PADE-APPROXIMATIONS IN NUMBER THEORY

F. Beukers
Mathematisch Instituut
Rijksuniversiteit Leiden
Wassenaarseweg 80
2333 AL Leiden
The Netherlands

INTRODUCTION. In 1873 Hermite [H] was the first to construct explicit simultaneous Padé-approximations to the system of functions $1, e^z, e^{2z}, \ldots, e^{nz}$ and discovered the transcendence of e. Later Lindemann [L] in 1882 extended Hermite's work to show that π is transcendental, thus providing the negative answer to the ancient problem of squaring the circle. An elegant exposition of these methods can be found in Siegel [Si1], Chapter I. The work of Hermite-Lindemann was largely extended by the work of Siegel and later Shidlovski on the algebraic independence of values of so-called E-functions. In this extension however, the authors use non-explicitly constructed rational approximations and we shall not proceed along these lines. We would like to refer interested readers to [Ba2], Chapter 11.

Very recently there has been an upsurge of interest in the use of Padé-approximations in irrationality and transcendence proofs, stimulated by Apéry's remarkable irrationality proof for $\zeta(3) = 1^{-3} + 2^{-3} + 3^{-3} + \ldots$. Literature on this proof can be found in R. Apéry [A], A.J. van der Poorten [P], E. Reyssat [R] and F. Beukers [Be1]. In attempts to generalize Apéry's method the role played by Padé-approximations in irrationality theory began to be re-appreciated. We first define what we mean by Padé-approximation.

Let $f_1, f_2, \ldots, f_k$ be a system of functions analytic around $z = 0$ and suppose $f_1(0) \neq 0$. We distinguish two kinds of Padé-approximations (see K. Mahler [M], H. Jager [J]) type I : polynomials $P_1(z),\dots,P_k(z)$ of degree $n_1,\dots,n_k$ such that

$$P_1(z)f_1(z) + \dots + P_k(z)f_k(z) = O(z^{N+k-1})$$

where $N = \Sigma n_i$

type II: polynomials $P_1(z),\dots,P_k(z)$ of degree $N-n_1,\dots,N-n_k$ with $\Sigma n_i = N$, such that

$$P_i(z)f_j(z) - P_j(z)f_i(z) = O(z^{N+1}) \qquad i,j=1,\dots,k .$$

Notice that if $k = 2$, then both types coincide, and we have in fact the classical Padé-table of the function $f_2(z)/f_1(z)$.

In the following we shall use the abbreviation P.A. for Padé-approximation.

G.V. Chudnovsky [C1] [C2] has constructed a very wide class of explicit type I and II-approximations of systems of generalised hypergeometric functions, which can be applied to obtain irrationality-results. To quote a few of them,

1) $\Gamma(\frac{1}{4})^4/\pi^2$ is irrational, [C2]

2) $\mathrm{dil}(a^{-1}) \notin \mathbb{Q}$ for $a \in \mathbb{Z}$, $|a| \geq 14$, [C1] [C2]

where

$$\mathrm{dil}(z) = \frac{z}{1^2} + \frac{z^2}{2^2} + \frac{z^3}{3^2} + \dots$$

3) $|\pi - \frac{p}{q}| > q^{-19.89}$ for all $\frac{p}{q} \in \mathbb{Q}$, $q > q_0$. [C1]

Also, the importance of Chudnovsky's work on the theoretical side of approximation theory should be stressed [C3] [C4]. One of the main features is the close connection that exists between P.A.'s to systems of functions satisfying a linear differential equation and the monodromy group of this differential equation.

In Section 1 of this note we give an impression of the applications of Padéfractions in irrationality theory by showing, $e^a \notin \mathbb{Q}$ for $a \in \mathbb{Q}$, $a \neq 0$.

In Section 2 we will review some of the results that have been obtained by application of P.A.'s of $(1-z)^{1/n}$ to some diophantine equations.

Despite the interest in P.A.'s that Apéry's irrationality proof for $\zeta(3)$ has aroused, it was hitherto unclear how to formulate Apéry's proof naturaly in terms of P.A.'s. In Section 3 we indicate how this might be achieved, although we must extend our définition of Padé-approximation a little.

SECTION 1.

THEOREM 1. *Let* $a \in \mathbb{Q}$, $a \neq 0$. *Then* e^a *is irrational.*

PROOF. Notice that it is sufficient to prove this theorem for $a \in \mathbb{N}$. The proof for $a \in \mathbb{Q}$ then follows easily by noticing that $(e^a)^{\mathrm{den}(a)} \notin \mathbb{Q}$ (where den(a) = denominator of a) and so we certainly have $e^a \notin \mathbb{Q}$. The $[n,n]$ P.A. of e^z can be found as follows. Consider

$$I_n(z) = z^{n+1}\int_0^1 e^{zt}\, P_n(t)dt \tag{1}$$

where $P_n(t)$ is the Legendre polynomial defined by $P_n(t) = \frac{1}{n!}\left(\frac{d}{dt}\right)^n t^n(1-t)^n$.

Notice that degree $P_n(t) = n$ and $P_n(t) \in \mathbb{Z}[t]$. By repeated partial integration we obtain,

$$I_n(z) = (-1)^n \frac{z^{2n+1}}{n!}\int_0^1 e^{zt}\, t^n(1-t)^n dt\ . \tag{2}$$

On the other hand, it is straightforward to see that

$$z^{n+1}\int_0^1 t^m e^{zt}dt = z^{n-m}\int_0^z e^t t^m dt$$

$$= Q_n(z)e^z + (-1)^{m+1}\, m!\ z^{n-m}$$

where $Q_n(z) \in \mathbb{Z}[z]$ has degree n. Therefore, term by term integration of (1) yields

$$I_n(z) = A_n(z) + B_n(z)e^z , \tag{3}$$

where $A_n(z), B_n(z) \in \mathbb{Z}[z]$ have degree $\leq n$.

We now substitute $z = a$. From (1) it is easy to see that $I_n(a) \neq 0$. Suppose $e^a = p/q \in \mathbb{Q}$, then (3) yields

$$\frac{1}{q} \leq |A_n(a)+B_n(a)\frac{p}{q}| = |I_n(a)| ,$$

which, for sufficiently large n, is in contradiction with the upper bound we obtain from (2),

$$|I_n(a)| < \frac{a^{2n+1}}{n!} e^a .$$

Hence e^a is irrational.

With a similar method it is also possible to show the irrationality of π^2 and the zeros of Bessel-functions of integer order, see [Be2]. We can also show the irrationality of $\log 2$ and $\pi/\sqrt{3}$ by using the Padé-table for $\log(1-z)$ and then substituting $z = -1$ and $z = e^{\pi i/3}$ respectively. Moreover, by refining the arguments in this case we can show theorems of the following type,

THEOREM 2. *For every* $\varepsilon > 0$ *there exist explicitly calculable numbers* $q_0(\varepsilon), q_1(\varepsilon)$ *such that*

$$|\log 2 - \frac{p}{q}| > |q|^{-4.660137\ldots-\varepsilon} \quad \text{for} \quad |q| > q_0(\varepsilon)$$

$$|\frac{\pi}{\sqrt{3}} - \frac{p}{q}| > |q|^{-8.30998\ldots-\varepsilon} \quad \text{for} \quad |q| > q_1(\varepsilon) .$$

For a clear derivation of these irrationality measures, see [A-R]. They were found independently by G.V. Chudnovsky [C1], and several others.

SECTION 2. In 1964 A. Baker [Ba] used P.A.'s to $(1-z)^{1/3}$ in order to prove

THEOREM 3. *For any* $\frac{p}{q} \in \mathbb{Q}$ *we have*

$$\left|\frac{p}{q} - \sqrt[3]{2}\right| > \frac{10^{-6}}{q^{2.955}} .$$

Multiplication of this inequality with $q(p^2+pq\sqrt[3]{2}+q^2\sqrt[3]{4})$ yields

$$|p^3 - 2q^3| > 10^{-6}\ q^{0.045} \quad \text{for any } p,q \in \mathbb{N} .$$

This implies that the diophantine equation $x^3 - 2y^3 = k$ (k given integer) has only finitely many solutions. Moreover, if x,y is a solution then $|y| < 10^{138}\ |k|^{23}$. Now let a,b,c,n be given integers with $n \geq 3$. In general we can use P.A.'s to $(1-z)^{1/n}$ in order to study the diophantine equation

$$(4) \qquad ax^n - by^n = c$$

in the unknown integers x,y . It is only possible however to give upper bounds for the number of solutions of (4) and not for the size of the solutions. In 1937 C.L. Siegel [Si2] was the first to study equation (4) in this way. By elaborating Siegel's methods one can show that if $n \geq 5$ and $c = 1$, equation (4) has at most 2 solutions (with $x,y \geq 0$ if n is even). See [D]. Very recently J. Evertse showed that if c is a prime-power then there are at most $2n + 6$ solutions (private communication).

In 1977 the author, using P.A.'s to $\sqrt{1-z}$, obtained, THEOREM 4. *For any* $x,r \in \mathbb{N}$ *we have*

$$\left|\frac{x}{2^r} - \sqrt{2}\right| > \frac{2^{-43.9}}{2^{1.8r}} .$$

Multiplication of this inequality with $2^{2r}(\sqrt{2}+x2^{-r})$ yields

$$|x^2 - 2^{2r+1}| > 2^{0.2r}\, 2^{-43.4}$$

from which we easily derive

COROLLARY. *Let* $D \in \mathbb{Z}$ *and let* $x,n \in \mathbb{N}$ *be a solution of the diophantine equation* $x^2 + D = 2^n$. *Then* $n < 435 + 10 \log|D|/\log 2$.

As a consequence we see that for given $D \in \mathbb{Z}$ the diophantine equation $x^2 + D = 2^n$ can be solved in finitely many steps. Moreover, after some technical considerations it is possible to show that $x^2 + D = 2^n$ has at most four solutions, unless $D = 7$ in which case the solutions read $(x,n) = (1,3),(3,4),(5,5),(11,7),(181,15)$. All this can be found in [Be3].

SECTION 3. The by now traditional way to prove the irrationality of $\zeta(3)$ can be sketched as follows. Define

$$a_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 .$$

Then there exist numbers $b_n \in [1,\ldots,n]^{-3}\mathbb{Z}$ (here $[1,\ldots,n]$ denotes the lcm.) with

$$(5) \qquad 0 < |a_n - b_n\zeta(3)| < 3(\sqrt{2}-1)^{4n} .$$

If $\zeta(3)$ were rational, say p/q then $|a_n - b_n\zeta(3)| \geq q^{-1}[1,\ldots,n]^{-3}$, contradicting the upper bound in (5) for n sufficiently large. For full details, see [R] or [Be1].

We will now show how the numbers a_n and b_n can be derived from Padé-type approximations. Define

$$L_k(z) = \frac{z}{1^k} + \frac{z^2}{2^k} + \frac{z^3}{3^k} + \dots .$$

Notice that $L_2(1) = \zeta(2)$ and $L_3(1) = \zeta(3)$. We look for polynomials $A_n(z), B_n(z), C_n(z), D_n(z)$ of degree n such that

$$(6) \qquad \begin{aligned} A_n(z)L_2(z) + B_n(z)L_1(z) + C_n(z) &= O(z^{2n+1}) \\ 2A_n(z)L_3(z) + B_n(z)L_2(z) + D_n(z) &= O(z^{2n+1}) \end{aligned}$$

and $B_n(1) = 0$. The four polynomials have $4(n+1)$ coefficients and the system (6) together with $B_n(1) = 0$ gives $2(2n+1) + 1 = 4n + 3$ linear conditions, so that the polynomials A_n, B_n, C_n, D_n really exist. Write

$$A_n(z) = \sum_{r=0}^{n} \alpha_r z^r \quad \text{and} \quad B_n(z) = \sum_{r=0}^{n} \beta_r z^r .$$

Since degree $C_n, D_n \le n$, the Taylor coefficient of z^m $(n+1 \le m \le 2n)$ in $A_nL_2 + B_nL_1$, $2A_nL_3 + B_nL_2$ respectively, must be zero, i.e.

$$(7) \qquad \begin{aligned} \sum_{r=0}^{n} \frac{\alpha_r}{(m-r)^2} + \frac{\beta_r}{m-r} &= 0 \\ \sum_{r=0}^{n} \frac{2\alpha_r}{(m-r)^3} + \frac{\beta_r}{(m-r)^2} &= 0 \end{aligned} \qquad m=n+1,\dots,2n .$$

Furthermore, $B_n(1) = 0$ implies $\Sigma\beta_r = 0$. This system of linear equations for α_r and β_r is easy to solve. Consider the rational function

$$R_n(t) = \sum_{r=0}^{n} \frac{\alpha_r}{(t-r)^2} + \frac{\beta_r}{t-r} = \frac{Q_n(t)}{t^2(t-1)^2\dots(t-n)^2} .$$

The conditions (7) now imply that $R_n(t)$ and its derivative are zero for $t = n+1, n+2, \ldots, 2n$. This implies that $Q_n(t)$ is a multiple of $(t-n-1)^2(t-n-2)^2\ldots(t-2n)^2$. If we put $Q_n(t)$ equal to this product then degree $Q_n(t) = 2n$, whereas the denominator of $R_n(t)$ has degree $2n + 2$. This automatically implies $\Sigma\beta_r = 0$. Therefore, the coefficients α_r, β_r can be obtained from the partial fraction expansion of

$$\frac{(t-n-1)^2(t-n-2)^2\ldots(t-2n)^2}{t^2(t-1)^2\ldots(t-n)^2} .$$

In particular it is easy to see that

$$\alpha_r = \binom{n}{r}^2 \binom{2n-r}{n}^2 .$$

Substitute $z = 1$ in (6) and use $B_n(1) = 0$. Then the second line yields

(8) $$2A_n(1)\zeta(3) + D_n(1) = \text{remainder}$$

where

$$A_n(1) = \sum_{r=0}^{n} \binom{n}{r}^2 \binom{2n-r}{n}^2 = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{n}^2 = a_n .$$

Thus we have recovered the number a_n from the approximations (6). It is now a matter of straightforward computation to show that the approximation (8) is actually the same as (5).

REFERENCES

[A-R] K. Alladi, M. Robinson, On certain irrational values of the logarithm, Lecture Notes in Math. 751, 1-9.

[A] R. Apéry, Irrationalité de $\zeta(2)$ et $\zeta(3)$. "Journées arithmétiques de Luminy", Astérisque n^o 61, 1979, 11-13.

[Ba1] A. Baker, Rational approximations to $\sqrt[3]{2}$ and other algebraic numbers, Quart. J. Math. Oxford, 15(1964), 375-383.

[Ba2] A. Baker, Transcendental Number Theory (Cambridge, 1975).

[Be1] F. Beukers, A note on the irrationality of $\zeta(2)$ and $\zeta(3)$, Bull. London Math. Soc., 11(1979), 268-272.

[Be2] F. Beukers, Legendre polynomials in irrationality proofs, Bull. Australian Math. Soc. (to appear).

[Be3] F. Beukers, The generalised Ramanujan-Nagell equation, Thesis, University of Leiden (1979), also to appear in Acta Arithmetica.

[C1] G.V. Chudnovsky, C.R. Acad. Sc. Paris, 288(1979), 607-609, 965-967, 1001-1003.

[C2] G.V. Chudnovsky, Padé-approximations to the generalized hypergeometric functions I, J. Math. pures et appl. 58(1979), 445-476.

[C3] G.V. Chudnovsky, Rational and Padé-approximations to solutions of linear differential equations and the monodromy theory, Lecture Notes in Physics 126, 136-169.

[C4] G.V. Chudnovsky, Padé-approximation and the Riemann monodromy problem, Proceedings of the NATO Advanced Study Institute, held at Cargèse, Corsica, France, June 24-July 7, 1979.

[D] Y. Domar, On the diophantine equation $|Ax^n-By^n| = 1$, $n \geq 5$, Math. Scand. 2(1954), 29-32.

[H] Ch. Hermite, Sur la fonction exponentielle, Oeuvres III, 150-181.

[J] H. Jager, A multidimensional generalization of the Padé table, Thesis, University of Amsterdam (1964).

[L] F. Lindemann, Ueber die Zahl π , Math. Ann. 20(1882), 213-225.

[M] K. Mahler, Application of some formulae by Hermite to the approximation of exponentials and logarithms, Math. Ann. 168(1976), 200-227.

[P] A.J. van der Poorten, A proof that Euler missed ... Apéry's proof of the irrationality of $\zeta(3)$, Math. Intelligencer, 1(1978), 195-203.

[R] E. Reyssat, Irrationalité de $\zeta(3)$ selon Apéry, Sém. Delange-Pisot-Poitou, 20e année, 1978/79, n^{o} 6.

[Si1] C.L. Siegel, Transcendental Numbers (Princeton 1949).

[Si2] C.L. Siegel, Die Gleichung $ax^n - by^n = c$, Math. Ann. 114(1937), 57-68.