HERMITE-PADÉ APPROXIMANTS AND APERY THEOREM

M. PRÉVOST AND V.N. SOROKIN

Abstract. A new Hermite-Padé construction for a system of generalized polylogarithms forming a cycle graph is proposed. Diophantine Apéry approximants for the value of the Riemann zeta function $\zeta(3)$ are obtained as a corollary.

Introduction.

In 1978, Apéry [1] presented Diophantine approximations for the values of the Riemann zeta function $\zeta(3)$ which proves the irrationality of this number. The analytic construction producing the Apéry approximations was first proposed independently by Beukers [2] and Gutnik [3]. This construction is a matrix Hermite-Padé problem for polylogarithms. The first author of this paper, M. Prévost, proposed [4] another proof of the Apéry theorem utilizing Padé approximations. The second author, V. Sorokin considered a classification of various Hermite-Padé problems for generalized polylogarithms leading to Apéry approximations. Namely, five new analytic constructions were proposed in [5, 6, 7]. Recently, as a result of joint work , the authors revealed one more and probably the last construction of such a type. This paper is devoted to study of this construction.

1. Analytic properties of generalized polylogarithms

Consider the following power series (generalized polylogarithms)

$$e(z)=\sum_{n=1}^{\infty}\frac{1}{z^n}\frac{1}{n},\quad f(z)=\sum_{n=1}^{\infty}\frac{1}{z^n}\frac{1}{n}\sum_{m=1}^{n}\frac{1}{m},\quad c(z)=\sum_{n=1}^{\infty}\frac{1}{z^n}\frac{1}{n^2}\sum_{m=1}^{n}\frac{1}{m} \tag{1}$$

These series had been already used in two Hermite-Padé problems in [6].

Series (1) converges in the domain $\{z\in\mathbf{C}:|z|>1\}$.

Note that $e(z)=-\log(1-1/z)$ is a logarithmic function holomorphic in the plane with the cut $\bar{\mathbf{C}}\setminus\Delta_0,\Delta_0=[0,1]$.

Functions (1) satisfy the differential equations

$$\frac{d}{dz}c(z)=-\frac{1}{z}f(z),\quad \frac{d}{dz}f(z)=\frac{1}{1-z}e(z),\quad \frac{d}{dz}e(z)=\frac{1}{z(1-z)} \tag{2}$$

Therefore, the functions $f(z)$ and $c(z)$ are holomorphic in the same domain.

A complete analytic continuation of functions (1) produces analytic multifunctions whose singularities are logarithm ramification points at $z=0,1,\infty$.

Weights for functions (1) are determined by the Sokhotsky formulas:

$$p(x)=\frac{1}{\pi}Im\,c(x-i.0),\quad \varepsilon^*(x)=\frac{1}{\pi}Im\,f(x-i.0),\ 0<x<1.$$

For the logarithms we have

$$\frac{1}{\pi} Im\, e(x - i.0) = 1.$$

then functions (1) are integrals of Cauchy type (actually, they are functions of the Markow type):

$$c(z) = \int_0^1 \frac{p(x)}{z-x} \mathrm{d}x, \qquad f(z) = \int_0^1 \frac{\varepsilon^*(x)}{z-x} \mathrm{d}x, \qquad e(z) = \int_0^1 \frac{1}{z-x} \mathrm{d}x \tag{3}$$

It was shown in [6] that

$$\varepsilon^*(x) = -\log(1-x), \quad x \in D_+ = \bar{\mathbf{C}} \setminus \Delta_+, \quad \Delta_+ = [1, +\infty]. \tag{4}$$

$$p(x) = \delta(x) + \varepsilon(x)\varepsilon^*(x) = \delta^*(1) - \delta^*(x), \quad x \in D_+ \tag{5}$$

where

$$\varepsilon(x) = -\log x,\ x \in D_- = \bar{\mathbf{C}} \setminus \Delta_-, \Delta_- = [-\infty, 0],$$

and

$$\delta(x) = \sum_{n=1}^{\infty} \frac{(1-z)^n}{n^2}$$

is a polylogarithm (a function holomorphic in D_-), in this case $\delta^*(x) = \delta(1-x)$.

Note that $p(x)$ is the so-called Roger polylogarithm.

Applying Sokhostky formulas to functions (4), (5), we get

$$\frac{1}{\pi} Im\, \varepsilon^*(t + i.0) = 1, \quad \frac{1}{\pi} Im\, p(t + i.0) = \varepsilon(t), \quad 1 < t < +\infty$$

In another form,

$$\frac{\varepsilon^*(x)}{x} = \int_1^\infty \frac{1}{t-x} \frac{dt}{t}, \qquad \frac{p(x)}{1-x} = \int_1^\infty \frac{\varepsilon(t)}{t-x} \frac{dt}{1-t}. \tag{6}$$

And finally,

$$\frac{1}{\pi} Im\, \varepsilon(s - i.0) = 1, -\infty < s < 0,$$

i.e.,

$$\frac{\varepsilon(t)}{1-t} = \int_{-\infty}^0 \frac{1}{t-s} \frac{ds}{1-s}. \tag{7}$$

Thus, functions (1) form Nikishin's system corresponding to the graph

$$1 \stackrel{\Delta_0}{\to} e \stackrel{\Delta_+}{\to} f \stackrel{\Delta_-}{\to} c.$$

This graph may be interpreted either as a system of differential equations (2) or as the sequence of integral representations (3), (6), (7) according to notations of [8]

2. Formulation of the Hermite-Padé problem

The function $c^*(z) = c(z) - \varepsilon(z)f(z)$ holomorphic in the domain $\bar{\mathbf{C}} \setminus [-\infty, 1]$ has the following derivative:

$$\frac{d}{dz}c^*(z) = -\frac{1}{1-z}\varepsilon(z)e(z)$$

Thus, the collection of analytic functions $\{c^*, \varepsilon e, e, \varepsilon, 1\}$ satisfies a system of differentialequations representable as the graph

$$\begin{array}{ccccccccc} & & & \varepsilon & & & & & \\ & \Delta_- & & & \Delta_0 & & & & \\ & \nearrow & & & \searrow & & & & \\ 1 & & & & & \varepsilon e & \longrightarrow & c^* & \qquad (*) \\ & \searrow & & & \nearrow & & & & \\ & \Delta_0 & & & \Delta_- & & & & \\ & & & e & & & & & \end{array}$$

Consider the linear form corresponding to this graph

$$\mathcal{C}_n(z) = C_n(z)c^*(z) + \tfrac{1}{(1-z)^n}\left\{ \begin{array}{ccc} & +\mathcal{E}_n(z)\varepsilon(z) & \\ B_n(z)\varepsilon(z)e(z) & & +J_n(z) \\ & +E_n(z)e(z) & \end{array} \right\}$$

where C_n is a polynomial of degree not greater than n and not identically equal to zero and $B_n, \mathcal{E}_n, E_n, J_n$ are polynomials of degree not greater than $2n$.

We require that the function $\mathcal{C}_n(z)$ has zero of order at least $(n+1)$ at infinity (here and further, up to a logarithmic singularity).

This condition is split into systems of equations:

$$c_n(z) = C_n(z)c^*(z) + \frac{1}{(1-z)^n}\{E_n(z)e(z) + J_n(z)\} = O(1/z^{n+1}), z \to \infty \qquad (8)$$

$$f_n(z) = -C_n(z)f(z) + \frac{1}{(1-z)^n}\{B_n(z)e(z) + \mathcal{E}_n(z)\} = O(1/z^{n+1}), \quad z \to \infty \qquad (9)$$

Further, consider the weight functions

$$\delta_n(x) = \frac{1}{\pi}Im\ \mathcal{C}_n(x - i.0) = C_n(x)\delta(x) + \frac{1}{(1-x)^n}\{B_n(x)\varepsilon(x) + E_n(x)\}, \quad 0 < x < 1,$$

$$f_n(x) = \frac{1}{\pi}Im\ \mathcal{C}_n(x - i.0) = -C_n(x)f(x) + \frac{1}{(1-x)^n}\{B_n(x)e(x) + \mathcal{E}_n(x)\}, \quad -\infty < x < 0.$$

We require that the functions $\mathcal{C}_n$ has a finite limit for $z \to 1$. This condition is split into two linear equations:

$$\mathcal{N}_n(x) = \mathcal{E}_n(x)\varepsilon(x) + J_n(x) = O((1-x)^n), \quad x \to 1, \tag{10}$$

$$B_n(x)\varepsilon(x) + E_n(x) = O(((1-x)^{n+1}), \quad x \to 1. \tag{11}$$

Replace (11) by a stronger condition on the weight function

$$\delta_n(x) = O((1-x)^{n+1}), \quad x \to 1, \tag{12}$$

which should provide fast convergence to zero for the values $\mathcal{C}_n(1)$ for $n \to \infty$. Then

$$\mathcal{C}_n(1) = C_n(1)c^*(1) + \gamma_n,$$

where $\mathcal{N}_n(x) = \gamma_n(1-x)^n + \cdots$.

In this case $c^*(1) = 2\zeta(3)$, i.e.,we get rational approximations to the values of the zeta-function.

Now consider weights for the functions $\delta_n(x)$ and $f_n(x)$:

$$\varepsilon_n^*(t) = \frac{1}{\pi} \ Im \ \delta_n(t - i.0), \quad -\infty < t < 0,$$

$$\varepsilon_n^*(t) = \frac{1}{\pi} \ Im \ f_n(t - i.0), \quad 0 < t < 1,$$

where

$$\varepsilon_n^*(t) = -C_n(t)\varepsilon^*(t) + \frac{B_n(t)}{(1-t)^n}.$$

We require that the difference of the functions $\delta_n(x)$ and $f_n(x)$ has zero of order not less than $n+1$ at the point $x = 0$, namely

$$C_n(x)\left(-(2\delta^*(x) - \frac{1}{2}\varepsilon^{*2}(x)) - (2\varepsilon(x) \pm \pi \ i)\varepsilon^*(x)\right)$$

$$+\frac{1}{(1-x)^n}\left\{B_n(x)\left(-\varepsilon^*(x) + (2\varepsilon(x) \pm \pi \ i)\right) + (E_n(x) - \mathcal{E}_n(x))\right\} = O(x^{n+1}\log x), \ x \to 0,$$

where the sign at the constant $\pm\pi \ i$ is determined by the half-plane x belongs to (the upper or lower one, respectively). The latter condition is split into two systems of linear equations:

$$\varepsilon_n^*(x) = -C_n(x)\varepsilon^*(x) + \frac{B_n(x)}{(1-x)^n} = O(x^{n+1}), \ x \to 0, \tag{13}$$

$$\omega_n^*(x) = -C_n(x)\omega^*(x) + \frac{1}{(1-x)^n}\left\{-B_n(x) \ \varepsilon^*(x) + (E_n(x) - \mathcal{E}_n(x))\right\} = O(x^{n+1}), \ x \to 0, \tag{14}$$

where

$$\omega_n^*(x) = 2\delta^*(x) - \frac{1}{2}\varepsilon^{*2}(x).$$

Thus se have revealed the analytic meaning of the hypothesis of the following Hermite-Padé problem for the graph (*).

Problem HP. *How to find a polynomial $C_n \neq 0$ of degree $\leq n$ and also polynomials $B_n, E_n, \mathcal{E}_n$ and J_n of degree $\leq 2n$ so that conditions (8)-(10), (12)-(14) hold?*

3. Solution to the Hermite-Padé problem

Denote by $e_n(x)$ the Legendre functions, in other words, the Padé approximations of the logarithmic function $e(x)$ at infinity, namely

$$e_n(x) = a_n(x)e(x) - b_n(x) = O(1/x^{n+1}),\ x \to \infty, \tag{15}$$

where a_n and b_n are polynomials of degree not greater than n and $a_n \neq 0$. Problem (15) has the unique (hereafter, up to normalization) solution and a_n are the Legendre polynomials. Normalize the solutions by the condition $a_n(0) = 1$. Assume

$$\varepsilon_n(x) = (1-x)^n\, e_n\left(\frac{1}{1-x}\right) = \alpha_n(x)\varepsilon(x) - \beta_n(x) = O((1-x)^{2n+1}),\ x \to 1.$$

Then $\varepsilon_n(x)$ is the Padé approximation of the logarithmic function $\varepsilon(x)$ at the point $x = 1$.

Theorem. *Problem **HP** has the unique (up to normalization) solution admitting the following integral representation:*

$$\mathcal{C}_n(z) = \int_z^\infty \frac{\varepsilon_n(x)e_n(x)(x-z)^n\, dx}{(1-x)^{2n+1}}. \tag{16}$$

Proof. Conditions of problem **HP** may be reduced to a system of $9n+5$ linear homogeneous equations for $9n+5$ unknown coefficients of the polynomials. But these equations are not independent. It is easy to see that the condition $E_n(0) = \mathcal{E}_n(0)$ follows from other conditions. Therefore, the system has non trivial solutions.

Let $\mathcal{C}_n(x)$ be an arbitrary solution to the problem **HP**. Applying the differential operator to this functions, we get

$$\frac{1}{n!}\left(\frac{d}{dx}\right)^{n+1} \mathcal{C}_n(x) = \frac{\widetilde{\mathcal{B}}_n(x)}{(1-x)^{2n+1}}$$

where $\widetilde{\mathcal{B}}_n(x)$ is a linear form

$$\widetilde{\mathcal{B}}_n(x) = \widetilde{B}_n(x)\varepsilon(x)e(x) + \begin{matrix} \widetilde{\mathcal{E}}_n(x)\varepsilon(x) \\ \\ \widetilde{E}_n(x)e(x) \end{matrix} + \widetilde{J}_n(x)$$

with some polynomials of degree $\leq 2n$. Note that the variable x does not occur in the denominator under differentiation because of conditions (13) and (14). The function $\mathcal{C}_n(x)$ may be uniquely reconstructed from the function $\widetilde{\mathcal{B}_n}(x)$ by using the integral operator

$$\mathcal{C}_n(z) = \int_z^\infty \frac{\widetilde{\mathcal{B}}_n(x)}{(1-x)^{2n+1}}(x-z)^n\, dx.$$

Relations (8) and (9) imply

$$\widetilde{\mathcal{B}}_n(x) = O\left(\frac{\log x}{x}\right),\ x \to \infty, \tag{17}$$

and (10) and (12) imply

$$\widetilde{\mathcal{B}}_n(x) = O((1-x)^{2n+1}\log(1-x)),\ x \to 1, \tag{18}$$

Obviously, the function $\widetilde{\mathcal{B}}_n(x) = \varepsilon_n(x)e_n(x)$ satisfies conditions (17), (18). It remains to establish that the solution to problem (17), (18) is unique. We do this in the following section.

4. Solution to the auxiliary problem

Perform the change of variables

$$R_n(x) = x^{2n}\widetilde{\mathcal{B}}_n(\tfrac{1}{x}) = \widetilde{B}_n(x)\varepsilon(x)\varepsilon^*(x) + \begin{matrix} \widetilde{\mathcal{E}}_n(x)\varepsilon(x) \\ \\ \widetilde{\mathcal{E}}^*_n(x)\varepsilon^*(x) \end{matrix} + \widetilde{J}_n(x)$$

Then we get the following problem.

Problem R. *How to find polynomials* $\widetilde{B}_n, \widetilde{\mathcal{E}}_n, \widetilde{\mathcal{E}}^*_n$ *and* $\widetilde{J}_n$ *of degree* $\leq 2n$ *and* $\widetilde{B}_n \neq 0$ *so that the following conditions hold:*

$$\begin{aligned} \widetilde{\varepsilon}_n(x) = \widetilde{B}_n(x)\varepsilon(x) + \widetilde{\mathcal{E}}^*_n(x) &= O\left((1-x)^{2n+1}\right),\ x \to 1, \\ \widetilde{\mathcal{E}}_n(x)\varepsilon(x) + \widetilde{J}_n(x) &= O\left((1-x)^{2n+1}\right),\ x \to 1, \\ \widetilde{\varepsilon}^*_n(x) = \widetilde{B}_n(x)\varepsilon^*(x) + \widetilde{\mathcal{E}}_n(x) &= O\left(x^{2n+1}\right),\ x \to 0, \\ \widetilde{\mathcal{E}}^*_n(x)\varepsilon^*(x) + \widetilde{J}_n(x) &= O\left(x^{2n+1}\right),\ x \to 0. \end{aligned}$$

Lemma. *The solution to problem **R** exists and is unique (up to normalization).*

Proof. Let $R_n(x)$ be an arbitrary even (with respect to the transform $x \to 1-x$) solution to problem **R**. We know that such nontrivial solutions exist. Conditions of the problem imply the orthogonality conditions

$$\int_{-\infty}^{0} \widetilde{\varepsilon}^*_n(t)t^m\, d\mu_n(t) - \int_1^{\infty} \widetilde{\varepsilon}_n(t)t^m\, d\mu_n(t) = 0,\ m = 0, \cdots, 2n,$$

where

$$d\mu_n(t) = \frac{dt}{t^{2n+1}(1-t)^{2n+1}}.$$

Since the solution is even, these relations are equivalent to

$$\int_1^{\infty} \widetilde{\varepsilon}_n(t)(2t-1)^{2m+1}\, d\mu_n(t) = 0,\ m = 0, \cdots, n-1.$$

Then from the properties of Nikishin's systems we get that the polynomials $\widetilde{B}_n(x)$ have the degree exactly $2n$, therefore, the solution is unique (up to normlization).

Applying similar considerations to an arbitrary odd solution, we get that it identically equals zero.

5. Connection with Apéry numbers

The theorem from section 3 implies

Corollary. The values $\mathcal{C}_n(1)$ are Diophantine Apéry approximations of the number $\zeta(3)$.

Proof. Transform integral (16) for $z = 1$ to the integral representation of the Apéry approximations obtained earlier in [6]. Changing the variables $x \to 1/x$, we get

$$\mathcal{C}_n(1) = \int_0^1 \frac{\varepsilon_n(x)\varepsilon_n^*(x)}{x^{n+1}(1-x)^{n+1}}\, dx,$$

where $\varepsilon_n^*(x) = \varepsilon_n(1-x)$. Utilize the following integral representations of the Legendre functions:

$$\varepsilon_n^*(x) = (-1)^n x^{2n+1} \int_0^1 \frac{u^n(1-u)^n}{(1-u\,x)^{n+1}}\, du,$$

$$\varepsilon_n(x) = (-1)^n \int_x^1 (1-\xi)^n \left(1-\frac{x}{\xi}\right) \frac{d\xi}{\xi}$$

and perform the change of variables $y = x\,\xi$. Then we get the desired integral representation

$$\mathcal{C}_n(1) = \int_0^1 \int_0^1 \int_0^1 \frac{\xi^n(1-\xi)^n y^n(1-y)^n u^n(1-u)^n\, d\xi\, dy\, du}{(1-y\xi)^{n+1}(1-uy\xi)^{n+1}}.$$

The corollary is proved.

The work was partly supported by INTAS, project 03-51-6637, the Russian Foundation for Basic Research, project 02-01-00565, and the Program of support for leading scientific schools, project BSh 1551.2003.1.

References

[1] R. Apéry, "Irrationalité de $\zeta(2)$ et $\zeta(3)$", *Astérisque*, **61**, 11-13, 1979.
[2] F. Beukers, "Padé approximations in number theory", *Lect. Notes Math.*, **888**, 90-99, 1981.
[3] L.A. Gutnik, "Irrationality of some values containing $\zeta(3)$", *Acta Arith.*, **42**, 3:255-264, 1983.
[4] M. Prévost, "A new proof of the irrationality of $\zeta(2)$ and $\zeta(3)$ using Padé approwimants", *J. Comput. Appl. Math.,* **67**, 2:219-235, 1996.
[5] V.N. Sorokin, " Hermite Padé approximations for Nikishin's systems and the irrationality of $\zeta(3)$," *Uspekhi Matem. Nauk*, **49**, 2:167-168.1994.
[6] V.N. Sorokin, "The Apéry theorem", *Vestn. Mosk. Univ., Matem. Mekhan.*, 3:48-53, 1998.
[7] V.N. Sorokin, "Cyclic graphs and the Apéry theorem"', Uspekhi Matem. Nauk, **57**, 3(345):99-134, 2002.
[8] A.A. Gonchar, E.A. Rakhmanov and V.N. Sorokin, "Hermite-Padé approximations for systems of functions of Markov-type", *Matem. Sborn.*, **188**, 5:33-58, 1997.

LMPA Joseph Liouville, Centre Universitaire de la Mi-Voix, Bât H. Poincaré, 50 rue F. Buisson, BP 699, 62228 Calais Cedex
E-mail address: prevost@lmpa.univ-littoral.fr

Moscow State University M.V. Lomonosov Mechanics-Mathematics Faculty Department of Theory of Functions and Functional Analysis Vorobievy Gory 119899 Moscow RUSSIA
E-mail address: vnsormm@mech.math.msu.su