $\varphi(\tau)$ have the structure similar to that of $J(\tau)$. For the numbers $z = \varphi(\tau)$ we obtain Ramanujan's period relations like in the elliptic case. Namely, we get $\underline{\underline{2}}$ new algebraic relations between values of

$$F(\tfrac{1}{24},\tfrac{5}{24};\tfrac{3}{4};z) \quad \text{and} \quad F'(\tfrac{1}{24},\tfrac{5}{24};\tfrac{3}{4};z)$$

for (0,3;2,4,6) case, and between values of

$$F(\tfrac{1}{12},\tfrac{1}{4};\tfrac{5}{6};z) \quad \text{and} \quad F'(\tfrac{1}{12},\tfrac{1}{4};\tfrac{5}{6};z)$$

for (0,3;2,6,6), where $z = \varphi(\tau)$.

When $\tau \in \mathbb{Q}(\sqrt{-d})$ and $-d$ is one of the 9 one class discriminants, we arrive at new Ramanujan-like period identities. There are 3 classes of hypergeometric functions for these two triangle subgroups, where the Clausen identity is satisfied and product of the periods can be expressed as a value of a single ${}_3F_2$ function.

Unfortunately, the convergence of series in these identities is not as fast as in the original Ramanujan case. The reason is obvious: for large values of $J(\tau)$,

$$2\pi i\tau \sim \log J;$$

because at the corresponding vertex of the triangle $\ell = \infty$. In compact case, $\varphi(\tau)$ does not grow that fast.

We now describe briefly an outline of a general theory of Ramanujan-like relations for arbitrary arithmetic groups Γ. In addition to the theory of complex multiplication for these groups (see [20]-[24], [31]) one needs an analog of Ramanujan's nonholomorphic function $s_2(\tau)$. Instead of looking at the Eisenstein series corresponding to Γ, we prefer to look directly at linear differential equation corresponding, according to a Schwarz theory to Γ. We follow the approach

presented in [33] for derivation of differential equations satisfied by Eisenstein series. In this approach we look at the derivatives of the automorphic function $\varphi(\tau)$ for the arithmetic group Γ normalized by its values at vertices. The function itself satisfies the third order (nonlinear) differential equation over $\bar{\mathbb{Q}}$ (see [30] and [31]) that follows from the determination of the Schwarzian $\{\varphi,\tau\}$ in terms of φ. An analog of $s_2(\tau)$ in (1.2) that is a nonholomorphic automorphic form for Γ is

$$-\frac{1}{\varphi'(\tau)}\cdot\left\{\frac{\varphi''(\tau)}{\varphi'(\tau)} - \frac{i}{\mathrm{Im}(\tau)}\right\} \tag{2.8}$$

For $\varphi(\tau) = J(\tau)$ one gets $s_2(\tau)$ in (2.8).

For example, let us look now at a quaternion triangle case (0,3;2,6,6). In this case, instead of an elliptic Schwarz formulas (1.10)-(1.11) one has the following representation of the (normalized) automorphic function $\varphi = \varphi(\tau)$ in H in terms of hypergeometric functions:

$$\frac{\tau+i(\sqrt{2}+\sqrt{3})}{\tau-i(\sqrt{2}+\sqrt{3})} = -\frac{3^{1/2}}{2^2\cdot 2^{1/6}}\cdot\left\{\frac{\Gamma(1/3)}{\sqrt{\pi}}\right\}^6\cdot\frac{F(\frac{1}{12},\frac{1}{4};\frac{5}{6};\varphi)}{\varphi^{1/6}\cdot F(\frac{1}{4},\frac{5}{12};\frac{7}{6};\varphi)}.$$

Thus the role of π in Ramanujan's period relations is occupied in (0,3;2,6,6)-case by the transcendence

$$\left\{\frac{\Gamma(1/3)}{\pi}\right\}^6.$$

In the case (0,3;2,4,6)-group the corresponding representation for $\varphi = \varphi(\tau)$ was actually derived in [28]:

$$\frac{(\sqrt{3}-1)\tau-i\sqrt{2}}{(\sqrt{3}-1)\tau+i\sqrt{2}} = -2(\sqrt{3}-\sqrt{2})\frac{\Gamma(-\frac{1}{24})\Gamma(-\frac{5}{24})}{\Gamma(-\frac{13}{24})\Gamma(-\frac{17}{24})}\cdot\varphi^{1/2}\cdot\frac{F(\frac{13}{24},\frac{17}{24};\frac{3}{2};\varphi)}{F(\frac{1}{24},\frac{5}{24};\frac{1}{2};\varphi)}.$$

Thus, in this case, we have a new Γ-factor:

$$\frac{\Gamma(\frac{1}{24})^4}{\{\Gamma(\frac{1}{3})\Gamma(\frac{1}{4})\}^2}.$$

Appearance of such constants can be explained from the point of view of Shimura-Deligne theory of periods of CM-varieties and values of the corresponding L-functions.

We will present the new Ramanujan-like identities for this and other arithmetic triangle subgroups corresponding to quaternion algebras elsewhere, together with applications to diophantine approximations in a fashion similar to that of Chapters 4,5 and [34].

We want to point out that, though there is no algebraic relation between automorphic functions corresponding to congruence subgroups of the elliptic modular and of quaternion groups, there are arithmetic and algebraic relations between the modular curves and the modular equations for these two classes of groups. These relations between L-functions of modular forms and modular curves are predicted by general Jaquet-Langlands theory and were investigated by Shimura in special cases (particularly in Hurwitz case) and were studied by Swinnerton-Dyer [31] and by D. Heihal [59].

The generalizations of Ramanujan identities allow us to express a variety of constants, such as π and other Γ-factors, as values of rapidly convergent series with nearly integral coefficients in an infinitude of ways (rather than a single expression) with convergence improving as the discriminant of the corresponding singular moduli increases. One can ask: what kind of constants allow these representations? Values of which hypergeometric functions can be represented by such quadratic identities?

We don't even know what kind of numbers can be expressed as values of hypergeometric functions at algebraic points. (Other than to say that they are "periods".)

New Ramanujan-like period identities let us study arithmetic nature of complicated transcendences using values of (more complicated) automorphic functions. This process gives us a hope that the best diophantine approximations to such constants as π (and also periods of elliptic curves) can be interpreted through values of modular and automorphic functions, much like the best approximations to cubic irrationalties can be interpreted through values of parabolic forms of high level (via-Weil-Taniyama conjecture).

In arithmetic applications there is one more side to Ramanujan identities--their p-adic interpretation that reveals the action of Frobenius on algebraic factors in period identities. These p-adic identities involve values of p-adic hypergeometric functions and p-adic (Morita) Γ-functions, cf. with [27].

3. Diophantine approximations to numbers and methods of Padé approximations.

The main tool in the diophantine approximations is the method of Padé-type approximations. These are auxiliary polynomial or rational approximations to functions, satisfying algebraic differential or functional equations, whose specializations at, say, algebraic points provide with unusually good approximations to numbers in question. These approximations are rational in nature and should not be confused with numerical approximations often provided by iteration of sequences of algebraic maps. The crucial property of Padé-type approximations lies in the matching of degree of approximation with the degree (as an algebraic function) of an approximant--thus Padé approximations are better than simple polynomial approximations and often better than Newton ones, with obvious known exceptions such as Newton approximations to $1/(1-x)$ and $\sqrt{1-x}$.

In order to be specific, we define one of the schemes of Padé approximations (the so-called Padé approximations of the second kind [35], [36]):

Definition 3.1: Let $f_1(x),\ldots,f_n(x)$ be analytic at $x = 0$. For a given (weight) $D \geq 0$, there exist polynomials $q(x), p_1(x),\ldots,p_n(x)$ of degree at most D such that

$$\mathrm{ord}_{x=0}\{q(x)\cdot f_i(x) - p_i(x)\} \geq D + 1 + [\tfrac{D}{n}]$$

for all $i = 1,\ldots,n$.

This means that all functions $f_i(x)$ are simultaneously approximated by $\frac{p_i(x)}{q(x)}$-rational functions with the common denominator. In arithmetic applications $f_i(x)$ usually satisfy linear differential equations with additional arithmetic conditions. In our current study of periods, these equations satisfy the global nilpotence conditions [35], [37]. According to one of our results from [38] this condition is equivalent to the statement that $f_i(x)$ have convergent expansions at $x = 0$ with algebraic and nearly integral coefficients, i.e. $f_i(x)$ are Siegel's G-functions (see below [35]). Padé approximations of Definition 3.1, or Padé approximations of the second kind are used to study arithmetic properties of $f_i(x)$ at algebraic (rational) $x = z$ close to 0, by specilizing approximations x to z. Though these approximations always exist, their arithmetic properties are virtually unknown with an exception of those cases when the close form expressions are found for Padé approximations. It is not merely a closed expression that is needed, but the control on the growth of the coefficients (local, i.e. nonarchimedian conditions) versus the order of convergence of approximations (an archimedian condition).

In the simplest case this local/global condition on Padé approximations can be stated as follows:

If $f_i(x) \in \mathbb{Q}[[x]]$ $(i = 1,\ldots,n)$, can one find Padé approximants $q(x)$, $p_1(x),\dots,p_n(x)$ of weight (degree) D from $\mathbb{Z}[x]$ such that heights of all polynomials are bounded by

$$C^D$$

for a constant $C > 1$ (independent of D)?

Unfortunately, this arithmetic condition is often not satisfied even for the simplest $f_i(x)$. This is the case when $n = 1$ and $f_1(x)$ is an algebraic function with more than 3 critical points, e.g. $f_1(x) = \sqrt{(x-a_1)\dots(x-a_{2k})}$. In this and other generic cases the heights of Padé approximants from $\mathbb{Z}[x]$ grow as fast as

$$C^{D^2}$$

for $D \to \infty$, which make them unsuited for applications in diophantine approximations.

The typical compromise then, is to consider Padé-type approximations that reduce significantly the number of conditions on the order of approximations. This allows us to control the growth of coefficients, but results here are merely the existence ones, based on counting argument, and as a consequence, the rate of such approximations is very poor and usually involves fantastically large constants, cf. [34],[39].

There is something good in hypergeometric and generalized hypergeometric functions that allow for closed form expressions of Padé approximations to them and some of their combinations, and makes local/global conditions to be satisfied. (One can mention two reasons for this: a) an explicit

expression for the monodromy group and contiguous relations; b) integral representations. In all cases, when explicit construction is possible, a) and b) play important roles, see [40-41].)

We are not going to present explicitly the corresponding systems of Padé approximations but we describe the most important cases when we have a closed form expression for simultaneous Pade approximations and the arithmetic condition on local-global behavior stated above.

These cases are (see also in [34], [40-41])

1) $f_i(x) = {}_2F_1(1,b_i;c_i;x)$ or
$f_i(x) = {}_2F_1(1,b;c;\omega_i x)$

(e.g. $f_i(x) = (1-x)^{\nu_i}$ or $f_i(x) = (1-\omega_i x)^{\nu}$: $i = 1,\dots,n$). They are essentially classical Hermite-Padé approximations, [40].

2) Whenever

$$f(x) = {}_{p+1}F_p\left({a_1,\dots,a_{p+1} \atop b_1,\dots,b_p}\Big| x\right)$$

and we look at simultaneous Padé approximation to p functions $f_1(x),\dots,f_p(x)$ defined as

$$(1:f_1:\dots:f_p) = (f:\delta_x f:\dots:\delta_x^p f_p)$$

for $\delta_x = x\frac{d}{dx}$, [34], [42]. This is a generalization of Euler-Gauss continued fraction expansions of $\frac{d}{dx}\log{}_2F_1(a,b;c;x)$.

An important counterpart to 2) is given by Padé approximations to

$$ {}_{p+1}F_p\left({a_1+i,a_2,\ldots,a_{p+1} \atop b_1,\ldots,b_p}\Big|\omega_j x\right):\ i=0,\ldots,p $$

the last system of Padé approximations for $p = 2$ was used by us in applications of Ramanujan identities, see [34], [42].

In cases 1) and 2) we had determined explicitly the asymptotic expansions of Padé approximants and of the corresponding remainder functions. Explicit expressions for the asymptotics of the denominators of coefficients of Padé approximants (i.e. of heights in the integral case) turned out to be very complicated arithmetic functions of rational parameters a_i, b_i, c_i. See examples in [43].

3) Padé approximations can be also explicitly determined for Picard generalizations of hypergeometric functions (Pochhammer integrals):

$$ F(\mu_0,\ldots,\mu_{d+1}) = \int_1^\infty t^{-\mu_0}(t-1)^{\mu_1}\prod_{i=2}^{d+1}(t-x_i)^{-\mu_i}\,dt. $$

Here μ_i are rational numbers and singularities x_i are linear functions of a single variable x. (These functions are general section of Picard function, where all x_i are treated as variables; in this case also one can construct multidimensional Padé approximations.)

In Picard cases there is a large class of arithmetic monodromy groups. In these cases an analog of complex multiplication theory, and theory of Ramanujan period relations can be constructed. This theory has interesting applications to diophantine approximations of numbers that arise from periods

of curves and algebraic varieties; most interesting ones are not yet fully explored. The generating functions of Padé approximants to Picard integrals are themselves expressed as periods of algebraic varieties.

4) Finally, there are explicit Padé approximations to multidimensional generalization of generalized hypergeometric functions, which are expressed as integrals over powers of polynomials in complex variables taken over polytops. We present only one example of a nice multidimensional integral arising as a natural generalization of Hermite-Lindemann proof [17] of transcendence of e and π. This formula can be considered as a multidimensional analog of operational calculus formula for Laplace transforms.

In this formula:

$$ I_\Delta = \int\cdots\int_\Delta e^{-\Sigma_{i=1}^n x_i y_i}\cdot P(y_1,\ldots,y_n)\prod_{i=1}^n dy_i, $$

where Δ is a polytope in n-dimensional space, P is a polynomial vanishing up to high orders at vertices of Δ. The integral I_Δ can be evaluated through values of P and its derivatives at vertices of Δ only. This is a generalization of the so called Hermite identity (see [17], [44]):

$$ I_\Delta = \sum_{\bar e\in V(\Delta)} e^{-(\bar x,\bar e)}\cdot\prod_{i=1}^n\left(\ell_i\left(\frac{\partial}{\partial y_1},\ldots,\frac{\partial}{\partial y_n}\right)-\ell(x_1,\ldots,x_n)\right)^{-1}\cdot P|_{\bar y=\bar e} $$

where $\ell_i(y_1,\ldots,y_n) = 0$ are hyperplane equations of sides of Δ intersecting at the vertex $\bar e \in V(\Delta)$.

The choice of the polytope and the polynomials vanishing up to high order in all its vertices determines a variety of Padé approximations to linear combinations of exponents. Hermite's simultaneous Padé approximation correspond to n = 1 and give Padé approximations to $e^{\omega_i x}$. Whenever Δ is an n-polytop (tetrahedron in $\mathbb{R}^n$), one obtains Padé approximations used by Hermite, Mahler and Siegel to estimate diophantine approximations to e^{α} and π, see [44]. An interesting intermediate case corresponds to the so called graded Padé approximations [40], [45] that provide sharp measures of diophantine approximations to such numbers as sin 1, $J_0(1)$, etc., though for the corresponding functions (sin x, $J_0(x)$ et. al.) no regular pattern continued fraction exists and arithmetic (local/global) conditions for rational approximations (as above) are not satisfied.

We can use explicit Padé approximations, particularly those to generalized hypergeometric functions to establish sharp irrationality measures for numbers that occur as values of these functions close to the point of nearly integral power series expansions.

We used Ramanujan's amazing generalized hypergeometric identities representing quadratic period relations and their generalizations, as described in Chapter 2, to derive new bounds on measures of diophantine approximations for a variety of numbers mainly connected with π. We are able to obtain these strong bounds because Ramanujan representations give expressions for classical constants as rapidly convergent

series with nearly integral (or integral) coefficients. Though from the point of view of numerical approximations these new series are not dramatically better than some of the classical ones, performing Padé approximations on them we arrive at incredibly good arithmetic rational approximations, often very close to the best rational approximations that only continued fraction expansion provides. (One cannot expect to get the continued fraction expansion this way, because no regular pattern emerges, at least experimentally, in the expansion of these numbers.)

What Ramanujan's and similar identities provide is a new identification of classical constants such as π or $\pi\cdot\sqrt{2}$ (or log $\epsilon_{\sqrt{k}}$) as values of nearly integral power series expansions of solutions of linear differential equations (primarily of hypergeometric and generalized hypergeometric type) very close to the point of expansion.

The closer we are to the origin, the better is the exponent of irrationality! To give a precise statement, we quote one of our results on G-functions [35].

We recall that a function $f(x) = \Sigma_{n=0}^{\infty} a_n x^n$ with $a_n \in \mathbb{Q}$ is called Siegel's G-function [35], [38], if a_n are nearly integral, i.e. f(x) converges around x = 0 and the common denominator of $a_0,\dots,a_n$ grows not faster than a geometric progression in n. All generalized hypergeometric functions with rational parameters are G-functions, as well as solutions of Picard-Fuchs equations.

One of our results relating diophantine approximations of G-functions with the global nilpotence of the corresponding differential equations states:

Theorem 3.2 [35]: Let $f_1(x),\ldots,f_n(x)$ be G-functions satisfying linear differential equations over $\mathbb{Q}(x)$. Let $r = a/b$, with integers a and b, be very close to the origin. Then we get the following lower bound for linear forms in $f_1(z),\ldots,f_n(z)$.

For arbitrary non-zero rational integers $H_1,\ldots,H_n$ and $H = \max\{|H_1|,\ldots,|H_n|\}$, if $H_1f_1(r) +\ldots+ H_nf_n(r) \neq 0$,

$$|H_1f_1(r) +\ldots+ H_nf_n(r)| > |H_1 \ldots H_n|^{-1}\cdot H^{1-\epsilon}$$

provided that r is very close to 0:

$$|b| \geq c_1\cdot|a|^{n(n-1+\epsilon)}$$

and $H \geq c_2$ with effective constants $c_i = c_i(f_1,\ldots,f_n,\tau,\epsilon)$. If r is not as close to 0, we get only

$$|H_1f_1(r) +\ldots+ H_nf_n(r)| > H^{\lambda-\epsilon}$$

for $\lambda = -(n-1)\log|b|/\log|b/a^n|$ (< 0).

This theorem shows what a qualitative difference in the exponent of diophantine approximation, the rate of convergence of a series representing a constant makes.

Looking at the rate of convergence of Ramanujan-like series and combining them with the G-function theorems, one obtains extremely strong results

all numbers connected with π have (Roth's) exponent of irrationality $2 + \epsilon$ for any $\epsilon > 0$.

Mock Proposition. If there exist infinitely many negative quadratic discriminants $-d$ with a fixed class number, then all elements of the field $\mathbb{Q}(\pi^2)$, irrational over $\mathbb{Q}$, have measures of irrationality with the exponent $2 + \epsilon$ for all $\epsilon > 0$.

The same result holds also for fields generated by $\log^2\epsilon_{\sqrt{k}}$ for real quadratic units $\epsilon_{\sqrt{k}}$.

Unfortunately, degrees of numbers $J(\tau)$ for $\tau = \frac{1+\sqrt{-d}}{2}$ grow to infinity as $d \to \infty$, thus we are confined (for a given degree) only to the finitude of cases of rapidly convergent series. This means that exponent stays away from 2. Also for all these cases, instead of general results we have to use explicit Padé approximations to generalized hypergeometric functions and Pochhammer integrals [34], [42].

We review briefly the best results on measures of diophantine approximations to number connected with π obtained using effective Padé approximations to generalized hypergeometric functions.

$$|\pi\sqrt{2} - \frac{p}{q}| > |q|^{-16.67\ldots}$$

for rational integers p, q: $|q| \geq q_0$. This is based on Ramanujan's series for $1/\sqrt{2}\pi$;

$$|\pi\sqrt{640{,}320} - \frac{p}{q}| > |q|^{-11.109\ldots}$$

For $\pi\sqrt{3}$ we use slightly different systems of Padé-type approximations [34], [46]:

$$|\pi\sqrt{3} - \frac{p}{q}| > |q|^{-5.791\ldots}$$

for $|q| \geq q_2$.

Finally, we can give one unconditional result for π:

Proposition 3.3: For arbitrary rational integers p,q with $|q| \geq 2$ we have:

$$|q\cdot\pi-p| > |q|^{-14.0}.$$

In the finite range $|q| \leq 10^{10^5}$ we had to look at the explicit continued fraction expansion of π.

In fact we know slightly better (conditional) measures of irrationality:

$$|\pi - \frac{p}{q}| > |w|^{-15.62\ldots}$$

for $|q| \geq q_3$, and the following one for π^2:

$$|\pi^2 - \frac{p}{q}| > |q|^{-7.81\ldots}$$

for $|q| \geq q_4$.

Let us look on the transcendence and algebraic independence results for periods and quasiperiods of hypergeometric curves, and the corresponding statements the transcendence of values of hypergeometric functions at algebraic points. (As it was already explained, with the exception of two classes of arithmetic triangle groups, in general, periods of a single

hypergeometric curve involve values of more than a single hypergeometric function.)

First of all, trivial cases have to be excluded: we remove from consideration all hypergeometric functions reducible to algebraic ones (exclude all equations on the Schwarz's list of finite groups when ratio of two solutions is an algebraic function). Next, only rational values of parameters are to be considered. (This corresponds to the G-functions or to a finite genus case; little to nothing is known about values of hypergeometric functions with algebraic irrational parameters.) Other hypergeometric functions reducible to elementary or genus 1 cases have to be excluded too; they are analyzed by other means.

Let us start with the simplest case of Legendre functions $F(\frac{1}{2},\frac{1}{2};1;z)$. In this case (or in any other case algebraically reducible to this one including triangle cases commensurable with the modular group) it is a consequence of our 1977 [26], [39] result that:

Theorem 3.4: For an arbitrary algebraic $z \neq 0,1$ two numbers $F(\frac{1}{2},\frac{1}{2};1;z)$ and $F(-\frac{1}{2},\frac{1}{2};1;z)$ are algebraically independent (over $\mathbb{Q}$).

Moreover, one can get the measure of algebraic independence of these two numbers: $\alpha = F(\frac{1}{2},\frac{1}{2};1;z)$, $\beta = F(-\frac{1}{2},\frac{1}{2};1;z)$. This measure is very close to the best possible: if $P(x,y) \in \mathbb{Z}[x,y]$, $P \not\equiv 0$, P has the (total) degree d and height (the maximum of absolute values of coefficients) $H > 1$, then [39]:

$$|P(\alpha,\beta)| > H^{-c_0 d^2 \log^2 (d+1)}$$

with $c_0 = c_0(\alpha,\beta,z) > 0$--an effective constant.

Theorem 3.4 can be supplemented with a stronger statement [26] that, whenever, $z \neq 0,1,\infty$ two of three numbers:

$$z, F(\tfrac{1}{2},\tfrac{1}{2},1;z) \quad \text{and} \quad F(-\tfrac{1}{2},\tfrac{1}{2};1;z)$$

are algebraically independent (over $\mathbb{Q}$).

It is natural to conjecture that for all other nontrivial (i.e. with an exception of the cases stated above) cases analog of Theorem 3.4 hold for hypergeometric functions corresponding to the hypergeometric curves of genus $g > 1$. We do not yet have such a strong result, though the analogue of Theorem 3.4 holds, if all hypergeometric functions giving periods of hypergeometric curves are considered. In particular, for all hypergeometric functions algebraically reducible to the triangle case $(0,3;2,4,6)$ (or $(0,3;2,6,6)$), the values $F(z)$ and $F'(z)$ are algebraically independent over $\mathbb{Q}$, whenever z is algebraic $\neq 0,1$ and F is the corresponding hypergeometric functions. (Note that if one considers in Theorem 3.4 and its generalization <u>all</u> branches of a given hypergeometric function and its derivatives, there are still only two algebraically independent numbers, whenever z is a singular moduli. This is due to the Ramanujan period relations.)

Let us look now at the general algebraic independence results for hypergeometric function when all branches of the function are involved. In this case we use the knowledge of the monodromy group of the hypergeometric equation and the uniformization of hypergeometric functions by functions meromorphic in $\mathbb{H}$. We use methods of [46] based on the uniformization $x = k^2(\tau)$ of general hypergeometric functions $F(x)$.

<u>Theorem 3.5:</u> Let $F_1(x)$ $(= F(a,b;c;x))$ and $F_2(x)$ $(= x^{1-c}\cdot F(a-c+1,b-c+1;2-c;x))$ be two algebraically independent (over $\mathbb{Q}(x)$) solutions of hypergeometric equation with rational a,b,c.

If F_1, F_2 do not correspond to genus $= 0,1$ cases, then for every algebraic $x \neq 0,1$ at least two of the numbers

$$F_1(x),\ F_1'(x),\ F_2(x),\ F_2'(x)$$

are algebraically independent.

There are at most 3 algebraically independent numbers among these and in singular moduli case there are exactly 2 algebraically independent ones.

4. The uniformization theory and accessory parameters for hyperelliptic curves.

In the nonarithmetic case, one is still interested in the arithmetic nature of groups (Fuchsian groups acting on H) uniformizing Riemann surfaces corresponding to algebraic curves of genus $g \geq 2$ defined over $\bar{\mathbb{Q}}$. In the uniformization problem we are using classical language of Schwarzians, Fuchsian linear differential equations, monodromy groups (representing groups of fractional transformations), and accessory parameters in the differential equations on which these groups depend. This language was used early in the development of the uniformization theory by Klein and Poincaré [47] when the continuity method was the main tool. The continuity method is no longer necessary to prove results in this theory. Nevertheless, the basis of classical theory--the one-to-one relationship between the uniformization of Riemann surfaces of genus $g \geq 2$ and certain linear differential equations of the second order depending on $3g-3$ complex accessory parameters can be used for the explicit numerical (or analytic, if possible) solution to the uniformization problem. In this approach one has to find a Fuchsian group, its fundamental domain and the uniformizing functions of a Riemann surface given by an explicit algebraic equation. There are several existing

numerical approaches to this problem (Myrberg method based on Koebe theorem, methods based on integral equations for an appropriate Riemann problem), but the best method requires the determination of accessory parameters. If the uniformization problem is solved via the accessory parameters, then one has to determine $6g - 6$ real variables in the second order Fuchsian equation, for which the monodromy group of this equation is represented by real 2×2 matrices (so that this group is Fuchsian). If one has an efficient algorithm that computes the monodromy group of a second order equation, this algorithm becomes a key subroutine in the program that searches for the choice of accessory parameters for which traces of products of all monodromy matrices have zero (or very close to zero within a given precision) imaginary parts. This iterative program can be viewed, e.g. as a descent method to find the global minima of squares of these imaginary parts. The minima (equal to zero) is unique and is achieved at Fuchsian accessory parameters.

We apply our methods of fast computation of solutions of linear differential equations (see Chapter 5 and [12]) to fast computations of monodromy groups. This allows us to apply the descent method to find the Fuchsian accessory parameters, when their number is not too large. In practice, one has to be content with $g \leq 6$, when the Riemann surface and the differentials on it are known. In this lecture we report on our computations for hyperelliptic curves of genus $g \geq 2$

which had been studied in detail because of the special role of hyperelliptic curves, and also because the number of accessory parameters is smaller (only $2g-1$ accessory parameters) and the Fuchsian equation depending on them has polynomial coefficients. Some time ago we have conducted similar computations for the Riemann surfaces of genera 0 and 1 with several prescribed branch points, see [12].

In our study of hyperelliptic curves we were particularly interested in possible closed form expressions of accessory parameters in terms of known functions. The Whittaker conjecture [11] expressed the accessory parameters as algebraic functions of the branch points. This elegant conjecture turned out to be incorrect, and only special cases of algebraic expressions for accessory parameters were found.

Let us describe briefly the role of accessory parameters.

Let us consider an algebraic curve given by an equation $P(x,y) = 0$ for an irreducible polynomial $P(x,y)$. One can interpret this curve as a ramified covering of a Riemann sphere $\mathbb{CP}^1$ by looking at y as a multivalued function of x. Let us assume that this algebraic curve has a genus $g \geq 2$, and let us denote the corresponding compact Riemann surface by Γ. If $\hat{\Gamma}$ denotes the universal covering surface of Γ, then the existence theorem of the uniformization theory implies the conformal equivalence of $\hat{\Gamma}$ to the upper-half plane H. This conformal equivalence with H manifests itself as the uniformization of Γ:

$$\varphi: H \to \Gamma.$$

The group of cover transformations of Γ is $G = \{T: \varphi(T\circ z) = \varphi(z) \text{ for all } z \in H\}$. Here G is a discrete group of Mobius transformations over $\mathbb{R}$, i.e. a Fuchsian group.

It is easier to look at the inverse map $z = \varphi^{-1}$ than on φ. The function z is multivalued on Γ and its monodromy is related to the action of G: $z \to \sigma(z)$ for $\sigma \in G$. Since G is a group of Mobius transformations and the Schwarzian derivatives $\{\cdot,x\}$ ($\{f,x\} = (\frac{f''}{f'})' - \frac{1}{2}(\frac{f''}{f'})^2$) are invariant under the fractional transformations, the function $\{z,x\}$ is a single-valued function on Γ.

This means that

$$\{z,x\} = 2R(x,y) \text{ on } \Gamma \tag{4.1}$$

for some rational function $R(x,y)$. The classical relationship between the Schwarzian and the second order differential equations states that a change of variables

$$y_1 = \left(\frac{dz}{dx}\right)^{-1/2}, y_2 = z\left(\frac{dz}{dx}\right)^{-1/2}, \quad z = y_2/y_1$$

reduces (4.1) to a second order linear differential equation on y_1 and y_2:

$$\left(\frac{d^2}{dx^2} + R(x,y)\right)y = 0. \tag{4.2}$$

Whenever z is locally one-to-one, $R(x,y)$ can be determined as [48]

$$R(x,y) = R_0(x,y) + Q(x,y),$$

where $Q(x,y)dx^2$ is a regular quadratic differential and $R_0(x,y)$ is a second derivative of an Abelian integral of the third kind (an explicit function that can be determined whenever Γ is known algebraically). If one knows the basis $\{Q_j(x,y): j = 1,\dots,3g-3\}$ of regular quadratic differentials on Γ, then an equation (4.2) can be represented as [48]:

$$\left(\frac{d^2}{dx^2} + R_0(x,y) + \sum_{i=1}^{3g-3} c_j \cdot Q_j(x,y)\right)y = 0 \qquad (4.3)$$

for 3g-3 complex parameters c_j $(j = 1,\dots,3g-3)$--the accessory parameters. For an analytic continuation of a (chosen) basis (y_1,y_2) of (4.3) around the loop γ of $\pi_1(\Gamma;\mathcal{O})$ we have a monodromy matrix M_γ:

$$(y_1,y_2)^t \underset{\gamma}{\to} M_\gamma (y_1,y_2)^t.$$

When the accessory parameters c_j $(j = 1,\dots,3g-3)$in (4.3) are Fuchsian, i.e. (4.3) indeed arises, as above, from the uniformization of Γ by G: $y_2/y_1 = z = \varphi^{-1}$ of $\varphi: H \to \Gamma$-- then the monodromy group of (4.3) generated by $\{M_\gamma\}$ is G, up to conjugation in $SL_2(\mathbb{C})$.

Consequently for a given Γ (and properly chosen R_0 and Q_j in (4.3)), there existsa unique group of c_j: $j = 1,\dots,3g-3$ (6g-6 real parameters) for which the corresponding monodromy group of (4.3) is Fuchsian (i.e. is a subgroup of $SL_2(\mathbb{R})$).

This description is slightly different for a hyperelliptic curve. In this case simple transformations allow us to

reduce (4.2) to an equation with coefficients polynomial in x. Following [49] one considers a general hyperelliptic equation of genus g:

$$y^2 = (x-e_1)\cdots(x-e_{2g+2}) \qquad (4.4)$$

Arguments of Schottky in the general case and Whittaker [49] for hyperelliptic curves, show that the Fuchsian group G of an equation (4.4) is a subgroup of index 2 of a group G^* (of self-inverse transformations, i.e. elliptic transformations of period 2). The group G can be easily found once G^* is known. The group G^* has genus 0 and thus its automorphic function can be expressed as a rational function of a single automorphic function. This means that in equation (4.2) corresponding to the hyperelliptic curve (4.4), the right hand side is a rational function of x with singularities only at e_i's. Analyzing the behavior of the uniformizing variable at ∞ one gets a clear description of (4.2) (Whittaker, [11]): The uniformizing variable z of (4.4) is represented as a ratio of two linearly independent solutions of the following second order linear differential equation

$$\left(\frac{d^2}{dx^2} + R(x)\right)y = 0, \quad \text{where}$$

$$R(x) = \frac{3}{16}\cdot\left\{\sum_{i=1}^{2g+2} \frac{1}{(x-e_i)^2} - \frac{2(g+1)x^{2g} - 2g\sum_{i=1}^{2g+2} e_i\cdot x^{2g-1} + \sum_{i=0}^{2g-2} C_i\cdot x^{2g-2-i}}{(x-e_1)\cdots(x-e_{2g+2})}\right\} \qquad (4.5)$$

Here $2g-1$ parameters $C_0,\dots,C_{2g-2}$ are accessory parameters. The Fuchsian accessory parameters C_i are determined uniquely by the condition that the monodromy group of (4.5) is Fuchsian, i.e. is represented by real 2×2 matrices.

Whittaker in [11], studying the birational transformations of curves (4.4) and the corresponding transformations of accessory parameters C_i in (4.5) found explicit expressions for uniformization of a special case of (4.4):

$$y^2 = x^5 + 1 \tag{4.6}$$

(here $g = 2$ and $e_6 = \infty$). The group G^* is conjugate to the group generated by the 5 transformations

$$\sigma_j^*: x \mapsto \frac{\alpha x - \epsilon_j}{\epsilon_j^* x - \alpha}: \; j = 0,\dots,4$$

for $\epsilon_j = \exp((4j+1)\pi i/10)$: $j = 0,\dots,4$ and $\alpha = \sqrt{\frac{\sqrt{5}+1}{2}}$. The group G itself is generated by $\sigma_0\sigma_1,\dots,\sigma_0\sigma_4$, and the corresponding linear differential equation in this case is reduced to a particular Gauss hypergeometric equation.

Whittaker called automorphic functions in the case (4.6) hyperlemmiscate functions, and on the basis of (4.6) and other similar examples had been led to a conjecture describing $R(x)$ in (4.5) explicitly:

Conjecture 4.1: Let $P(x) = (x-e_1)\dots(x-e_{2g+2})$. Then the Fuchsian accessory parameters $C_0,\dots,C_{2g-2}$ in (4.5) are polynomial in e_i, and we have the following explicit expression

for $R(x)$ in (4.5):

$$R(x) = \frac{3}{16}\cdot\left\{\left(\frac{P'(x)}{P(x)}\right)^2 - \frac{2g+2}{2g+1}\cdot\frac{P''(x)}{P(x)}\right\}. \tag{4.7}$$

Whittaker's students and collaborators studied this conjecture. Recent results on the Whittaker conjecture belong to Rankin [50], who showed the truth of the Whittaker conjecture for a large number of equations associated with one of the finite groups on the Schwarz's list, when (4.5) is related to the hypergeometric equation.

We ran numerical checks for random curves (4.4) of genus g, $2 \le g \le 10$ (particularly with integral e_i); the Conjecture 4.1 is incorrect.

Our numerical experiments with Fuchsian accessory parameters C_i in (4.5) suggest that the form of C_i predicted by (4.7) in Conjecture 4.1 holds only in cases when the differential equation (4.5) is reducible by an algebraic transformation to Gauss hypergeometric equations. Consequently, if our evidence holds, for a given g there would be only finitely many hyperelliptic equations (4.4) for which the accessory parameters C_i in (4.5) are expressed as polynomials of e_j as in (4.7). Still there are quite a few such equations. Rankin [50] presented some interesting examples, e.g. $y^2 = x(x^{10}+11x^5-1)$ (cf. with Apéry's equation for recurrences approximating $\zeta(2)$, see last paragraph of [46] and [12], and $3y^2 - 2y^3 = 4x^3 - 3x^4$; $y^m = x^p(x^q-1)^r$; $y^m = \left(\frac{x^8+14x^5+1}{x^4(x^4-1)^4}\right)^3$ etc.

Even if accessory parameters are not globaly algebraic (as functions of branch points in case of (4.4)), one is still very much interested in their arithmetic properties. Our particular attention is focused on the following problem [12]:

Problem: Let us consider a linear differential equation (4.5) for the hyperelliptic equation (4.4) that gives a Fuchsian uniformization (or the corresponding equation (4.2) for the general algebraic curve). Let us assume that this algebraic curve is defined over $\bar{\mathbb{Q}}$ (or even $\mathbb{Q}$). Are the accessory parameters algebraic? If they are not algebraic, are they algebraically independent?

The algebraicity of accessory parameters is known only in cases when the answer to the Whittaker conjecture is positive and, in general, when the equation (4.5) is reducible to a Gauss hypergeometric equation.

We conducted multiprecision computations of accessory parameters for a generic curve (4.4) with integral e_i. Our computations seem to indicate the transcendence of accessory parameters. Namely, we determined the absence of algebraic relations of moderate degrees (100 at most) with moderate size integral coefficients (up to 10^{100}). We do not know of any examples of algebraicity of accessory parameters other than in cases arising from Schwarz triangle functions or other arithmetic subgroups.

This class of problem is a part of a more general problem

on the transcendence of elements of the monodromy matrices of Fuchsian linear differential equations. Next to nothing is known about this problem. E.g. among the second order Fuchsian linear differential equations the transcendence of elements of the monodromy matrices had been studied only in case when a monodromy group of this equation is commutative.

We are able to prove some transcendental results for monodromy groups of Fuchsian differential equations. Our methods stem from the uniformization of these Fuchsian equations by special subgroups of $SL_2(\mathbb{C})$ using the uniformization of algebraic curves by special functions in H having only 3 branch points [46]. The only condition we impose is the existence of a G-function solution for this differential equation.

Theorem 4.2: Let a Fuchsian differential equation $L[\frac{d}{dx},x]y = 0$ of order n be satisfied by a transcendental G-function $f(x)$ that does not satisfy any equation over $\bar{\mathbb{Q}}(x)$ of order less than n. If $L[\frac{d}{dx},x] \in \bar{\mathbb{Q}}(x)[\frac{d}{dx}]$, and $\mathfrak{M}$ is a monodromy group of L, corresponding to a choice of a fundamental system of solutions $y_1,\ldots,y_n$ of $Ly = 0$ with algebraic initial conditions $y_j^{(i)}(x_1)$ for an algebraic x_1, then at least one element of one matrix from $\mathfrak{M}$ is transcendental, provided that the base point x_1 of $L[.]y = 0$ is nonsingular.

5. Bignum and high precision computations of solutions of linear differential equations.

Many computations associated with number theory require manipulations with integers or real numbers of large sizes. These numbers are called bignums and bigfloats as opposite to the fixed precision numbers that are primitives in computer operations and most high-level languages. Obviously, the speed of manipulation with bignums and bigfloats starts to depend on their sizes. In order to highlight this dependence we will differentiate between operational and boolean (bit) complexity. By operational complexity of computations we understand the number of operations (additions, multiplications, comparisons and copyings) independently of the sizes of numbers involved. On the other hand, by the boolean (or bit) complexity of computations we understand the total number of operations required to complete a given program. Here the primitive operations are understood to be: additions, multiplications, comparisons, storage and retrieval of short numbers (say, single digit or single bit numbers). (This definition is most closely associated with modern serial machine, where each of the primitive operations takes about the same time.)

The operational complexity does serve as a good measurement of computational time only if all intermediate operands

have bounded length. This is the case of: a) floating point computations with a fixed precision M; or b) computations in modular arithmetic, where all results are considered mod a fixed number. (Still the computational time will differ from the operational complexity by a factor equal to the complexity of multiplication of primitive numbers in cases a) or b).)

To understand the differences between operations on short and long numbers one should consider the multiplication of bignums.

Let us denote by $M_{bit}(n)$ the bit complexity of multiplication of two n-bit integers. The best known upper bound on $M_{bit}(n)$ belongs to Schonhage-Strassen [51]:

$$M_{bit}(n) = O(n \log n \operatorname{loglog} n).$$

It is widely assumed that $M_{bit}(n) = O(n \log n)$ (and this was, in fact, proved for a large set of n). It is also conjectured, though without too much of supporting evidence, that the last bound is tight: $M_{bit}(n) = \Omega(n \log n)$. For related complexity investigation see [52], [53].

The $M_{bit}(n)$ quantity is indeed crucial. In comparison, a bit complexity of addition is relatively simple: $M_{add}(n) = O(n)$ (in the scalar case; in parallel case, or on VLSI, it is not that simple). The multiplication of bigfloats is, obviously, as bit complex as that of bignums.

A (normalized) bigfloat of precision n is denoted by $.b_0 \ldots b_n \times \text{SIGN} \times B^{\text{EXP}}$ $(= \text{SIGN}\cdot\{b_0\cdot B^{\text{EXP}-1} + b_1\cdot B^{\text{EXP}-2} + \ldots\})$, where B is the base: $B = 2$, or, say, $B = 10$, SIGN is the sign: 0, ± 1 and EXP is the exponent. The normalization condition is $b_0=1$ for SIGN$\neq$0 and the overflow conditions is EXP $= O(n\log n)$.

All algebraic operations on bigfloats have bit complexity of the same order of magnitude as a multiplication. For example, let $B(n)$ denote one of the following bit complexities: division of n-bit bigfloat numbers, square root extraction, or raising to the fixed (rational) power. Then $B(n) = O(M_{bit}(n))$, and $M_{bit}(n) = O(B(n))$. (The determination of the best constant factors in these equations is an interesting problem. E.g. if $D(n)$ is a bit complexity of inverting an n-bit bigfloat, then the Newton iterations to $1/x-y = 0$ give a bound $D(n) \leq 4\cdot M_{bit}(n) + O(n)$. What is the best constant instead of 4? cf. [53].)

To determine the complexity of transcendental operations on bigfloats is not that easy. A general problem of bigfloat computations can be formulated as follows:

Problem: Let $F(x)$ be a function defined by algebraic, differential (functional or integral) equations or their combinations, with initial or boundary conditions at rational points or arbitrary bigfloats of full n-bit precision. Let this function $F(x)$ be defined uniquely in the neighborhood of a bigfloat x_0 as an analytic function. Compute $F(x_0)$ with the full n-bit precision with the minimal bit complexity.

Though this question is a practical one, and high precision computations are often needed in mathematical physics, particularly in computation of instabilities and attractors, most often high precision computations are carried out for particular constants of number-theoretic interest. Computations of γ (the Euler constant) and, particularly, π became a major undertaking in computational mathematics, often expensive. Low operational complexity algorithms, for computation of the logarithmic function based on Landen's transformation for elliptic functions (the Gauss arithmetic-geometric mean iteration) were proposed by Salamin and improved by Brent [54] and Borwein-Borwein [5]. These low operational complexity algorithms are translated into low bit complexity algorithms for computation of values of the logarithmic function. Namely, to compute the value of $\log x$ (and to be nonambiguous, we can assume that $|x-1| < 1/2$), with the full n-bit precision, starting from the n-bit x, one needs bit-complexity of at most $O(M_{bit}(n)\cdot \log n)$. In particular, n digits of π can be computed in that many short (bit) operations. These particular methods are used in recent computation of π.

For special functions there is no low operational complexity methods of computation that are known to give low bit complexity. That is why we propose low bit complexity methods of computations of values of special functions, not related to any rapidly convergent analytic transformations. Our methods differ in bit complexity by $\log^2 n$ or $\log^3 n$ from the bit complexity of algebraic computations.

(It is an interesting problem [5], whether the computation of n digits of a classical transcendence, like π, can be done in $O(M_{bit}(n))$ only. Probably not, but a proof will be more difficult than the transcendence proof.)

Our new low bit (boolean) complexity algorithms of evaluation solutions of (linear) differential equations are based on the reduction to fast evaluation of solutions of linear difference equations with polynomial coefficients. We start first with the reduction from linear differential equations to the difference ones, and then describe low operational and bit complexity methods of solution of linear difference equations.

Let us look at an arbitrary linear differential equation with rational (polynomial) coefficients, either in the scalar form

$$a_m y^{(m)} + a_{m-1} y^{(m-1)} + \dots + a_1 y' + a_0 y = 0, \tag{5.1}$$

or, in the general matrix form,

$$\frac{d}{dx} Y(x) = A(x) \cdot Y(x),$$

where $a_i \in \mathbb{C}(x)$, and $A(x) \in M_n(\mathbb{C}(x))$. We are interested in the evaluation of solutions of (5.1) or (5.2) with an arbitrary precision using the method of (formal) power series expansions [56]. A solution of (5.1) or (5.2) is determined by its

initial conditions at $x = x_0$. If $x = x_0$ is not a singular point of (5.1) or (5.2), then the solution $y = y(x)$ of (5.1) or $y(x)$ of (5.2) is uniquely determined by its initial conditions $y(x_0), \dots, y^{(m-1)}(x_0)$ in the case of (5.1), or $n \times n$, matrix $Y(x_0)$, in the case of (5.2), respectively. If $x = x_0$ is a regular singular point of (5.1) or (5.2), then the initial conditions take the form of a few leading terms in the expansion of $y(x)$ or $Y(x)$ in powers of $(x-x_0)^\alpha$ for various local exponents α. Finally, if $x = x_0$ is an irregular singularity of (5.1) or (5.2), one can interpret as initial conditions at $x = x_0$ few first terms in the asymptotic expansions that are linear combinations of formal powers series from $\mathbb{C}[[x-x_0]]$ times functions like $\exp(Q((x-x_0)^{-\alpha}))$ for $Q(x) \in \mathbb{C}[x]$. In all these cases, having specified initial conditions for $y(x)$ or $Y(x)$ at $x = x_0$, we want to evaluate within a given precision ℓ $y(x)$ or $Y(x)$ at another point $x = x_1$. For all practical purposes we assume that values of x_0 and x_1 are given correctly with the precision of ℓ bits (or decimal digits), or as rational or algebraic numbers of sizes less than ℓ. To determine values at $x = x_1$ from those at $x = x_0$ one has to specify a path from x_0 to x_1 on the Riemann surface (or its universal covering) of $y(x)$ or $Y(x)$. This problem of analytic continuation had been studied by us in [12], and we will return to it shortly. Now we look at the most important case when $x = x_1$ lies within the disc of convergence of power series expansions defining $y(x)$ or $Y(x)$. [In the case of a regular point $x = x_0$ this is simply the disc of convergence of $y(x)$ or $Y(x)$. If $x = x_0$ is a regular or irregular singularity, and the expansion at $x = x_0$ looks like $y(x) = \varphi_0(x)\cdot y_0(x)$, where $y_0(x) \in \mathbb{C}[[x-x_0]]$, and $\varphi_0(x) = (x-x_0)^{\alpha}$ or $\varphi_0(x) = \exp(Q((x-x_0)^{-\alpha}))$, then $x = x_1$ should lie within the disc of convergence of $y_0(x)$.]

In the cases of (5.1) and (5.2), when the coefficients are rational functions and the set S of singularities of (5.1) or (5.2) is a discrete (finite) set, the radius of convergence of $y(x)$ or $Y(x)$, if nonzero, is bounded from below by the distance from the point $x = x_0$ to the nearest point in S. We consider now only the case of nonzero radius of convergence (though, the general case can be treated in the same framework using the generalized Borel transform). For this reason we now consider the case when $x = x_0$ is a regular or a regular singular point only. For (5.1), the basis of solutions at $x = x_0$ can be expressed in terms of regular expansions

$$y(x,\alpha) = (x-x_0)^{\alpha} \sum_{n=0}^{\infty} y_n(\alpha)\cdot(x-x_0)^n \tag{5.3}$$

The coefficients $y_n(\alpha)$ are determined for $n \geq 0$ from the initial conditions at $x = x_0$ and the linear recurrence

$$\sum_{j=0}^{\min(n,d)} y_{n-j}(\alpha)\cdot f_j(\alpha+n-j) = 0, \tag{5.4}$$

$n = 1,2,3,\ldots$, with the explicit expression of coefficients $f_j(\beta)$ in terms of a_i in (5.1) as follows: if $a_j = Q_j(x)\cdot(x-x_0)^j$ $(j = 0,\ldots,m)$, then for $f(x,\alpha) \overset{\text{def}}{=} \alpha(\alpha-1)\ldots(\alpha-m+1)\cdot Q_m(x) +\ldots+ \alpha Q_1(x) + Q_0(x)$, we put $f(x,\alpha) = f_0(\alpha) + (x-x_0)f_1(\alpha) +\ldots+ (x-x_0)^d f_d(\alpha)$, where d is the bound of degrees for all polynomials $Q_j(x)$. The exponent α in (5.3) satisfies the indicial equation $f_0(\alpha) = 0$ [56].

Similarly, we look at regular solutions $Y(x)$ of (5.2) having a regular expansion of the form

$$Y(x) = \{\sum_{N=0}^{\infty} C_N(x-x_0)^N\}\cdot(x-x_0)^W, \tag{5.5}$$

where C_0 $(\in M_{n\times n}(\mathbb{C}))$ is the initial condition for $Y(x)$ at $x = x_0$, and the (matrix) coefficients C_N of (5.5) are determined from the matrix linear recurrence of length d (the maximal degree of the rational function in $A(x)$ in (5.2)). To derive this recurrence, let us consider a case of a regular point $x = x_0$ when $W = 0$. For $A(x)$ from (5.2) let us put $A(x) = A_0(x)/d(x)$, where $A_0(x) \in M_{n\times n}(\mathbb{C}[x])$, $d(x) \in \mathbb{C}[x]$. We put $d(x) = \Sigma_{j=0}^{d} d_j(x-x_0)^j$, where $d_0 \neq 0$, and $A_0(x) = \Sigma_{j=0}^{d-1} A_j(x-x_0)^j$. Then we have the following recurrence on C_N:

$$C_{N+1}\cdot(N+1)\cdot d_0 = \sum_{i=0}^{\min\{N,d-1\}} A_i\cdot C_{N-i} - \sum_{i=0}^{\min\{N,d-1\}} d_{i+1}(N-i)\cdot C_{N-i}. \tag{5.6}$$

Here $d_j = d^{(j)}(x_0)/j!$, $A_i = A^{(i)}(x_0)/i!$. In applications, in order to evaluate the power series expansion (5.3) or (5.5), we should look at recurrences satisfied by $C_N(x-x_0)^N$ instead of C_N. E.g., the recurrence for $C_N(x-x_0)^N$ has the form

$$C_{N+1}(x-x_0)^{N+1} = \{\sum_{i=0}^{\min\{N,d-1\}} \alpha_{i;N}\, C_{N-i}(x-x_0)^{N-i}\}/\delta_0, \quad (5.7)$$

where $\delta_0 = (N+1)\cdot d_0$, $\alpha_{i;N} = (x-x_0)\cdot[A_i(x-x_0)^i - d_{i+1}(N-i)(x-x_0)^i]$: $i = 0,\ldots,\min\{N,d-1\}$. This recurrence (like (5.4) or (5.6)) can be written in the matrix form. We write this matrix recurrence that computes simultaneously d consecutive coefficients $C_{N-i}(x-x_0)^{N-i}$: $i = 0,\ldots,d-1$ ($i \le \min\{N,d-1\}$), and simultaneously, the N-th partial sum $Y_N(x) = \Sigma_{i=0}^{N-1} C_i(x-x_0)^i$ of $Y(x)$. This new recurrence follows from (5.7), if to add one more formula:

$$Y_{N+1}(x) = Y_N(x) + C_N(x-x_0)^N. \quad (5.8)$$

To represent the matrix recurrence, we introduce a $n(d+1) \times n$ matrix
$\mathcal{Y}_N = (Y_n, C_N\cdot(x-x_0)^N,\ldots,C_{N-(d-1)}\cdot(x-x_0)^{N-d+1,t}$ From (5.7)-(5.8) we deduce a matrix recurrence:

$$\mathcal{Y}_{N+1} = \mathfrak{a}(N)\cdot\mathcal{Y}_N,$$

where $\mathfrak{a}(N)$ is a $n(d+1) \times n(d+1)$ matrix consisting of blocks of $n \times n$ matrices: $\mathfrak{a}(N) = (B_{ij}(N))_{i,j=1}^{d+1}$. Here $B_{1,1}(N) = 1$, $B_{1,2}(N) = 1$, $B_{i,j}(N) = 0$ for $j = 3,\ldots,d+1$; $B_{k,\ell}(N) = \delta_{k-1,\ell}$ for $k \ge 3$, $B_{2,1}(N) = 0$, and

$$B_{2,j}(N) = \alpha_{j-2;N}/\delta_0: j = 2,\ldots,d+1, \quad (5.10)$$

in the notations of (5.7). Similarly, in the case of equation (5.1), we get a matrix recurrence

$$\mathcal{Y}_{N+1} = \mathfrak{a}(N)\cdot\mathcal{Y}_N,$$

for $(d+1)\times(d+1)$ matrix $\mathfrak{a}(N)$ and $\mathcal{Y}_N = (y_N(x), y_N(\alpha)\cdot(x-x_0)^N,\ldots,y_{N-d+1}(\alpha)(x-x_0)^{N-d+1})$, for $\mathcal{Y}_N(x) = (x-x_0)^\alpha\cdot\Sigma_{i=0}^{N-1} y_i(\alpha)\cdot(x-x_0)^i$.

Let us determine now the computational cost of deriving the matrix recurrence (5.9) from the original linear differential equations (5.1) and (5.2), and the dependence of $\mathfrak{a}(N)$ in (5.9) on N, x_0 and x. To derive (5.9) we need to know the original coefficients of (5.1) or (5.2) and to know α or W. In all important cases α is a fixed rational number (and W is a diagonal matrix with rational number entries). To compute coefficients in (5.4) or (5.6) one needs to determine the coefficients of the translated polynomials in (5.1) or (5.2) after the translation $x \to x + x_0$. Thus the total number of operations to compute the coefficients matrix $\mathfrak{a}(N)$ in (5.9) is $O(n^2 d\log d)$ in the case of (5.2), and is $O(n\, d\log d)$ in the case of (5.1) (with $d \ge n$ in the case of (5.1)). The matrix $\mathfrak{a}(N)$ in (5.9) can be represented as $\mathfrak{a}(N) = \mathfrak{a}_0(N)/d_0(N)$, where $\mathfrak{a}_0(N)$ and $d_0(N)$ are polynomial in N, x_0 and X. In the case of (5.2) (with $W = 0$), $\mathfrak{a}(N)$ is an $n(d+1) \times n(d+1)$ matrix, with $\mathfrak{a}_0(N)$, $d_0(N)$ linear in N, polynomial in x_0 and $x - x_0$ of degrees at most d in x_0, and with $\mathfrak{a}_0(N)$ of degree d in $(x-x_0)$ ($d_0(N)$ is independent of x). More precisely, $\mathfrak{a}_0(N)$, as a polynomial in $x - x_0$ and x_0 has a total degree of at most d. In the case of the equation (5.1), $\mathfrak{a}_0(N)$ and $d_0(N)$ are polynomial of degree n in N; they are polynomials in $x - x_0$ and x_0. $d_0(N)$ is polynomial in x_0 of degree at most d, and $\mathfrak{a}_0(N)$ is polynomial in $x - x_0$ and x_0 of total degree at most d.

The growth of coefficients in the power series expansions $y(x)$ or $Y(x)$ can be estimated e.g. from the recurrences (5.2) or (5.4) (or (5.9)) according to the Poincare-Perron theorem (see Perron [57]), which states that the asymptotics of solutions of recurrences are determined by roots of the limit characteristic equation of a linear recurrence with constant coefficients that one deduces from (5.2) or (5.4) as $N \to \infty$. An explicit form of (5.2) or (5.4) shows that these roots of the limit characteristic equation are $1/(s-x_0)$ for $s \in S$ (the set of singularities of (5.1) or (5.2)), when x_0 is a regular point. More precisely, an asymptotic analysis of recurrences (5.2) or (5.4) (or (5.9)) gives the following leading term of the asymptotics of coefficients $y_N(\alpha)$ or C_N:

$$|y_N(\alpha)| \leq \gamma_1 \cdot N^{\nu} . \mathrm{dist}^*(x_0,S)^{-N}, \tag{5.10}$$

$$\|C_N\| \leq \gamma_2 \cdot N^{\mu} . \mathrm{dist}^*(x_0,S)^{-N}. \tag{5.11}$$

Here $\mathrm{dist}^*(x_0,S) = \min\{|s-x_0|: s \in S, x_0 \neq S\}$, and $\|.\|$ is a c_0-norm of $n \times n$ matrices. Asymptotic bounds (5.10) or (5.11) hold for arbitrary equations (5.1) or (5.2) with any initial conditions at regular (singularity) $x = x_0$. In general (i.e. for generic equations (5.1) or (5.2) or for generic initial conditions), one can complement (5.10) and (5.11) with a similar lower bound for $N \geq N_0$, but with different constants γ_1 and γ_2. Bounds (5.10) and (5.11) show that whenever $\delta > 0$, and $|x_1-x_0|/\mathrm{dist}^*(x_0,S) < 1-\delta$, one obtains the value of $y(x_1)$ or $Y(x_1)$ from the matrix $\mathcal{Y}_N$ at $x = x_1$ because $y_N(x_1)$ or $Y_N(x_1)$ converges to $y(x_1)$ or $Y(x_1)$, respectively, as a geometric progression in N. To evaluate $y(x_1)$ or $Y(x_1)$ with the precision ℓ one needs the value of $\mathcal{Y}_N$ with the precision $\ell + O(\log \ell)$ for $N = O(-\ell \log(\frac{|x_1-x_0|}{\mathrm{dist}^*(x_0,S)}))$, if $|x_1-x_0|/\mathrm{dist}^*(x_0,S) < 1-\delta$ (for fixed $\delta > 0$).

Now we can touch upon the problem of analytic continuation studied in detail in [12]. The key is the standard superposition formula that expresses the solution $Y(x;x_0)$ of (5.2) normalized at $x = x_0$: $Y(x;x_0)|_{x=x_0} = I_n$ has the form of chain rule:

$$Y(x;x_0) = Y(x;x_1) \cdot Y(x_1;x_0) \tag{5.12}$$

for any three points x_0,x_1,x in $\mathbb{CP}^1$. This superposition formula leads to the chain rule of evaluation of an arbitrary solution $Y(x)$ with the initial conditions $Y(x)|_{x=x_0} = A$ at $x = x_0$ at a point x_{fin}, which is the end-point of a path $\overline{x_0x_1 \ldots x_m x_{m+1}}$ in $\mathbb{CP}^1$ with $m+2$ vertices $x_0,x_1,\ldots,x_m,x_{m+1} = x_{fin}$:

$$Y(x)|_{x=x_{fin}} = Y(x_{m+1};x_m) \ldots Y(x_2;x_1)Y(x_1;x_0) \cdot A. \tag{5.13}$$

A rule (5.13) is unambiguous, whenever x_{i+1} lies within the disc of convergence of $Y(x;x_i)$, e.g. whenever for a fixed $\delta > 0$, $|x_{i+1}-x_i|/\mathrm{dist}^*(x_i,S) < 1-\delta$.

Let us estimate now the bit complexity of computing of $Y(x_1;x_0)$ using the matrix recurrence (5.9). As above, we assume that for a fixed $\delta > 0$, $|x_1-x_0|/\mathrm{dist}^*(x_0,S) < 1-\delta$. As it was shown above, to compute the $n \times n$ matrix with precision

$\ell + O(\log \ell)$ one can compute $\mathcal{Y}_N$ for $N = O(\ell/\log(|x_1-x_0|/\mathrm{dist}^*(x_0,S)))$ from (5.9) with initial conditions $Y(x)|_{x=x_0} = I_n$. (I.e. $C_0 = I_n$.)

To estimate the bit complexity of computations of $\mathfrak{a}_N$ one needs to know the amount of bit information in the representation of x_0 and x_1. We can represent x_0 and x_1 as rational numbers p/q of logarithmic size$(p,q) = \log(\max\{1,|p|,|q|\})$, or as b-bit binary (floating point) numbers $.\underbrace{b_1b_2b_3 \ldots b_b}_{b} \times 2^{EXP}$ (for $EXP = O(b)$). The matrix of coefficients $\mathfrak{a}(N)$ in (5.9) depends polynomially on x_0 and $x_1 - x_0$. Convenient parameters can be x_0 and $x_1 - x_0$ or $\frac{x_1-x_0}{x_0}$ (if $x_0 \neq 0$). Let us look at x_0, $x_1 - x_0$ written with a common denominator: $x_0 = \frac{X_0}{D}$, $x_1-x_0 = \frac{X_{10}}{D}$ (e.g. $D = 2^b$ in the binary representation). We define then $b = \log_2\max\{1,|D|,|X_0|,|X_{10}|\}$ as the (logarithmic) size of x_0, x_1-x_0. Then $\mathfrak{a}(N) = \mathfrak{a}_0(N)/d_0(N)$ where $\mathfrak{a}_0(N), d_0(N)$ are linear in N with coefficients that are $O(d\cdot b)$-bit integers. Let $e = -\log_2(|x_1-x_0)/\mathrm{dist}^*(x_0,S)) > 0$ be the measure of closeness of x_1 to x_0. Then we need to compute $\mathcal{Y}_N$ for $N = O(\ell/e)$ with the precision $\ell + O(\log\ell)$.

To compute $\mathcal{Y}_N$ from the matrix recurrence (5.9) we use the fast binary-splitting technique of Theorem 6.1. This way we arrive at the following local evaluation of the solution $Y(x)$ of (5.2), where we specify all the dependencies on the sizes of coefficients involved.

Theorem 5.1: Let (5.2) be a fixed matrix linear differential equation, where all elements of $A(x)$ are rational functions of

the total degree at most d, and coefficients of these functions are rational numbers of sizes at most k (or arbitrary complex numbers represented by their binary approximations as k-bit binary floats). Let $x = x_0$ be a regular or regular singular point of (5.2) represented by a rational number of (logarithmic) size of at most b, or by a b-bit binary floating point number. Let $Y(x)$ be a solution of (5.2) with fixed initial conditions at $x = x_0$, and let x_1 be another binary b-bit number, lying within the disc of convergence of $Y(x)$. Then to compute $Y(x)$ at $x = x_1$ with precision of ℓ leading digits one needs at most

$$c_1 \cdot M_{atrix}(n) \cdot \frac{db+k}{e} \cdot M_{bit}(\ell) \cdot \log^2 \ell + O(\ell)$$

bit-operations. Here $e = -\log_2(|x_1-x_0|/\mathrm{dist}^*(x_0,S)) > 0$, $M_{atrix}(n)$ is the number of operations needed for $n \times n$ matrix multiplications, $M_{bit}(\ell)$ $(= \ell\cdot\log(\ell)\cdot\log\log(\ell))$ is the number of operations for ℓ-bit multiplication, and c_1 is an absolute constant. The constant in $O(\ell)$ depends on initial conditions of $Y(x)$.

Theorem 5.1 is the basis of bit-burst method of fast evaluation of any solution of a linear differential equation at any point on its Riemann surface. According to this method the evaluation of any branch of any solution with the precision of ℓ leading bits (digits) requires at worst only $O(M_{bit}(\ell)\cdot\log^3\ell\cdot(1+o(\ell))$ total bit operations.

Starting from Theorem 5.1 and using the chain rule (5.13) of analytic continuation along any fixed path, we can evaluate $Y(x)$ (the solution of (5.2) with given initial conditions) everywhere on its Riemann surface. First, the recipes for the optimal choice of polygon $\overline{x_0x_1 \ldots x_mx_{m+1}}$ homotopic to a given path γ in $\mathbb{CP}^1\backslash S$ are presented in [12], so as to minimize the total number of operations of evaluations of $Y(x)$ at intermediate points x_i. Second, we are using Theorem 5.1 in a clever way to evaluate $Y(x)$ from $x = x_0$ to $x = x_1$ making several steps between x_0 and x_1 (again using the chain rule (5.13)), releasing consecutive blocks of x_1 in bursts. We call this method "bit-burst" method. In this approach one, in order to evaluate $Y(x)$ at a bigfloat $x = \Sigma_{N=0}^{\infty} b_N \cdot B^{-N}$, evaluates $Y(x)$ consecutively at $x_i = x_i + \Sigma_{N=2^i}^{2^{i+1}} b_N \cdot B^{-N}$ by analytic continuation of $Y(x)$ (rule (5.13)): "adding more bits of a number, but computing with the same accuracy". In this approach we are matching the distance from the evaluation point to the point of expansion with the size of the point of evaluation. In particular, we move initial conditions from $x = x_0$ to a nearby point $x = x_0'$ of bounded size, and then evaluate $Y(x)$ at $x = x_1$ starting from $x = x_0'$ in bit-bursts. This way we arrive at the following general theorem that gives an upper bound on the bit complexity:

<u>Theorem 5.2</u>: Let (5.2) be a given linear differential equation with rational function coefficients, and $Y(x)$ be its arbitrary (regular) solution with initial conditions at

$x = x_{in}$, where x_{in} is an K-bit number. Given a path γ from x_{in} to an K-bit number x_{fin} (on the Riemann surface of $Y(x)$) of length L, one can evaluate $Y(x)|_{x=x_{fin}}$ at $x = x_{fin}$ with the full K-bit precision one needs at most

$$O(M_{bit}(K)(\log^3 K + \log L))$$

bit operations.

The bit-burst method in the general form of Theorem 5.2 should be improved: one would like to see $\log^3 K$ replaced by $\log K$ always. We don't know how to do it in general, but sometimes the complexity can be lowered:

I. If x_{in} and x_{fin} are fixed rational numbers (or are given as big floats with $o(K)$ bits), then the computation of $Y(x)$ with the <u>full K bits of precision</u> has bit complexity at most $O(M_{bit}(K) \cdot (\log^2 K + \log L))$. (This is the case of computation of classical constants.)

II. If the differential equation (5.1) or (5.2) possesses special arithmetic properties, bit-complexity can be lowered. E.g. if the equation (5.1) is globally nilpotent or (5.1)-(5.2) possesses a solution which is either an E-function or a G-function, then the general bit bound of Theorem 5.2 can be lowered to

$$O(M_{bit}(K) \cdot (\log^2 K + \log L)).$$

If, further, like in I, x_{in} and x_{fin} are given only by a few bits (i.e. are fixed rational numbers), and $Y(x)$ is built from

E-or G-functions, then K significant digits of $Y(x_0)$ can be computed in

$$O(M_{bit}(K)\cdot(\log K + \log L))$$

bit operations.

This bound is unsurpassed by any other algorithm even for elementary functions, like the exponent, where low operational complexity algorithms are well known [5], [54-55].

Remark: The bit-burst method should not be confused with the popular bit-serial (or bit-by-bit) method, where each new bit of the evaluated function is added for a new input bit of the value. Instead, we output the full precision of evaluated function for a bit burst of an input value, and the number of input bits increases geometrically. Note a significant difference with the Newton method of computations, where input bits are also introduced in bursts, but at each step only an appropriate precision of evaluation is needed, thus making the algorithm self-correcting. Of course, the Newton method is efficient only for computation of algebraic functions. The bit burst method is particularly fast when the result is needed with full precision, but the amount of input bits is limited. This is the case, e.g. of computations of solutions of linear differential equations at fixed rational points, or of computations of invariants of differential equations such as monodromy group.

6. Fast solution of matrix difference equations and parallel algorithms.

As we saw in the proof of Theorem 5.1 the main component in the fast evaluation of solutions of differential equations using the power series method is the solution of a difference equation with rational function coefficients. Instead of the recurrence (5.9) we look at the matrix difference equation written as follows

$$A_{N+1} = C(N)\cdot A_N \tag{6.1}$$

where $C(N)$ is $n \times n$ matrix rationally dependent on N. The solution A_N of (6.1) with initial conditions $A_N|_{N=0} = A_0$ has a symbolic representation

$$A_N = \overleftarrow{\prod_{i=0}^{N-1}} C(i)\cdot A_0, \text{ or}$$
$$A_N = C(N-1)\cdot C(N-2)\cdot\ldots\cdot C(1)\cdot C(0)\cdot A_0, \tag{6.2}$$

i.e. the order of the terms in the product is reversed. The fast method of computation of A_N in (6.2) is known as the binary-splitting method (or divide and conquer method). This is a known technique to accelerate the solution of linear recurrences, and also a well known instrument in applications of the Chinese remainder theorem and interpolation. In this method a binary tree (when $N = 2^k$; otherwise more complicated trees associated with addition chain methods are used) of operations is constructed so that multiplication of terms in (6.2) proceeds in the way that operands have slow growing sizes. (This method is opposite to the obvious method of computation of (6.2) with consecutive multiplications by $C(i)$).

With notations of (6.2) we introduce the following auxiliary variables

$$\mathcal{G}_{L;K} = \overleftarrow{\prod_{j=K}^{L-1}} C(j) \qquad (6.3)$$

for $L > K$. We put $\mathcal{G}_{K;K} = I_n$. In these notations A_N is

$$A_N = \mathcal{G}_{N;0} \cdot A_0. \qquad (6.4)$$

There is a simple chain rule of computations of $\mathcal{G}_{L;K}$ which is the basis of any splitting method including the binary splitting method:

$$\mathcal{G}_{L;K} = \mathcal{G}_{L;M} \cdot \mathcal{G}_{M;K} \quad \text{for} \quad L \geq M \geq K. \qquad (6.5)$$

The chain rule (6.5) provides us with an immediate algorithm of computations of $\mathcal{G}_{L;K}$ from the binary expansion of L and K. This method is at its best when N is a power of 2, $N = 2^k$.

The algorithm consists of the outer loop over all ℓ from 0 to k and the inner loop over all $(k-\ell)$-bit (binary) integers. We start at $\ell = 0$ with the initialization:

$$\mathcal{G}_{K+1;K} = C(K) \text{ for all } K \text{ in } K = 0,\ldots,2^k-1 \qquad (6.6)$$

At the ℓ-th step we have determined all $\mathcal{G}_{2^\ell(K+1);2^\ell \cdot K}$ for all $(k-\ell)$-bit integers K: $0 \leq K \leq 2^{\ell-k} - 1$. At the step $\ell + 1$ we use the rule (6.5), and obtain:

$$\mathcal{G}_{2^{\ell+1}(K+1);2^{\ell+1}K} = \mathcal{G}_{2^\ell(2K+2);2^\ell(2K+1)} \cdot \mathcal{G}_{2^\ell(2K+1);2^{\ell+1}\cdot K} \qquad (6.7)$$

for $0 \leq K \leq 2^{k-\ell-1} - 1$, i.e. $\mathcal{G}_{2^{\ell+1}(K+1);2^{\ell+1}\cdot K}$ are determined for all $(k-\ell-1)$-bit integers.

Finally, at step k at $\ell = k$, according to (6.4), A_N is determined as: $A_N = \mathcal{G}_{2^k\cdot(K+1);2^k\cdot K} = \mathcal{G}_{2^k;0}$ at $K = 0$.

For $\ell = 1,\ldots,k$, at ℓ-th step of this algorithm we perform in (6.7) $2^{k-\ell}$ matrix multiplications. Binary-splitting method is efficient if the computations are conducted with increased precision, i.e. all bits of information in computations of A_N via (6.1)-(6.2) are preserved. In the typical case, when $C(N)$ is rational in N, the total (memory) space needed to hold all bits of A_N (as a rational number, i.e. numerator and denominator) is $O(n^2 \cdot N \cdot \log_2 N)$. [Only in special cases, when the recurrence (6.1) represents a recurrence associated with a globally nilpotent Fuchsian linear differential equation the memory requirements are $O(N)$.] It is important to notice that the binary-splitting method of computation of A_N in (6.6)-(6.7) requires about the same amount of memory. Indeed, to compute the ℓ-th step of the algorithm only the previous step is needed with a total amount of memory space of $O(N \log_2 N)$.

The total amount of operations depends on the cost of the multiplication of ℓ-bit numbers $M_{bit}(\ell)$. Under the assumptions above, that $C(N)$ is a rational function of N, let

$$C(N) = \frac{C_0(N)}{d(N)} \tag{6.8}$$

where $C_0(N) \in M_{n\times n}(\mathbb{Z}[N]), d(N) \in \mathbb{Z}[N]$. Let in (6.8) d be the maximal degree of polynomials in $C_0(N)$ and $d(N)$. With (6.8) substituted in (6.2) the solution A_N of (6.1) can be represented as

$$A_N = \frac{\overleftarrow{\prod_{i=0}^{N-1}} C_0(i)}{\prod_{i=0}^{N-1} d(i)} A_0. \tag{6.9}$$

We compute $\overleftarrow{\prod}_{i=0}^{N-1} C_0(i)$ using the binary-splitting algorithm of (6.6)-(6.7) for $C_0(N)$ instead of $C(N)$. Let h be the maximum of sizes of coefficients of polynomials in $C_0(N)$ and $d(N)$. Then, as it follows from the iterative scheme (6.7), all $G_{2^\ell\cdot(K+1);2^\ell\cdot K}$ for $0 \le K \le 2^{k-\ell} - 1$ are integers of sizes bounded by $O((d+h)\cdot 2^\ell\cdot \log_2 N)$ (for $\ell \le k$). [The constant under $O(.)$ depends only on n and logarithmically so; by choosing instead of h the maximum of sizes of polynomials in $C_0(N)$ and different norm of the matrix, the dependence on n can be removed.] Consequently, at every step from $\ell = 1$ to k the total number of bit operations is bounded by $2^{\ell-k} \times M_{bit}(2^\ell \log_2 N)\cdot(d+h)$. The total number of bit-operations (with $\log_2 N = k$) is $O(M_{matrix}(n)\cdot(d+h)\cdot M_{bit}(N)\cdot \log^2 N)$. This

gives the numerator in (6.9). The denominator is computed the same way with the total number of operations again bounded by $O((d+h)\cdot M_{bit}(N)\cdot \log^2 N)$. [If necessary, the denominator can be computed faster using the distribution of prime ideals in the Galois group of polynomial $d(N)$. The bit complexity becomes $O(M_{bit}(N).\log N)$.]

When $N = 2^k$ we use binary-splitting method as presented above; if N is arbitrary the corresponding addition chain tree (used for fast computation of N in $O(\log N)$ additions only) is applied.

<u>Theorem 6.1</u>: Let in the recurrence (6.1), $C(\cdot)$ be $n \times n$ matrix whose rational function entries have sizes and degrees bounded by s. Then the bit complexity of computations of A_N in (6.2) is bounded by

$$O(M_{bit}(N)\cdot M_{atrix}(n)\cdot s\cdot \log^2 N),$$

where $M_{atrix}(n)$ is the operational complexity of $n \times n$ matrix multiplication.

(Only the term $\log^2 N$ is not the best possible. Apparently, under additional arithmetic assumptions it can be improved. This is the case when all A_N are "nearly integral," i.e. the generating function $Y(x) = \Sigma_{N=0}^{\infty} A_N x^N$ is a G-function. In this case, and also in the case when $Y(x)$ is a E-function, the $\log^2 N$ term in (6.2) can be replaced by

$$\log N\cdot(\log\log N)^{1+\epsilon}.)$$

Our best results on the upper bounds of the operational complexities of computations of solutions of the recurrence (6.1) are not as good as in Theorem 6.1.

Since in operational complexity count the sizes of operands are irrelevant, binary-splitting methods or a trivial method of computation of (6.2) give the same bound: $O(N\cdot(n^2 . d + n^{\mu}))$, where $\mu < 2.5$ is the exponent in the matrix multiplication problem.

The "binary splitting" method:

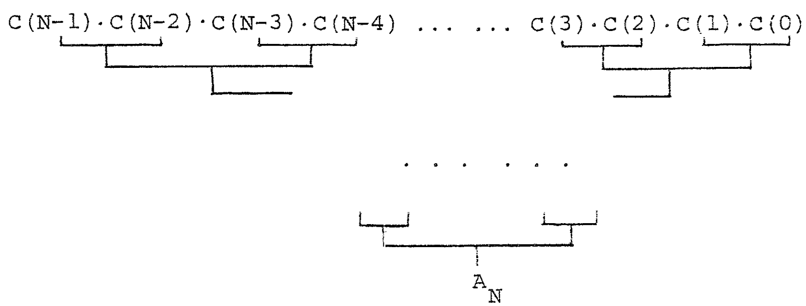

Fig. 2

The linear dependence on N is unsatisfactory. To speed up the computation of (6.2) we use a linear acceleration instead of geometric one. We start with introduction of some auxiliary cost functions. The operational complexity (cost) of multiplication of polynomials of degree d is $M_{poly}(d) = O(d\log d)$; the cost of evaluation of a polynomial $p(x)$ of degree d at $d(+1)$ consecutive points: $0,1,\ldots,d-1(d)$--is $M_{eval}(d) = O(d \log^2 d)$, with $O(d\log d)$ essential multiplications. (There is, though, a simple algorithm that

requires $O(d^2)$ additions.) Similarly $M_{shift}(d)$ is the cost of a shift of a polynomial $p(x)$ to $p(x+h)$ for $p(x)$ of degree d.

Let us look at new recurrences following from (6.1)

$$A_{N+k} = C_k(N)\cdot A_N, \tag{6.10}$$

where for any $k \geq 1$

$$C_k(N) = C(N) \cdots C(N+k-1). \tag{6.11}$$

Let us consider, as above, $N = 2^a$ (again similar arguments apply to an arbitrary N):

$$\begin{aligned} A_{N+2^c} &= C_{2^c}(N)\cdot A_N, \\ C_{2^c}(N) &= C_{2^{c-1}}(N+2^{c-1})\cdot C_{2^{c-1}}(N). \end{aligned} \tag{6.12}$$

To compute $C_{2k}(N) = C_k(N+k)\cdot C_k(N)$ from $C_k(N)$, as a matrix with entrees rational in N, one needs: a) to compute the coefficients of polynomial/rational expansion of $C_k(N+k)$ in (powers of) N--it takes $O(n\, M_{shift}(d_k))$ operations: b) to multiply the rational/polynomial entries of $C_k(N+k)$ and $C_k(N)$--this takes $O(M_{atrix}(n) \cdot M_{pol}(d_k))$. Here d_k is the maximum of the degrees of elements of $C_k(N)$.

Now we remark that the scheme (6.10) for a given k is nothing but a version of a general scheme (6.1), but: (i) with a different matrix C--the size of it is the same, but the degrees of its polynomial/rational function entries are different; (ii) the length of the recurrence is k times shorter. The maximum, d_k, of degrees of rational/polynomial entries of

$C_k(N)$ is $\leq kd$; say, for simplicity, $d_k = kd$.

Let us estimate now the total cost of the trivial algorithm of computation of (6.2) in the case of recurrence (6.10) (or a binary-splitting version of this algorithm). To evaluate $A_N = A_{t\cdot k}$ in t steps, starting from A_0 following the rule (6.10) one needs: α) consecutive evaluation of $C_k(0),\ldots,C_k((t-1)\cdot k)$ at $t = N/k$ points. Once the rational function $C_k(\cdot)$ is known, the total operational complexity of α) is $O(n^2\cdot M_{eval}(d_k)\cdot[\frac{N}{kd_k}+1])$. Secondly, one needs β) t consecutive matrix multiplications $C_k(j\cdot k)\cdot A_{j\cdot k}$ for $j = 0,\ldots,t-1$ with the total cost of $O(M_{atrix}(n)\cdot t)$.

The cost of determination of $C_k(\cdot)$ for $k = 2^c$, as follows from the discussion above, is $\Sigma_{b=0}^{c-1}\{O(M_{atrix}(n)\cdot M_{pol}(2^b\cdot d)) + O(n^2\cdot M_{shift}(2^b\cdot d))\}$, i.e. the cost of determination of $C_k(\cdot)$ is $O(M_{atrix}(n)\cdot M_{pol}(kd) + n^2 M_{shift}(kd))$. This result is true not only for $k = 2^c$, but for any k.

Thus, whenever $N \geq k^2 d$ (so that $[N/kd_k] \geq 1$), the total cost of computation of A_N using from the scheme (6.10) is:

$$O(M_{atrix}(n)\cdot\{M_{pol}(kd)+\frac{N}{k}\} + n^2\{M_{shift}(kd)+M_{eval}(kd)\cdot\frac{N}{k^2 d}\}).$$

Substituting $M_{pol}(x) = O(x\log x)$, $M_{shift}(x) = O(x\log x)$, $M_{eval}(x) = O(x\log^2 x)$ we can derive the following upper bound on the total cost of computations of A_N using (6.10) with, say, $k = [\sqrt{N/d}] + o(\sqrt{N})$:

$$O(M_{atrix}(n)\sqrt{Nd}\log N + n^2\sqrt{Nd}\log^2 N)$$

for $N/d \gg 1$. This bound can be slightly improved, also the cost of a single multiplication has to be multiplied by the corresponding weight (if long numbers are involved). We arrive at the following result:

<u>Theorem 6.2</u>: Let in the recurrence

$$A_{N+1} = C(N)\cdot A_N,$$

$C(\cdot)$ be $n\times n$ matrix with rational coefficients with degrees bounded by d. Then the operational complexity (total number of operations) is bounded by

$$O(M_{atrix}(n)\cdot\sqrt{Nd}\log N + n^2\sqrt{Nd}\log^2 N),$$

with at most $O(M_{atrix}(n)\cdot\sqrt{Nd}\log N)$ multiplications.

A far cry from Theorem 6.1 is $O(\sqrt{N}\ \ldots)$ number of operations. Can it be reduced in general? Consider an example:

<u>Example</u>: Let us look on operations in modular arithmetic mod M for a fixed M. Then the cost of an operation is bounded by $O(M_{bit}(\log M)) = O(\log M\cdot\log\log M\cdot\log\log\log M)$.

To factor M one has only to compute in modular arithmetic $N!\ (\mathrm{mod}\ M)$ for $N = [\sqrt{M}]$, i.e. a solution of a simple recurrence:

$$A_{N+1} = N\cdot A_N \bmod M,$$

and then check for g.c.d. of $N!\,(\mathrm{mod}\ M)$ and M.

Theorem 6.2 implies that we got a simple method of factorizing M in $O(\sqrt[4]{M}.\log^2 M)$ bit operations. This is asymptotically equivalent to two popular factorization methods: Pollard's ρ-method and Shank's quadratic form composition method. Our method, though, is completely deterministic.

Unlike computations in modular arithmetic, where we have difficulties bounding complexity (as $N \to \infty$) of computation of solutions of (6.1), in the case of fixed precision computations the situation is different. In this case we can use asymptotic expansions of solutions of (6.1) in powers of N near ∞. Using these Borel (Birkhoff's) expansions we get the following bound:

Proposition 6.3: The matrix solution A_N of (6.1) can be computed within the precision M of leading digits in

$$O(\log^3(MN)\cdot M_{bit}(M))$$

bit (boolean) operations.

To prove Proposition 6.3 we use $1/N$ expansions of solutions of difference equations, and asymptotic (inverse factorial) series representation of A_N for large N. In modular arithmetic mod p we do not have similar expansions and the problem of low complexity computations of (6.2) is an open one, cf.[12].

Proposition 6.3 is interesting because many constants

that arise as values of functions not satisfying linear differential equations are well approximated by constants that are values of functions satisfying difference and differential equations. (Among such constants we can name values of Γ-function, Barnes Γ-functions, values of ζ- and L-functions associated with algebraic number fields and modular forms.)

How about nonlinear differential equations? One would like to hope that in this case, as for linear differential equations, the bit-complexity of multiprecision computations is close to the optimal one, perhaps differing by a power of $\log n$ only. Unfortunately we cannot in general prove such tight bounds because the Galois group (a rather sophisticated object in the nonlinear case) is more difficult to compute.

(For nonlinear equations with some structure, like generalized Ricatti equations, Painleve-type equations,... etc. low complexity algorithms can be constructed.)

The best bound that we have so far for n-bit precision computations of solutions of nonlinear differential equations is

$$O(\sqrt{n}\ \log^2 n\cdot M_{bit}(n)).$$

As far as massive parallelism is concerned, the full n-bit precision computations of values of nonlinear differential equations (whose right hand side is built from "known" function, that can be themselves solutions of differential equations) can be completed in only $O(\log n)$ steps with the total circuitry of $n^{O(1)}$ processors.

How efficient are these methods practically for moderate values of n compared with the standard finite-difference (Runge-Kutta,... etc.) methods of solution?

One of the advantages of the proposed methods is the low storage requirements even in the case when solutions have to be tabulated and not evaluated at a single point. Indeed, our methods determine not only a value of a function, but provide with good approximating polynomials and rational functions that can be used in large domains. Tabulation of these objects requires much less computational effort, when precision n is fixed.

Let us look now at parallel versions of these algorithms. For computations in the parallel environment one has to balance the following parameters: a) the number of parallel steps (the depth of the circuit); b) the total number of bit (boolean) operations; and c) the number of (micro) processors. For the moment we put aside important consideration of: d) communication network; 3) local-global storage and retrieval requirements; and, in the case of VLSI designs, that of f) area-time complexity.

For a problem with I or O of $O(n)$ bits, e.g. in the evaluation with the full n-bit precision of the value of an algebraic or transcendental function, one can hope for the depth of $O(\log n)$. With this depth one would like also to have parallel algorithms of total circuitry (number of processors) to be optimal as well: on the order of

$O(M_{bit}(n)\cdot\log n)$.

There are serious obstacles in generation of such fast parallel algorithms from the serial algorithms with the same bit complexity. Iterative algorithms, that are used to construct low bit (and operational) complexity serial algorithms, do not provide depth $O(\log n)$. These iterative algorithms, particularly the arithmetic-geometric mean, and also the Newton iteration method, can get the depth at best $O(\log^2 n)$.

The problem actually starts with computation of elementary functions (most notably that of division). Popular iterative methods (Sieveking-Kung,...) give depth $O(\log^2 n)$, though with only $O(M_{bit}(n))$ of total circuitry for elementary function computations. Only recently new algorithms of Reif, Bini, Schonhage and Pan, see review [58] for references, made it possible to compute elementary functions in $O(\log n)$ (or $O(\log n\cdot\log\log n)$ in Reif and Beame-Cook-Hoover algorithms) depth circuits. Unfortunately, the total circuitry significantly increases: the typical number of processor becomes $n^{O(1)}$, see [58].

The methods that we propose, bit-burst algorithms for computations of arbitrary linear differential equations, have always depth $O(\log n)$, even though, in general, the total bit operation count is $O(M_{bit}(n)\cdot\log^3 n)$. In fact, for E- and G-functions the total bit operation count with the same depth is only $O(M_{bit}(n)\cdot\log^2 n)$.

References

[1] S. Ramanujan, Collected Papers, Cambridge, 1927, 23-39.

[2] G.H. Hardy, Ramanujan, Cambridge, 1940.

[3] G.N. Watson, Some singular moduli (I);(II);(III);(IV); Quart. J. Math. Oxford, 3 (1932), 81-98; 189-212; Proc. London Math. Soc. 40 (1936), 83-142; Acta Arithmetica, 1 (1936), 284-323.

[4] J.M. Borwein, P.B. Borwein, Pi and the AGM, Wiley, 1987.

[5] E.T. Whittaker, G.N. Watson, A Course of Modern Analysis, 4 ed., Cambridge, 1927.

[6] A. Weil, Elliptic Functions According to Eisenstein and Kronecker, Springer, 1976.

[7] C.L. Siegel, Bestimmung der elliptischen Modulfunktionen durch eine Transformations gleichung, Abh. Math. Sem. Univ. Hamburg, 27 (1964), 32-38.

[8] A. Weil, Sur les Périodes des Intégrales Abéliennes, Comm. Pure Appl. Math., 29 (1976), 813-819.

[9] D. Masser, Elliptic Functions and Transcendence, Lecture Notes Math., v. 437, Springer, 1975.

[10] S. Lefschetz, On certain numerical invariants of algebraic varieties with application to Abelian varieties, Trans. Amer. Math. Soc., 22 (1921), 327-482.

[11] E.T. Whittaker, On hyperlemniscate functions, a family of automorphic functions, J. London Math. Soc., 4 (1929), 274-278.

[12] D.V. Chudnovsky, G.V. Chudnovsky, Computer assisted number theory, Lecture Notes Math., Springer, 1240, 1987, 1-68.

[13] Ch. Hermite, Sur la Théorie des Équations Modulaires, C.R. Acad. Sci. Paris., 48 (1859), 940-1079-1097; 49 (1859), 16-110-141.

[14] H.M. Stark, Class-numbers of complex quadratic fields, Lecture Notes Math., Springer, v. 320, 1973, 153-174.

[15] D. Shanks, Dihedral quartic approximation and series for π, J. Number Theory, 14 (1982), 397-423.

[16] R. Fricke, Die Elliptischen Funktionen und ihre Anwendungen, v. 1, Teubner, 1916.

[17] A. Baker, Transcendental Number Theory, Cambridge, 1979.

[18] H.M. Stark, A transcendence theorem for class number problems, I; II; Ann. Math. 94 (1971), 153-173; 96(1972), 174-209.

[19] C.L. Siegel, Zum Beweise des Starkschen Satzes, Invent. Math., 5 (1968), 180-191.

[20] G. Shimura, Automorphic forms and the periods of Abelian varieties, J. Math. Soc. Japan, 31 (1979), 561-59.

[21] G. Shimura, The arithmetic of certain zeta functions and automorphic forms on orthogonal groups, Ann. of Math., 111 (1980), 313-375.

[22] P. Deligne, Valeurs de fonctions L et périodes d'integrales, Proc. Symp. Pure Math., v. 33, Part 2, Amer. Math. Soc., Providence, R.I., 313-346.

[23] P. Deligne, Cycles de Hodge absolus et périodes des intégrales des variétés abéliennes, Soc. Math. de France, Memoire, N° 2, 1980, 23-33.

[24] G. Shimura, Introduction to the Arithmetic Theory of Automorphic Forms, Princeton, University Press, 1971.

[25] G. Shimura, Y. Taniyama, Complex Multiplication of Abelian Varieties and Its Applications to Number Theory, Publications of the Mathematical Society of Japan, N° 6, 1961.

[26] G.V. Chudnovsky, Algebraic independence of values of exponential and elliptic functions. Proceedings of the International Congress of Mathematicians, Helsinki, 1979, Acad. Sci. Tennice, Helsinki, 1980, v. 1, 339-350.

[27] B.H. Gross, N. Koblitz, Gauss sums and the p-adic Γ-function, Ann. of Math., 109 (1979), 569-581.

[28] R. Fricke, F. Klein, Vorlesungen über die Theorie der Automorphen Functionen, bd. 2, Tenbner, 1926.

[29] R. Morris, On the automorphic functions of the group $(0,3;\ell_1,\ell_2,\ell_3)$, Trans. Amer. Math. Soc., 7 (1906), 425-448.

[30] K. Takeuchi, Arithmetic triangle groups, J. Math. Soc. Japan, 29 (1977), 91-106.

[31] H.P.F. Swinnerton-Dyer, Arithmetic groups in Discrete Groups and Automorphic Functions, Academic Press, 1977, 377-401.

[32] J.I. Hutchinson, On the automorphic functions of the group (0,3;2,6,6), Trans. Amer. Math. Soc., 5 (1904), 447-460.

[33] D.V. Chudnovsky, G.V. Chudnovsky, Note on Eisenstein's system of differential equations, in Classical and Quantum Models and Arithmetic Problems, M. Dekker, 1984, 99-116.

[34] D.V. Chudnovsky, G.V. Chudnovsky, The use of computer algebra for diophantine and differential equations, in Computer Algebra as a Tool for Research in Mathematics and Physics, Proceeding of the International Conference, 1984, New York University, M. Dekker, 1987.

[35] D.V. Chudnovsky, G.V. Chudnovsky, Applications of Padé approximations to diophantine inequalities in values of G-functions, Lecture Notes Math., v. 1135, Springer, 1985, 9-50.

[36] K. Mahler, Perfect systems, Compositio Math., 19 (1968), 95-166.

[37] D.V. Chudnovsky, G.V. Chudnovsky, Applications of Padé approximations to the Grothendieck conjecture on linear differential equations, Lecture Notes Math., v. 1135, Springer, 1985, 51-100.

[38] C.L. Siegel, Über einige Anwendungen diophantischer Approximationen, Abh. Preuss. Akad. Wiss. Phys. Math. Kl. 1, 1929.

[39] G.V. Chudnovsky, Contributions to the Theory of Transcendental Numbers, Mathematical Surveys and Monographs, v. 19, Amer. Math. Soc., Providence, R.I., 1984.

[40] G.V. Chudnovsky, Padé approximation and the Riemann monodromy problem, in Bifurcation Phenomena in Mathematical Physics and Related Topics, D. Reidel, Boston, 1980, 448-510.

[41] G.V. Chudnovsky, Rational and Padé approximation to solutions of linear differential equations and the monodromy theory. Lecture Notes Physics, v. 126, Springer, 1980, 136-169.

[42] G.V. Chudnovsky, Padé approximations to the generalized hypergeometric functions. I., J. Math. Pures et Appliques, Paris, 58 (1979), 445-476.

[43] G.V. Chudnovsky, On the method of Thue-Siegel, Ann. of Math., 117 (1983), 325-382.

[44] C.L. Siegel, Transcendental Numbers, Princeton University Press, 1949.

[45] G.V. Chudnovsky, Rational approximations to linear forms of exponentials and binomials, Proc. Nat'l Acad. Sci. U.S.A., 80 (1983), 3138-3141.

[46] D.V. Chudnovsky, G.V. Chudnovsky, A random walk in higher arithmetic, Adv. Appl. Math., 7 (1986), 101-122.

[47] H. Poincare, Sur les groupes des équations lineaires, Acta. Math., 4 (1884) 201-312.

[48] D.A. Hejhel, Monodromy groups and Poincaré series, Bull. Amer. Math. Soc., 84 (1978), 339-376.

[49] E.T. Whittaker, On the connexion of algebraic functions with automorphic functions, Phil. Trans., 122 A(1898), 1-32.

[50] R.A. Rankin, The differential equations associated with the uniformization of certain algebraic curves, Proc. Roy. Soc. Edinburgh, 65 (1958) 35-62.

[51] A. Schonhage, V. Strassen, Schnelle Multiplikation grosser Zahlen, Computing, 7 (1971), 281-292.

[52] D.V. Chudnovsky, G.V. Chudnovsky, Algebraic complexities and algebraic curves over finite fields, Proc. Natl. Acad. Sci. USA, 84 (1987), 1739-1743.

[53] A. Schonhage, Equation solving in terms of computational complexity, Proc. International Congress of Mathematicians, Berkeley, 1986.

[54] R.P. Brent, Multiple-precision zero-finding methods and the complexity of elementary function evaluation, in Analytic Computational Complexity, J.F. Traub, Ed., Academic Press, 1975, 151-176.

[55] R.P. Brent, The complexity of multiple-precision arithmetic, in Complexity of Computational Problem Solving, R.S. Anderson and R.P. Brent, Eds. Univ. of Queensland Press, Brisbane, Australia, 1975, 126-165.

[56] D.V. Chudnovsky, G.V. Chudnovsky, On expansions of algebraic functions in power and Puiseux series, I; II; J. of Complexity, 2 (1986), 271-294; 3 (1987), 1-25.

[57] O. Perron, Die Lehre von den Kettenbrüchen, Teubner, 1929.

[58] D. Bini, V. Pan, Polynomial division and its computational complexity, J. of Complexity, 2 (1986), 179-203.

[59] D. A. Hejhal, A classical approach to a well-known spectral correspondence on quaternion groups, Lecture Notes Math., v. 1135, Springer, 1985, 127-196.

Multiple q-Series and $U(n)$ generalizations of Ramanujan's ${}_1\Psi_1$ sum

by

S.C. Milne*

Department of Mathematics
University of Kentucky
Lexington, Kentucky 40506 - 0027

*Partially supported by NSF Grant DMS-8604232