

Keldysh Institute • Publication search Keldysh Institute preprints • Preprint No. 114, 2017

ISSN 2071-2898 (Print) ISSN 2071-2901 (Online)

#### Varin V.P.

Factorial transformation for some classical combinatorial sequences

**Recommended form of bibliographic references:** Varin V.P. Factorial transformation for some classical combinatorial sequences // Keldysh Institute Preprints. 2017. No. 114. 31 p. doi:<u>10.20948/prepr-2017-114-e</u> URL: <u>http://library.keldysh.ru/preprint.asp?id=2017-114&lg=e</u>

# РОССИЙСКАЯ АКАДЕМИЯ НАУК ОРДЕНА ЛЕНИНА ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ ИМЕНИ М.В. КЕЛДЫША

V.P. Varin

# FACTORIAL TRANSFORMATION FOR SOME CLASSICAL COMBINATORIAL SEQUENCES

Moscow, 2017

# УДК 521.1+531.314

V.P. Varin. Factorial transformation for some classical combinatorial sequences. Preprint of the Keldysh Institute of Applied Mathematics of RAS, Moscow, 2017.

Factorial transformation known from Euler's time is a very powerful tool for summation of divergent power series. We use factorial series for summation of ordinary power generating functions for some classical combinatorial sequences. These sequences increase very rapidly, so OGFs for them diverge and mostly unknown in a closed form. We demonstrate that factorial series for them are summable and expressed in known functions. We consider among others Stirling, Bernoulli, Bell, Euler and Tangent numbers. We compare factorial transformation with other summation techniques such as Pade approximations, transformation to continued fractions, and Borel integral summation. This allowed us to derive some new identities for GFs and express integral representations of them in a closed form.

**Key words.** Factorial transformation; factorial series; continued fractions; Stirling, Bernoulli, Bell, Euler and Tangent numbers; divergent power series; generating functions.

В.П. Варин. Факториальное преобразование некоторых классических комбинаторных последовательностей. Препринт Института прикладной математики им. М.В. Келдыша РАН, Москва, 2017.

Факториальное преобразование известное со времен Эйлера является весьма эффективным инструментом суммирования расходящихся степенных рядов. Мы используем факториальные ряды для суммирования обычных производящих функций для некоторых классических комбинаторных последовательностей. Эти последовательности очень быстро растут, поэтому ОПФ для них расходятся и в основном неизвестны в замкнутой форме. Показано, что факториальные ряды для них суммируются и выражаются в известных функциях. Рассматриваются числа Стирлинга, Бернулли, Белла, Эйлера и тенгенциальные, и некоторые другие числа. Факториальное преобразование сравнивается с другими методами суммирования, такими как Паде-аппроксимации, преобразованием к цепным дробям, и интегральным суммированием Бореля. Это позволило вывести некоторые новые тождества для производящих функций и выразить их интегральные представления в явном виде.

Ключевые слова. Факториальное преобразование; факториальные ряды; цепные дроби; числа Стирлинга, Бернулли, Белла, Эйлера и тенгенциальные; расходящиеся степенные ряды; производящие функции.

© ИПМ им. М.В. Келдыша РАН. Москва, 2017 г.

#### § 1. Introduction

It seems strange, but convergence and divergence of power series (PS) seem to be a source of controversy in mathematics up to present days. Since the famous maxim by Abel that "Divergent series are the invention of the devil...", many prominent mathematicians had obtained important results that prove the contrary, i.e., they are as useful and meaningful as convergent ones.

The notion that if a formal PS diverges, then the offending object (a solution to an ODE, a first integral, a normalizing transformation, etc.) somehow does not exist is firmly engrained in popular mathematical thinking. For example, we read in a modern paper on convergence of some PS that "if the series converges, then this expansion corresponds to some solution to the equation". This statement rivals the famous saying by M.Twain that "Clams will lie quiet if music be played to them".

According to Borel-(Ritt) theorem, any formal PS is an asymptotic series for some analytical function in a sector. Of course, this function is not unique, since any algorithm we use to obtain an asymptotic series will necessarily miss an exponentially small (i.e., flat) function that can be added to that analytical function. Thus an asymptotic series is a class of equivalence of some functions rather then a function in itself.

The operative word in the above statement is "asymptotic", and convergence or divergence have nothing to do with it.

Hence if a formal PS obtained as a result of an asymptotic solution of a problem converges, then this does not necessarily mean that the solution (or whatever it is we are looking for) is analytical.

Here is a (trivial) example,  $y(x) = x + C \exp(-1/x)$ , that satisfies the equation  $x^2 y'(x) = y(x) - x + x^2$ . Thus the first integral of this equation is not analytical at the origin despite the obvious convergence of the PS. And an existence of a flat addition to a formal PS does not imply the divergence of the PS.

Although, as we will see, in some cases asymptotic series and flat functions are inextricably mixed together.

The question whether a divergent PS represents an analytical function is trivial by Borel theorem. On the other hand, the question what is the function that it does represent, is usually very difficult, and should be treated case by case.

This is what we will be doing in this paper.

We will consider some rather unusual power generating functions (GF) for some well known classical combinatorial sequences. As is well known, power GFs come (mostly) in two kinds, i.e., ordinary GFs, and EGFs, or exponential GFs (where the *n*-th number is divided by n!). This division by n! guarantees the convergence of EGF and produces an analytical generating function for the sequence.

In this paper, we will consider the sequences of Bernoulli, Bell, Stirling, Euler and Tangent numbers. The EGFs for these numbers are well known.

However, these numbers increase very rapidly, so ordinary GFs for them will necessarily diverge in any neighborhood of the origin (apart from GF for Stirling numbers of the second kind, which converges). Some of these GFs appear in combinatorial problems (see [1, p. 85]) as formal PS, and then continued fractions (CF) are used for their summation (see Sect. 2).

In a recent book a GF for Bernoulli numbers was introduced, as well as a number of related functions [2, p. 239]. It was demonstrated that these functions have remarkable combinatorial properties, and have some interesting applications.

In Sect. 2, we will fill some blanks left in the presentation of GF for Bernoulli numbers in [2]. Namely, it was demonstrated in [2, p. 241] that a formal PS for Bernoulli numbers satisfies a functional equation, and the PS is unique. The functional equation takes the form

$$\beta_1\left(\frac{x}{1-x}\right) - \beta_1(x) = x^2. \tag{1}$$

It was proved (see [3]) that this formal PS does represent an analytical function, since a convergent CF expansion of the GF exists. And it was established that this analytical GF and its generalizations are related to the Gamma function.

We will demonstrate just how close these functions are related to the  $\Gamma$  function, and find these functions in an explicit and closed form (in fact, in several forms). We will also produce an exponentially small addition to these functions, and thus find a general solution to Eq (1) and related equations.

In the following sections, we will treat Bell, Stirling, Euler and Tangent numbers as promised, and produce analytical GFs for these numbers that have divergent PS expansions.

For this, we will need to use summation of divergent PS. There are many forms of such summation including, ironically, the Abel summation. However, they are mostly applicable to convergent PS outside their radius of convergence.

One of the most powerful summation techniques is transformation of PS to a CF. Pade approximations of truncated series are closely related to this method. Often, it is possible to prove (constructively) this way the existence of an analytical function with a given power expansion. For example, Euler used this technique for summation of the series of factorials (see [4]).

However, it is seldom possible to express a CF in a closed form even if it exists. So, instead, we use another powerful summation technique closely related to Borel summation and Laplace transform, and called *factorial transformation* or transformation to factorial series (FS).

Factorial transformation is certainly known (see [5]), but surprisingly not well known. This deplorable fact was also mentioned in [6], where a survey of applications of FS can be found. So far, FS are used mainly as an acceleration of convergence technique for PS.

It turns out that classical combinatorial sequences have a close rapport with the factorial transformation in such a way that the resulting FS are expressed in a closed form. This is invariably linked with a hypergeometric pattern in a FS that can be recognized with the help of a CAS (we use Maple). The details will be given in appropriate places.

To conclude this introductory section, a few words on notation that we use.

There are many different notations for Stirling numbers and even for binomial coefficients. We avoid combinatorial notation with square and curly brackets and use traditional self-explanatory one that is also portable (i.e., from CAS to LaTeX).

### § 2. Bernoulli numbers and their GFs

Bernoulli numbers are ubiquitous in mathematics, and, consequently, they have been studied very thoroughly. The classical EGF for these numbers takes the form

$$\frac{x}{\exp(x) - 1} = \sum_{n=0}^{\infty} \frac{B(n)}{n!} x^n.$$
 (2)

Thus

$$B(0) = 1$$
,  $B(1) = -1/2$ ,  $B(2) = 1/12$ ,  $B(3) = 0$ ,...

Sometimes, another function for EGF is used, namely,  $x \exp(x)/(\exp(x) - 1)$ . Since these functions differ by x, it follows that, for this EGF, B(1) = 1/2, and no other difference. Multiplying both sides of (2) by  $\exp(x)$ , and collecting similar terms, we obtain the identity

$$\sum_{m=0}^{n-1} C(n,m) B(m) = 0, \quad n > 1,$$
(3)

where C(n, m) is the binomial coefficient.

As it was pointed out in [2] (in the Appendix written by Don Zagier), usually only one of generating functions, i.e., GF or EGF, is useful or interesting. For Bernoulli numbers, this is not the case. Consider a formal PS

$$\beta_1(x) = \sum_{n=0}^{\infty} B(n) x^{n+1},$$
(4)

which obviously diverges. Here and below, the equality is taken in asymptotical sense where appropriate.

We have used the same notation  $\beta_1(x)$  in (4) as in (1), since this PS satisfies (1) formally. It can be proved as in [2, p. 241] by substitution of rhs(4) in (1) and collecting similar terms using binomial identity. Then we would obtain the identity (3).

The solution is unique as a formal PS, but we know that a divergent PS does not represent a unique function. So where do we find a flat addition to (4)? It turns out that it is very simple in this case.

**Proposition 1.** Let f(x) = f(x+p) be any p-periodic function,  $p \in \mathbb{C}$ . Then the function

$$\tilde{\beta}_1(x) = \beta_1(x) + f\left(\frac{p}{x}\right)$$

is the general solution to the functional equation (1).

**Proof** is obvious, since f(p/x) satisfies the homogenous Eq (1); and the contrary, the relation b(x/(1-x)) = b(x) implies 1-periodicity of the function c(x) = b(1/x).

Thus, for example, the function  $\exp(2\pi i/x)$  gives a flat addition to GF (4) in the upper half-plane.

The problem remains how to calibrate, so to speak, an asymptotic series (4) such that a unique analytical function be produced that represents GF (4) for Bernoulli numbers.

It can be done as follows.

An obvious change of variable in (1) gives the equation

$$\beta_1(x) - \beta_1\left(\frac{x}{1+x}\right) = \frac{x^2}{(1+x)^2}.$$

We substitute here  $x \in \{1, 1/2, 1/3, ...\}$ , i.e., the harmonic series, and sum up these identities. They telescope, and we obtain

$$\beta_1(1) = \frac{\pi^2}{6} - 1.$$

Similarly, we can start with x = 1/n instead of x = 1, and obtain

$$\beta_1\left(\frac{1}{n}\right) = \sum_{k=n}^{\infty} \frac{1}{(k+1)^2} = \Psi(1, n+1),$$

where  $\Psi(x) = \frac{d \log \Gamma(x)}{dx}$  is digamma (or psi-) function, and  $\Psi(k, x) = \frac{d \log \Psi(k-1,x)}{dx}$ ,  $k \in \mathbb{N}, \Psi(0, x) = \Psi(x)$  are polygamma functions.

Now it is easy to guess that n can be replaced by 1/x in the above identity.

**Proposition 2.** The function  $\Psi(1, 1 + 1/x)$  is a special solution to Eq (1) with the asymptotic expansion (4).

**Proof.** We can either verify this fact by substitution using well known identities for polygamma functions, or we can differentiate twice the following obvious identity

$$\log \Gamma(t) - \log \Gamma(t+1) = -\log t,$$

with respect to t, then substitute there t = 1/x. As is well known,  $\Psi(1, 1+X) \approx 1/X$  as  $X \to +\infty$ , and  $\Psi(1, 2) = \pi^2/6 - 1$ . Thus  $\Psi(1, 1+1/x) \approx x$  as  $x \to +0$ , and it is calibrated (of fixed at a regular point) as needed.

As a corollary, we obtain asymptotic expansion of the function  $\Psi(1, 1 + X)$  at  $X = +\infty$  explicitly. Of course, this expansion is well known, but Bernoulli numbers are rarely mentioned.

In [2], several generalizations were given for the function  $\beta_1(x)$  that we briefly consider.

First, the functions  $\beta_k(x)$  were introduced that satisfy equations

$$\beta_k\left(\frac{x}{1-x}\right) - \beta_k(x) = k \, x^{k+1}, \quad k \in \mathbb{N}.$$
(5)

Similar argument as above gives formal PS for  $\beta_k(x)$ 

$$\beta_k(x) = \sum_{n=0}^{\infty} C(n+k-1,n) B(n) x^{n+k};$$
(6)

and it was established in [2, p. 242] that these functions satisfy recurrent ODEs

$$x^2 \frac{d\beta_k(x)}{dx} = k\beta_{k+1}(x), \quad k \in \mathbb{N}.$$

It turns out that these functions also have simple explicit forms

$$\beta_k(x) = \frac{(-1)^{k-1}}{(k-1)!} \Psi(k, 1+1/x), \quad k \in \mathbb{N},$$
(7)

which can be verified easily.

We cannot use the previous formula for generalization of the functions  $\beta_k(x)$  for integers k < 0 for obvious reasons, and the use of (6) for k < 1 simply gives Bernoulli polynomials  $(-1)^k B(-k, -1/x)$ .

So another combinatorial way to introduce these functions was used in [2]. Namely,

$$\gamma_k(x) = \sum_{n \ge \max(1, -k)}^{\infty} \frac{(n-1)!}{(n+k)!} B(n+k) x^n, \quad k \in \mathbb{Z}.$$

Then we obtain (see [2])

$$\gamma_{-k}(x) = (k-1)! \beta_k(x), \quad k > 0,$$

and, in particular,  $\gamma_{-1}(x) = \beta_1(x)$ .

The functions  $\gamma_k(x)$  satisfy recurrent ODEs

$$x^{2} \frac{d\gamma_{k}(x)}{dx} = \gamma_{k-1}(x) - \frac{B(k)}{k!} x, \quad k \ge 0,$$
(8)

as well as a series of functional equations similar to (5) (see [2, p. 243]).

The function  $\gamma_0(x)$  has the asymptotic expansion

$$\gamma_0(x) = \sum_{n=1}^{\infty} \frac{B(n)}{n} x^n.$$
(9)

We substitute k = 0 in (8) and obtain

$$x^2 \frac{d\gamma_0(x)}{dx} = \beta_1(x) - x,$$

from which we obtain

$$\gamma_0(x) = -\Psi(1/x) - \log(x) - x + \text{const.}$$
 (10)

Since asymptotic expansion of the function  $\Psi(X)$  as  $X \to \infty$  is known, then const = 0 in the above formula.

Now we can verify explicitly that

$$\gamma_0\left(\frac{x}{1-x}\right) - \gamma_0(x) = \log(1-x) + x,\tag{11}$$

in accordance with [2, Eq (A.14)].

The same argument gives explicit

$$\gamma_1(x) = \frac{\log x}{x} + \frac{1}{x} + \log \Gamma\left(\frac{1}{x}\right) - \frac{1}{2}\log x - \frac{1}{2}\log(2\pi),$$

where we already fixed an arbitrary constant using known asymptotic properties of  $\Gamma$  function.

And finally, we give the function  $\gamma_2(x)$ .

We need to solve the ODE

$$x^2 \frac{d\gamma_2(x)}{dx} = \gamma_1(x) - \frac{x}{12}$$

This gives

$$\gamma_2(x) = \frac{\log x}{2x} + \frac{1}{2x} + \frac{\log 2}{2x} + \frac{\log \pi}{2x} - \frac{\log x}{12} - \frac{\log x}{2x^2} - \frac{3}{4x^2} - \int_0^{1/x} \log \Gamma(t) + \text{const.}$$

And here the problem is how to fix an arbitrary constant.

We cannot use a known asymptotic expansion of  $\log \Gamma(t)$  at the origin, and then integrate it, since it is not a local problem (due to 1/x as upper limit). And we cannot use the ODE above or the functional equation [2, Eq (A.15)], since the arbitrary constant is cancelled.

A numerical estimate gives const  $\approx 0.24875447$ .

On the other hand

$$\gamma_2(x) = \int_0^x \left(\frac{\gamma_1(x)}{x^2} - \frac{1}{12x}\right) dx,$$

since the function under this integral  $\approx -x/360$ .

So one way or another, we obtain a quadrature.

It is clear that we cannot keep solving ODEs (8) without problems. Inevitably, new transcendents will appear that we know nothing about.

As an application, we compute some values of Riemann  $\zeta$  function with a totally unorthodox use of the Euler-Maclaurin summation formula. This formula takes the form (see [7, p. 518])

$$\sum_{k=m}^{n} f(k) = \int_{m}^{n} f(x) \, dx + \frac{1}{2} \left( f(m) + f(n) \right) + \sum_{j=1}^{N} \frac{B(2j)}{(2j)!} \left( f^{(2j-1)}(n) - f^{(2j-1)}(m) \right) + \frac{1}{(2N+1)!} \int_{m}^{n} B(2N+1, x - [x]) \, f^{(2N-1)}(x) \, dx.$$
(12)

Here B(n, x) is Bernoulli polynomial, [x] is the integer part of x,  $f^{(n)}(x)$  is the n-th derivative of an analytical function f(x) for which both sides of (12) make sense.

The formula (12) is precise, since it is with the remainder. Usually this formula is used for very effective numerical approximations.

First, we consider the harmonic series and compute the Euler's constant  $\gamma$ . This means we take f(x) = 1/x in (12).

We make the following operations on (12):

(a) drop the remainder altogether;

(b) subtract  $\log n$  from both sides;

(c) substitute m = 1 and  $n = \infty$ ;

(d) take the limit  $N \to \infty$ .

After some simplifications we obtain the identity

$$\gamma = \frac{1}{2} + \sum_{j=1}^{\infty} \frac{B(2j)}{2j}.$$

which seems absurd, but it is not (and belongs to Euler).

We recall the formula (9) and introduce a new function

$$\zeta_0(x) = \gamma_0(x) + 1 + \log(x) = 1 - x - \Psi(1/x).$$
(13)

Then  $\gamma = \zeta_0(1) = -\Psi(1)$ , which is correct.

Thus we simply used an asymptotic expansion (9), then put there x = 1. This substitution is not correct formally, but if we had transformed the series (9) into a CF, or simply used a closed form (10) already known and then put there x = 1, then this operation is perfectly legitimate.

Now we consider the Basel problem, i.e., we compute the value  $\zeta(2)$ . This means we take  $f(x) = 1/x^2$ .

We make the same operations on (12) omitting (b). Thus we obtain the identity

$$\zeta(2) = 1 + \sum_{j=0}^{\infty} B(j) = 1 + \beta_1(1) = 1 + \Psi(1, 1+1),$$

which is correct (see (4) and Prop. 2).

Several CF expansions of GF for Bernoulli numbers are given in [2, p. 258], and they converge at appropriate arguments. Thus these CF expansions also give CFs for  $\pi^2$  (this fact is missing in [2]).

Now we find the Apéry constant  $\zeta(3)$  in the same way, i.e., we take  $f(x) = 1/x^3$ . Similar manipulations as above give

$$\zeta(3) = 1 + \frac{1}{2} \sum_{j=1}^{\infty} (2j+1) B(2j).$$
(14)

Now we need to make sense of the rhs() of this formula.

**Proposition 3.** A finite number of multiplications/divisions by x, power transformations  $x \to x^r$ ,  $r \in \mathbb{Q}$ , and differentiation with respect to x of (4) gives asymptotic series that are expressed through polygamma functions.

**Proof** is obvious. First we apply prescribed transformations to the formal PS (4), and then to its representation  $\Psi(1, 1 + 1/x)$  in the same order.

Thus we obtain

$$\sum_{j=1}^{\infty} (2j+1) B(2j) x^{2j} = x - 1 - \frac{1}{x^2} \Psi(2, 1+1/x).$$

Putting here x = 1 and using this in (14) gives  $\zeta(3) = \zeta(3)$ , which is certainly true.

If we knew how to express the values of polygamma functions without  $\zeta$  function, we would have obtained a representation of the Apéry constant.

We believe that if a formal procedure keeps giving correct results, then there is some truth behind it. In this case, the liberal use of divergent series above can be easily justified. We introduce a set of functions  $\zeta_n(x), n \in \mathbb{N}$  by the following formula

$$\zeta_n(x) = 1 + \frac{(-1)^{n+1}}{n! \, x^n} \,\Psi\left(n, 1 + \frac{1}{x}\right) = 1 + \frac{1}{n!} \sum_{j=0}^{\infty} \left(\prod_{k=1}^{n-1} (j+k)\right) B(j) \, x^j, \quad (15)$$

which is a reformulation of asymptotic expansions of polygamma functions. Then we see that our informal procedure with the Euler-Maclaurin summation formula keeps giving  $\zeta_n(1) = \zeta(n+1), n \in \mathbb{N}$ , and  $\zeta_0(1) = \gamma$  as in (13) (compare also with (6)).

The divergent PS on the right of (15) can also be used numerically for small x quite effectively, just like original Euler-Maclaurin formula is very effective for numerical evaluations. For example, for n = 2 and x = 0.5, summation up to the smallest term in rhs(15) gives  $\approx 1.308414$ , which gives the error less than 0.0002 to the true value given by lhs(15). For x = 0.1, the error is less than  $10^{-24}$ .

Another way to use (15) numerically is to obtain a CF expansion of rhs(15). But this should be treated carefully, since slight modifications of the formula may produce very different CFs, which are far from being equivalent.

As an example, we find a new CF expansion for the Apéry constant.

For this, we take the function

$$\tilde{\zeta}_2(x) = \frac{x}{2} - \frac{1}{2} + \zeta_2(x) = 1 + \frac{1}{2}x^2 - \frac{1}{12}x^4 + \frac{1}{2}x^6 - \frac{3}{20}x^8 + \frac{5}{12}x^{10} + \dots,$$

and find its regular CF expansion

$$\tilde{\zeta}_{2}(x) = 1 + \frac{x^{2}}{4 + \frac{4x^{2}}{3 + \frac{2x^{2}}{1 + \frac{6x^{2}}{5 + \dots}}}} = b_{0} + \frac{a_{1}x^{2}}{b_{1} + \frac{a_{2}x^{2}}{b_{2} + \dots}} = b_{0} + \frac{\mathbf{K}}{\mathbf{k}_{1} + \frac{a_{n}x^{2}}{b_{n}}}, \quad (16)$$

where we have used Gauss' notation for CF expansion. Then we put x = 1 in this formula and obtain

$$\zeta(3) = b_0 + \mathop{\mathbf{K}}\limits_{n=1}^{\infty} \frac{a_n}{b_n},\tag{17}$$

where partial denominators  $b_n$ , n = 0, 1, 2, ... in (17) are 1, 4, 3, 1, 5, 1, 7, 1, 9, 1, 11, 1, 13, ..., which coincides with the partial denominators in CF expansion of tan(1) (see [8, A093178]), i.e., 1, 1, 1, 3, 1, 5, 1, 7, 1, 9, 1, 11, 1, 13, ... And partial numerators  $a_n$ , n = 1, 2, 3, ... in (17) are

 $1, 4, 2, 6, 9, 18, 24, 40, 50, 75, 90, 126, 147, 196, 224, 288, 324, 405, \ldots,$ 

which coincides with the sequence [8, A028724], that has interesting combinatorial properties, and that is given by a simple formula, i.e.

$$a(n) = \frac{1}{2} \left[ \frac{n}{2} \right] \left[ \frac{n-1}{2} \right] \left[ \frac{n-2}{2} \right] = \{0, 0, 0, 1, 2, 6, 9, 18, 24, 40, 50, \ldots\}, n \in \mathbb{N}.$$

We remark that similar manipulations with the polylog function

$$\operatorname{Li}_3(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^3}$$

produce very ugly CFs with extremely rapidly growing coefficients. And if we simply express  $\zeta(3)$  in CF form, then we obtain the partial numerators  $1, 1, 1, \ldots$ , and the partial denominators

$$1, 4, 1, 18, 1, 1, 1, 4, 1, 9, 9, 2, 1, 1, 1, 2, 7, 1, 1, 7, 11, \ldots,$$

which [8] does not recognize other than as it is [8, A013631].

Now we have to prove that CF (17) with the properties describe above actually gives  $\zeta(3)$ .

**Theorem 1.** The formula (17) is true.

**Proof.** First, we simplify (17) separating odd and even indices. Simple calculation gives  $a_n$  for n > 2, and  $b_n$  for  $n \ge 2$ :

$$a_n = \begin{cases} \frac{1}{16} (n-1) (n+1)^2, & n \text{ odd} \\ \frac{1}{16} n^2 (n+2), & n \text{ even}, \end{cases} \quad b_n = \begin{cases} n+1, n-1 \text{ odd} \\ 1, n-1 \text{ even} \end{cases}$$

Next, we make equivalence transform (see [9, p. 127]) of CF (17) and write it in the *simple* form, i.e.,

$$\zeta(3) = 1 + \mathbf{K}_{n=1}^{\infty} \frac{1}{b_n c_n} = 1 + \mathbf{K}_{n=1}^{\infty} \frac{1}{q_n},$$
(18)

where

$$c_1 = \frac{1}{a_1}, \quad c_2 = \frac{a_1}{a_2}, \quad c_n = \frac{1}{a_n c_{n-1}}, \quad n > 2.$$
 (19)

We have

$$\{c_1, c_2, \ldots\} = \{1, \frac{1}{4}, 2, \frac{1}{12}, \frac{4}{3}, \frac{1}{24}, 1, \ldots\}.$$

It turns out that the recurrence relation for  $c_n$  can be solved explicitly. We skip some calculations, and obtain

$$c_n = 4^{(-1)^{n+1}} \prod_{j=3}^n a_j^{(-1)^{n+1+j}}, \quad n \in \mathbb{N},$$

where empty product, as usual, equals 1. This formula is identical with (19) for n > 1, as it is easily verified. Now, skipping again some calculations, we obtain surprisingly simple formula for the partial denominators  $q_n = b_n c_n$  in (18)

$$q_n = \frac{8}{n+1}, n \text{ odd}; \qquad q_n = \frac{2(n+1)}{n(n+2)}, n \text{ even},$$

which gives

$$\{q_1, q_2, \ldots\} = \{4, \frac{3}{4}, 2, \frac{5}{12}, \frac{4}{3}, \frac{7}{24}, 1, \frac{9}{40}, \ldots\}.$$

It immediately follows that both CFs, (17) and (18), converge, since the sum of the partial denominators  $q_n$  gives two harmonic series, and thus diverges.

Well known formulas for convergents of a CF, i.e.,  $P_n/Q_n$  take the form

$$P_{-1} = 1, P_0 = 1, P_n = q_n P_{n-1} + P_{n-2}, n \ge 1$$
  

$$Q_{-1} = 0, Q_0 = 1, Q_n = q_n Q_{n-1} + Q_{n-2}, n \ge 1.$$
(20)

We again treat odd and even indices separately, and, skipping some cumbersome calculations, obtain the formula for  $Q_n$ :

$$Q_n = \frac{1}{2}(n+1)(n+3), n \text{ odd}; \qquad Q_n = \frac{1}{4}(n+2)^2, n \text{ even}.$$

Now we use the formula for convergents

$$\frac{P_n}{Q_n} - \frac{P_{n-1}}{Q_{n-1}} = \frac{(-1)^{n+1}}{Q_n Q_{n-1}},$$

which gives the convergent (as we already know) series as  $n \to \infty$ 

$$\frac{P_n}{Q_n} = 1 + \sum_{k=1}^n \frac{(-1)^{k+1}}{Q_k Q_{k-1}}.$$

Converting this to partial fractions, this sum evaluates to  $\zeta(3)$  as  $n \to \infty$ , which completes the proof.

Since denominators  $Q_n$  of convergents  $P_n/Q_n$  are found by simple formulas, one would expect the same for the numerators  $P_n$ , since they are found by the same recurrent formula (20). But this is not the case. We have

$$P_n = \begin{cases} 1+m(m+1)(\Psi(2,m+1)+2\zeta(3)) & n \text{ odd} \\ \frac{1}{2}(m+1)^2(\Psi(2,m)+2\zeta(3)) + \frac{(m+1)^2}{m^3} + \frac{m+2}{2(m+1)}, & n \text{ even}, \end{cases}$$

where m = [(n+1)/2]. This formula seems transcendental, but it is not, since

$$\Psi(2, m+1) + 2\zeta(3) = 2\sum_{n=1}^{m} \frac{1}{n^3}$$

Thus  $P_n$  are not integers.

As far as we know, there was only one CF for  $\zeta(3)$  with a regular pattern found by Stieltjes and Ramanujan (see [10, p. 46], and references there). Apéry apparently tried to use it for the proof of irrationality of  $\zeta(3)$ , but the convergence rate was found not fast enough.

Unfortunately, our CF (17) converges even more slowly, as numerical experiments reveal, so we would not use it further. On the other hand, there are ways of improving convergence of CFs, so CF (17) may be useful.

Another promising development would be to obtain similar representations for  $\gamma$  and  $\zeta(2n+1)$ , n > 1, but this is beyond the scope of the present paper.

Finally, we remark that what we did with the Euler-Maclaurin summation formula is very similar to the Ramanujan summation (see [11]). It seems worth the effort to try to "abuse" the Euler-Boole summation formula (see [11, p. 135]) in the same way, but again, we have to abstain for now.

#### § 3. Factorial transformation

First, we define the Pochhammer symbol

$$(x)_a = \Gamma(x+a)/\Gamma(x), \quad x, a \in \mathbb{C}.$$

For clarity

$$(x)_n = \prod_{k=0}^{n-1} (x+k), \quad n \in \mathbb{N}.$$

Then we have for  $n, m \in \mathbb{N}_0$ 

$$(x)_n = \sum_{m=0}^n (-1)^{n+m} S_1(n,m) x^m, \quad x^n = \sum_{m=0}^n (-1)^{n+m} S_2(n,m) (x)_m, \quad (21)$$

where  $S_1(n,m)$ ,  $S_2(n,m)$  are Stirling numbers of the first and second kind.

The Stirling numbers have fundamental combinatorial properties on a par with the factorial. Thus (21) can be proved based on their combinatorial definitions just like it was done in [2, Prop. 2.6]. But we take a shortcut and define Stirling numbers by the relations (21), i.e., we use them as GFs for Stirling numbers.

So we use here the *signed* Stirling numbers (unlike [2], where unsigned numbers are used). Thus  $S_1(n,m) = (-1)^{n+m} |S_1(n,m)|$ , and for  $S_2(n,m) \ge 0$  this makes no difference. Some formulas look simpler with signed numbers and some with unsigned, but we have to make a choice.

Now we define the function

$$Q(x,n) = (-1)^n \Gamma\left(1 + \frac{1}{x}\right) / \Gamma\left(n + 1 + \frac{1}{x}\right) = \frac{(-1)^n}{x} \left(\frac{1}{x}\right)_{n+1}^{-1}.$$

**Theorem 2.** For any formal PS, both series

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} Q(x,n) \sum_{m=0}^{n} (-1)^m S_1(n,m) a_m,$$
(22)

have the same asymptotic expansion as  $x \to 0$ ,  $\operatorname{Re}(x) > 0$ .

**Proof** can be found in the book [5], where FS were developed for asymptotic series at infinity. The formula (22) is simply a reformulation for PS at the origin.

But since [5] is a rarity, and uses obsolete notation, we give here a separate proof.

We can take (22) as the definition of *factorial transformation*, and rhs(22) as definition of FS.

First, we give some different forms of the function Q(x, n). We have

$$Q(x,n) = \sum_{k=0}^{n} \frac{(-1)^{n+k} C(n,k)}{(1+kx) n!} = \frac{(-1)^n}{x n!} B\left(n+1,\frac{1}{x}\right) = \sum_{j=0}^{\infty} (-1)^j S_2(j,n) x^j,$$
(23)

where, only in this formula, B(u, v) is the Euler Beta function (we will not use it).

All equalities in (23) are verified easily except the last one, which gives a GF for Stirling numbers of the second kind. The last formula is proved as follows.

We expand the fraction 1/(1+kx) in the first equality (23) in PS in x, then change the summation order. Then, comparing with the last equality (23), we obtain the identity

$$S_2(m,n) = \frac{1}{n!} \sum_{k=0}^n (-1)^{n+k} C(n,k) k^m.$$
(24)

This identity is well known (see [2, p. 30]), where it is proved by demonstration that both sides of (24) satisfy the same recurrent relation. This is a rather cumbersome but straightforward calculation (see [2, p. 32]).

The recurrent relation for the numbers  $S_2(m, n)$  is

$$S_2(n+1,m) = S_2(n,m-1) + m S_2(n,m),$$
(25)

 $S_2(0,0) = 1$ ,  $S_2(0,1) = S_2(1,0) = 0$ , and, obviously,  $S_2(n,m) = 0$  for n < m. Thus summation in the last sum in (23) starts at j = n.

The formula (25) is obtained from the definition (21) if we express  $x^{n+1}$  in two ways: (a) as in (21) with  $n \to n+1$ ; and (b)  $x^n$  by (21) times x. Then we use an obvious identity for Pochhammer symbol

$$x(x)_m = (x)_{m+1} - m(x)_m.$$

Now that we have this useful GF for Stirling numbers  $S_2(n,m)$ , we use it in (22) as Q(x,n) instead of the original one. Then, after some manipulations with the sums, we obtain the necessary and sufficient condition for (22) to be true, i.e.,

$$\sum_{j=1}^{n} S_2(n,j) S_1(j,m) = \delta_{n,m},$$

where  $\delta_{n,m}$  is the Kronecker symbol.

This is a well known orthogonality condition for Stirling numbers, and it follows directly from the definition (21) if we use the first formula in the second.

Since the GF Q(x,n) for Stirling numbers  $S_2(n,m)$  is a rational function, we deduce that  $S_2(n,m)$  (unlike  $S_1(n,m)$ ) cannot grow too rapidly as  $n \to \infty$ .

It was shown in [12] that if FS at infinity (i.e., x replaced by 1/x in (22)) converges, then it converges for 0 < const < Re(x). This means that rhs(22) (if convergent) converges in a circle of some radius  $R_0$  with the center at  $x_0 = R_0 > 0$ .

As we proved, both sides of (22) give the same asymptotic PS as  $x \to 0$ . Moreover, it was proved in [12] that the remainders after truncation at the *n*-th term for both series in (22) have the same order of smallness, i.e., both PS and FS give approximations of the same order to the value of the function.

Further, it was shown in [12] that "most of the ordinary functions of analysis" can be expressed as convergent FS. The author went so far as to claim that "the theory of asymptotic series is reduced to the level of the theory of convergent series".

We would not go that far in our assessments, but as we will see, FS (22) gives a very powerful summation tool on a par with CF if not better.

The last statement can be made more specific.

It is well known that convergents to a CF are equivalent to a (N-1, N), (N, N), or (N, N-1) Pade approximant of the original PS.

It is easy to see that the partial FS (22) to the order N is equal to a rational function with the denominator

$$D_N(x) = \left| \frac{x^N}{Q(x,N)} \right| = \prod_{k=1}^N (1+kx),$$

which is a polynomial of degree N, and the numerator which is also a polynomial of degree N.

Thus a partial FS is similar both to Pade approximation and to a partial CF approximation. But, unlike them both, FS expansion almost certainly converges, and can frequently be expressed in a closed form as we will see shortly.

We remark that FS were known for at least a 100 years before Pade approximations. Then, as we mentioned, they were mostly forgotten. But in the last 50 years or so, powerful algorithms were developed for summation of binomial sums and for hypergeometric sums in general (see [13]). This means that if a series has a hypergeometric pattern, then it can be recognized with the help of a modern CAS, and the series summed up in a closed form.

As an example, we consider the following sum

$$S(x) = \sum_{n=0}^{\infty} \frac{Q(x,n)}{n!},$$

which is not valid at first glance, since the functions Q(x, n) have singularities at x = -1/k, k = 1, 2, ..., n. Thus singularities keep piling up, and this sum cannot converge to an analytical function at the origin. But it converges nonetheless for  $\operatorname{Re}(x) > 0$ , since it is easy to see that S(x) is a FS for the following asymptotic series

$$\sum_{m=0}^{\infty} \left( \sum_{n=0}^{m} \frac{S_2(m,n)}{n!} \right) (-1)^m x^m = F([], [1+1/x], -1) = \Gamma\left(1+\frac{1}{x}\right) J_{1/x}(2), (26)$$

where F([], [], z) is a generalized hypergeometric function, and  $J_v(z)$  is the Bessel function of the first kind.

We remark that CAS as a rule cannot expand Bessel or hypergeometric functions in PS with respect to parameters. But there is a trick for this. If we take F([], [1+1/x], z), expand it in PS in z, then substitute z = -1, and then expand in PS in x, then we obtain lhs(26).

Thus (26) can be read both ways, i.e., from left to right it is a GF for the sequence

$$a(m) = (-1)^m \sum_{n=0}^m \frac{S_2(m,n)}{n!}, \quad m \in \mathbb{N}_0,$$

and from right to left it is an asymptotic expansion for the Bessel function with respect to a parameter.

If fact, since Q(0,n) = 0, n > 0, and  $Q(x,n) \simeq (-1)^n/n!$  as  $x \to \infty$ , FS indeed have a tendency to converge for  $\operatorname{Re}(x) > 0$ . We give another example

$$\sum_{n=0}^{\infty} Q(x,n) \, n! = \sum_{m=0}^{\infty} \left( \sum_{n=0}^{m} S_2(m,n) \, n! \right) \, (-1)^m \, x^m = F([1,1], [1+1/x], -1).$$

Now we construct FSs for Bernoulli numbers.

But first we need to point out that a divergent asymptotic PS may produce different functions for  $\operatorname{Re}(x) > 0$  and for  $\operatorname{Re}(x) < 0$  (or in other sectors, i.e., a situation similar to the Stokes phenomenon). In fact, this property can be taken as a characteristic one for divergent series. Thus we produce two FS for Bernoulli numbers. First, we consider

$$G(x) = \sum_{n=0}^{\infty} (-1)^n B(n) x^n.$$
 (27)

As we will see,  $\beta_1(x) \neq x G(-x)$  (see (4)).

We apply Theorem 2, and obtain

$$G(x) = \sum_{n=0}^{\infty} Q(x,n) \sum_{m=0}^{n} S_1(n,m) B(m) = \sum_{n=0}^{\infty} Q(x,n) a_0(n).$$

Now we have to make sense of the inner sum in this formula, i.e.,  $a_0(n)$ , which is, in fact, a *Stirling transform* of the sequence B(m), m = 0, 1, ...

Since this is a demonstration of a technique, we will go in some detail.

We consider the sequences

$$\{a_1(n)\} = \{(n+1)! a_0(n)\}; \ \{a_2(n)\} = \left\{\frac{a_1(n+1)}{a_1(n)}\right\}, \ n \in \mathbb{N}_0,$$

then we observe that  $a_2(n) = -(n+1)^2$ . Going back, we obtain the identity

$$a_0(n) = \sum_{m=0}^n S_1(n,m) B(m) = (-1)^n \frac{n!}{n+1}.$$
 (28)

To prove (28), we take inverse Stirling transform of both sides of (28) and obtain

$$B(n) = \sum_{m=0}^{n} S_2(n,m) (-1)^m \frac{m!}{m+1},$$

which is a known identity (see [2, Theorem 2.8]).

Thus we obtain

$$G(x) = F([1,1,1], [2,1+\frac{1}{x}], 1) = \frac{1}{x} \Psi\left(1,\frac{1}{x}\right),$$
(29)

where the hypergeometric function converges at the argument z = 1, since 1 + 1 + 1 < 2 + (1 + 1/x) implies x > 0. The second equality in (29) will be explained later in this section.

Thus G(x) is defined in the right half-plane as hypergeometric function and analytically continued to the left half-plane by rhs(29).

Now we take GF  $\beta_1(x)$  in (4) as an asymptotic series and apply Theorem 2. We obtain FS

$$\beta_1(x) = x + x \sum_{n=1}^{\infty} Q(x,n) \sum_{m=1}^{n} (-1)^m S_1(n,m) B(m),$$

where summation is started at n = 1 for convenience.

Similar but more lengthy manipulations with the inner sum reveal the identity

$$\sum_{m=1}^{n} (-1)^n S_1(n,m) B(m) = (-1)^{n+1} \frac{(n-1)!}{n+1}, \quad n \in \mathbb{N},$$
(30)

which implies

$$B(n) = \sum_{m=1}^{n} S_2(n,m) (-1)^{n+m+1} \frac{(m-1)!}{m+1}, \quad n \in \mathbb{N}.$$

These identities can be proved as before (see [2, Theorem 2.8]).

Combining these formulas with Prop. 2, we obtain (along with the promised FS) a hypergeometric representation of the polygamma function

$$\Psi\left(1,1+\frac{1}{x}\right) = x - \frac{x^2}{2(1+x)}F\left([1,1,2],\left[3,2+\frac{1}{x}\right],1\right),\tag{31}$$

which is new (as far as we know).

Convergency test for the hypergeometric function in (31) gives x < -1, and 0 < x, which is in accordance with lhs(31).

However, rhs(31) is not the same function as lhs(31), since the latter is defined, say, at x = -2/3, but rhs(31) is not. So lhs(31) is an analytical continuation of rhs(31). In addition, rhs(31) converges very slowly for big x, so this formula is more useful from right to left than otherwise.

Now we recall the original functional Eq (1). Since both sides of (31) satisfy Eq (1) formally, we can take it either as a hypergeometric transformation formula (probably a new one), or we simply obtain another representation of the polygamma function

$$\Psi\left(1,1+\frac{1}{x}\right) = x - \frac{x^2}{2(1-x)} \left(2x - F\left([1,1,2], \left[3,1+\frac{1}{x}\right],1\right)\right), \quad (32)$$

which converges for Re(x) > 0 (x = 1 is not a singularity: F([1, 1, 2], [3, 2], 1) = F([1, 1], [3], 1) = 2).

Using the substitution  $x \to x/(1+x)$  in (1) repeatedly, we obtain a sequence of equations such that their lhs() telescope. Thus we obtain another representation

$$\beta_1(x) = x^2 \sum_{n=1}^{\infty} \frac{1}{(1+nx)^2} = \Psi\left(1, 1+\frac{1}{x}\right).$$

Similar summation works for  $\beta_k(x)$  in (5), and even for  $\gamma_0(x)$  (see (10), (11)), but not further. Thus we can write other representations of lhs(32) through  $F([1,1,2], [3, n+1/x], 1), n \in \mathbb{N}.$  Finally, capitalizing on the property of Bernoulli numbers, we see that  $G(x) = x + \beta_1(x)/x = x + \Psi(1, 1 + 1/x)/x$ , which gives rhs(29). Thus we have yet another hypergeometric transformation formula

$$x F([1,1,1], [2,x+1], 1) = 1 + x - \frac{x}{2(x+1)} F([1,1,2], [3,2+x], 1).$$

where we put  $x \to 1/x$ . This formula transforms poorly convergent function on the left as  $x \to 0$  to better convergent one on the right.

# § 4. Bell numbers

The two well known EGFs for Stirling numbers are

$$\sum_{n=0}^{\infty} \frac{S_1(n,m)}{n!} x^n = \frac{\log^m (1+x)}{m!}, \quad \sum_{n=0}^{\infty} \frac{S_2(n,m)}{n!} x^n = \frac{(\exp(x) - 1)^m}{m!}.$$
 (33)

If we sum up the first of these sums from m = 0 to  $m = \infty$ , we will obtain  $S_1(0,0) = 1$ ,  $S_1(1,0) + S_1(1,1) = 1$ , and the following identity

$$\sum_{m=0}^{n} S_1(n,m) = 0, \quad n > 1.$$

But the summation of the second EGF in (33) produces the identity

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left( \sum_{m=0}^{n} S_2(n,m) \right) x^n = \exp(\exp(x) - 1),$$
(34)

which is, in fact, a well known EGF for Bell numbers, which we denote B(n) (not to confuse with Bernoulli numbers B(n)).

Thus Bell numbers are less "fundamental" than the Stirling numbers. Still, they are very important in combinatorics, so we will try to produce an ordinary GF for these numbers.

The first attempt at summation of the GF for Bell numbers as a FS fails. We only give it as an example that FS not always converge.

Thus we consider

$$U(x) = \sum_{n=0}^{\infty} B(n) x^n = 1 + \sum_{n=1}^{\infty} Q(x,n) \sum_{m=1}^{n} (-1)^m S_1(n,m) B(m),$$

where summation in FS starts at n = 1 for convenience.

Comparative investigation of integer sequences, which is done with the help of online encyclopedia [8], gives the representation

$$U(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{1 + x (n+1)} Q(x, n) n! L_n(-1),$$

where  $L_n(x)$  is the Laguerre polynomial.

The sequence  $n! L_n(-1)$  is recognized by [8], and we have

$$L_n(-1) \simeq \frac{\exp(-1/2 + 2n^{1/2})}{2n^{1/4}\pi^{1/2}}, \quad n \to \infty$$

(see [8, A002720]). Here  $y \simeq x$  means y = x (1 + o(1)) as  $x \to \infty$ . Thus U(x) does not exist as an FS.

The second option is

$$W(x) = \sum_{n=0}^{\infty} (-1)^n \operatorname{B}(n) x^n = \sum_{n=0}^{\infty} Q(x,n) \sum_{m=0}^n S_1(n,m) \operatorname{B}(m) = \sum_{n=0}^{\infty} Q(x,n),$$

where the last equality follows from the representation (34) and orthogonality property of Stirling numbers.

Thus we have

$$W(x) = \exp(-1) F\left(\left[\frac{1}{x}\right], \left[1 + \frac{1}{x}\right], 1\right) = \frac{1}{x} \int_0^1 \exp(t - 1) t^{1/x - 1} dt, \qquad (35)$$

where rhs(35) follows from the integral representation of  $\Gamma$  and incomplete  $\Gamma$  functions, that also express W(x).

To verify lhs(35) with a CAS, we need the trick that was already described. Here we take  $\exp(-z) F([1/x], [1+1/x], z)$ , expand it in PS in z, then substitute z = 1, and then expand in PS in x.

The GF (35) also explains why the first attempt at construction of a GF for Bell numbers failed. The function W(x) as an integral rhs(35) is not defined for Re(x) < 0. But W(x) as lhs(35) is an analytical function defined in the complex plane except the poles  $x_n = -1/n$ ,  $n \in \mathbb{N}$ . Thus W(x) is also valid for x < 0,  $x \neq x_n$ , and, technically, gives the GF U(x), except that it does not provide asymptotic expansions in any sector with the line  $x \in (-\infty, 0)$ .

#### § 5. Euler and Tangent numbers

These numbers are put together for a very good reason, as we will see. The two well known EGFs for these numbers are

$$\sum_{n=0}^{\infty} \frac{E(n)}{n!} x^n = \frac{1}{\cosh(x)}, \quad \sum_{n=0}^{\infty} \frac{T(n)}{n!} x^n = 1 + \tanh(x).$$
(36)

There are several rather complicated explicit formulas for Euler numbers E(n) that can be found in literature (that we skip).

Tangent numbers T(n) are usually defined slightly differently than in (36), but this again is done for a purpose. We have

$$T(n) = \frac{4^{n+1} - 2^{n+1}}{n+1} B(n+1), \quad n \in \mathbb{N}_0,$$

where, and only in this formula, Bernoulli number B(1) = 1/2 (see (2)).

**Proposition 4.** The sequences of Euler and Tangent numbers are binomial transforms of each other, i.e., for  $n \in \mathbb{N}_0$ , we have

$$T(n) = \sum_{m=0}^{n} C(n,m) \, (-1)^m \, E(m), \quad E(n) = \sum_{m=0}^{n} C(n,m) \, (-1)^m \, T(m).$$
(37)

**Proof.** An obvious identity for EGFs (36), namely

$$\exp(-x)\left(1 + \tanh(x)\right) = \frac{1}{\cosh(x)}$$

gives the second equality (37) either by the well known multiplication rule for EFGs, or by simply collecting similar terms. Then the first equality (37) follows automatically by the property of binomial transform being an involution.

We denote formal GFs for Euler and Tangent numbers as  $E_g(x)$  and  $T_g(x)$ . First, we consider FS for Euler numbers. Theorem 2 gives

$$E_g(x) = \sum_{n=0}^{\infty} E(n) x^n = \sum_{n=0}^{\infty} Q(x,n) \sum_{m=0}^{n} (-1)^m S_1(n,m) E(m) = \sum_{n=0}^{\infty} Q(x,n) a_0(n),$$

where  $\{a_0(n)\}\$  is the Stirling transform of the sequence  $\{(-1)^n E(n)\}$ .

The sequence  $\{a_0(n)/n!\}$ , as it turns out, has a recognizable pattern

$$\left\{\frac{a_0(n)}{n!}\right\} = \left\{1, 0, \frac{-1}{2}, \frac{1}{2}, \frac{-1}{4}, 0, \frac{1}{8}, \frac{-1}{8}, \frac{1}{16}, 0, \frac{-1}{32}, \frac{1}{32}, \frac{-1}{64}, 0, \ldots\right\}, \ n \in \mathbb{N}_0,$$

which can be summarized as follows

$$a_0(n) = \frac{|(n \mod 4) - 1| (-1)^{[(n+1)/2] + [n/4]}}{2^{[(n+1)/2]}} n!, \quad n \in \mathbb{N}_0$$

Thus we need to consider four different cases of hypergeometric summation dependent on  $(n \mod 4)$ .

In order to simplify the formulas, we introduce the functions

$$g:(x,k) \longrightarrow \frac{(k+1)x+1}{4x}, \frac{(k+2)x+1}{4x}, \frac{(k+3)x+1}{4x}, \frac{(k+4)x+1}{4x}, \frac{(k+4)x+1}$$

which produces a sequence of 4 rational functions, and

$$f:(x,k) \longrightarrow F([1,g(1,k-1)],[g(x,k)],-1/4),$$

which gives a hypergeometric function dependent on g(x, k).

Then, skipping some cumbersome calculations and collecting formulas, we obtain

$$E_g(x) = 1 - Q(x,2) f(x,2) + 3 Q(x,3) f(x,3) - 6 Q(x,4) f(x,4).$$

For clarity, we give just one term of this expression, Q(x, 2) f(x, 2),

$$\frac{x^2}{(x+1)(2x+1)} F\left(\left[\frac{3}{4}, 1, 1, \frac{5}{4}, \frac{3}{2}\right], \left[\frac{3x+1}{4x}, \frac{4x+1}{4x}, \frac{5x+1}{4x}, \frac{6x+1}{4x}\right], \frac{-1}{4}\right).$$

Except for the obvious singularities  $x_n = -1/n$ ,  $n \in \mathbb{N}$ , the GF  $E_g(x)$  is well defined in the complex plane due to the argument z = -1/4, which guarantees rapid convergence.

Now we consider FS for Tangent numbers. Theorem 2 gives

$$T_g(x) = \sum_{n=0}^{\infty} T(n) x^n = \sum_{n=0}^{\infty} Q(x,n) \sum_{m=0}^{n} (-1)^m S_1(n,m) T(m) = \sum_{n=0}^{\infty} Q(x,n) b_0(n),$$

where  $\{b_0(n)\}\$  is the Stirling transform of the sequence  $\{(-1)^n T(n)\}$ .

The sequence  $\{b_0(n)/2/n!\}$  also has a recognizable pattern. In fact, it is the same as for Euler numbers only shifted by 2 and with minus sign

$$\left\{\frac{b_0(n)}{2\,n!}\right\} = \left\{\frac{1}{2}, \frac{-1}{2}, \frac{1}{4}, 0, \frac{-1}{8}, \frac{1}{8}, \frac{-1}{16}, 0, \frac{1}{32}, \frac{-1}{32}, \frac{1}{64}, 0, \ldots\right\}, \ n \in \mathbb{N}_0,$$

which can be summarized as follows

$$b_0(n) = \frac{\left| (n-2 \mod 4) - 1 \right| (-1)^{\left[ (n-1)/2 \right] + \left[ (n-2)/4 \right]}}{2^{\left[ (n+1)/2 \right]}} n!, \quad n \in \mathbb{N}_0.$$

Thus, as before, we need to consider four different cases of hypergeometric summation dependent on  $(n \mod 4)$ .

Skipping some cumbersome calculations and collecting formulas, we obtain

 $T_g(x) = 1 - Q(x, 1) f(x, 1) + Q(x, 2) f(x, 2) - 6 Q(x, 4) f(x, 4).$ 

As it is often the case, asymptotic PS are very useful for numerical evaluation of a function at a singularity even if they diverge. For example, summation of PS for Euler numbers up to the smallest term for x = 0.1, which takes 16 terms, gives the error less then  $1.01 \times 10^{-6}$  compared to the true value  $E_g(0.1) = 0.990449430463732220$ . The same test for Tangent PS for x = 0.1, which takes 14 terms, gives the error less then  $0.95 \times 10^{-6}$  compared to the true value  $T_g(0.1) = 1.098138472266119760$ .

At the same time, GFs  $E_g(x)$  and  $T_g(x)$  give an example of how differently formal PS behave before and after summation. Such an example, we believe, was not given before.

**Proposition 5.** Let the sequences  $\{g_n\}$  and  $\{h_n\}$  be binomial transforms of each other, and let g(x) and h(x) be their respective formal GFs, i.e.,

$$g(x) = \sum_{n=0}^{\infty} g_n x^n, \quad h(x) = \sum_{n=0}^{\infty} h_n x^n,$$

then

$$g(x) = \frac{1}{1-x} h\left(\frac{-x}{1-x}\right),\tag{38}$$

considered as formal PS transformation.

**Proof.** The identity (38) for GFs is, of course, well known (see [14], where an obsolete form of binomial transform is used). It is clearly an involution and easily confirmed with formal manipulation of PS.  $\blacksquare$ 

Thus one would expect the same property (38) hold for  $E_g(x)$  and  $T_g(x)$ , except that their FS are not defined for  $\operatorname{Re}(x) < 0$ .

We denote alternative versions of GFs for Euler and Tangent numbers as  $E_g^-(x)$  and  $T_g^-(x)$ , i.e.,

$$E_g^{-}(x) = \sum_{n=0}^{\infty} (-1)^n E(n) x^n, \quad T_g^{-}(x) = \sum_{n=0}^{\infty} (-1)^n T(n) x^n$$

keeping in mind that FS for them are still defined for  $\operatorname{Re}(x) > 0$ .

Actually,  $E_g^-(x) = E_g(x)$ , and

$$T_g^{-}(x) = 1 + Q(x,1) f(x,1) - Q(x,2) f(x,2) + 6 Q(x,4) f(x,4),$$

i.e,  $T_g(x) + T_g^-(x) = 2$ , which follows from  $\cosh(x)$  being even, and  $\tanh(x)$  being odd (see (36)).

Thus, instead of (38), we have two identities

$$E_g(x) = \frac{1}{1+x} T_g\left(\frac{x}{1+x}\right), \quad T_g^-(x) = \frac{1}{1+x} E_g^-\left(\frac{x}{1+x}\right),$$

which are not involutions, i.e., not reversable in the sense  $E_g \leftrightarrow T_g$ .

We can solve this system of four functional equations and obtain the identities

$$E_g(x) = \frac{2}{1+x} - \frac{1}{1+2x} E_g\left(\frac{x}{1+2x}\right), \ T_g(x) = 2 - \frac{1}{1+2x} T_g\left(\frac{x}{1+2x}\right).$$
(39)

These identities can be considered either as hypergeometric transformation formulas, or as functional equations that generate Euler and Tangent numbers just like Eq (1) generated Bernoulli numbers.

#### § 6. Stirling numbers of the first kind

Stirling numbers of the second kind, that are considered in [14] for some strange reason as being more "fundamental" than their counterpart of the first kind, already have very useful GFs Q(x, n) (see (23)). Here we construct similar GFs for sequences  $\{S_1(n, m), n \in \mathbb{N}_0\}, m \in \mathbb{N}_0$ .

The functions Q(x, n) are rational. Hence changing sign of x basically makes no difference, since their PS converge at the origin. For GFs for  $S_1(m, n)$ , this is manifestly not the case, since, for example,  $S_1(n, 1) = (-1)^{n-1} (n-1)!$ , n > 0. Thus their PS diverge at the origin, and we may need to consider two separate FS summations as we did for the Bell numbers.

Let us denote formal GFs for the sequences  $\{S_1(n,m), n \in \mathbb{N}_0\}$  as

$$Y_m(x) = \sum_{n=0}^{\infty} S_1(n,m) x^n, \quad m \in \mathbb{N}_0.$$
 (40)

Note that the summation here starts actually at n = m, i.e.,  $Y_m(x) \simeq x^m$ .

Thus, obviously,  $Y_0(x) \equiv 1$ , and  $Y_1(x)$  satisfies the famous Euler equation (see [4])

$$x^2 Y_1'(x) + Y_1(x) = x,$$

that was used countless times as an example of an ODE having a divergent asymptotic PS solution.

Thus we already have a summed up GF  $Y_1(x)$  as

$$Y_1(x) = \exp\left(\frac{1}{x}\right) \operatorname{Ei}\left(1, \frac{1}{x}\right) = \exp\left(\frac{1}{x}\right) \int_{1}^{\infty} \frac{1}{t} \exp\left(\frac{-t}{x}\right) dt, \quad \operatorname{Re}(x) > 0,$$

where Ei() is the integral exponent function.

**Proposition 5.** The GFs  $Y_m(x)$  satisfy the following recurrent ODEs

$$x^{2} Y'_{m}(x) + Y_{m}(x) = x Y_{m-1}(x), \quad m \in \mathbb{N}.$$
 (41)

**Proof.** We already have the base of induction. Let (41) hold for k < m. Substitution of (40) in (41) gives the recurrent relation for the numbers  $S_1(n, m)$ 

$$S_1(n+1,m) = S_1(n,m-1) - n S_1(n,m),$$

which follows directly from the definition (21).  $\blacksquare$ 

The general solution to (41) is

$$Y_m(x) = \exp\left(\frac{1}{x}\right) \left(\int_0^x \frac{1}{t} Y_{m-1}(t) \exp\left(\frac{-1}{t}\right) dt + \text{const}\right),\tag{42}$$

where an arbitrary const = 0, since  $Y_m(x) \approx x^m$ . In particular,

$$Y_2(x) = \exp\left(\frac{1}{x}\right) \int_0^x \frac{1}{t} \operatorname{Ei}\left(1, \frac{1}{t}\right) \, dt.$$
(43)

Just for the record, the GFs Q(x, n) also satisfy recurrent ODEs

$$x(1+nx)Q'(x,n) + x^2Q'(x,n-1) = Q(x,n), \quad n \in \mathbb{N},$$

that can be used to prove the first equality in (23).

The ODEs (41) and their solutions  $Y_m(x)$  give an example of how (almost) inextricably flat functions can be linked to analytical functions represented by divergent PS.

Problems begin when we try to express  $Y_m(x)$  in explicit form.

We have a head start with  $Y_1(x)$ , since there are three century of study behind it. But let us express it first as a general solution to Euler ODE and then find an arbitrary constant. We have

$$Y_1(x) = \exp\left(\frac{1}{x}\right) \left(\int \frac{1}{x} \exp\left(\frac{-1}{x}\right) dx + C\right) = \exp\left(\frac{1}{x}\right) \left(\int \frac{1}{x} F\left([], [], \frac{-1}{x}\right) dx + C\right)$$
$$= \exp\left(\frac{1}{x}\right) \left(C + \log x + \frac{1}{x} F\left([1, 1], [2, 2], \frac{-1}{x}\right)\right), \quad C = -\gamma,$$

where  $\gamma$  is, not surprisingly, the Euler constant. In fact,  $Y_1(1) = \delta$  is the Euler-Gompertz constant (see [10, p. 423]), and we have

$$\delta = \mathbf{e} \operatorname{Ei}(1,1) = \mathbf{e} \left( F([1,1],[2,2],-1) - \gamma \right) = 0.59634736232319407434.$$

Thus we face a non-local problem of finding an asymptotics of an integral, or an asymptotics of a hypergeometric function at infinity in order to match the local asymptotics of the function at the origin (compare with  $\gamma_2(x)$  in Sect. 2).

We can move exactly one step further along this path before we run out of explicit arbitrary constants. We have

$$Y_2(x) = \mathbf{e}^{1/x} \left( \frac{\pi^2}{12} + \frac{\gamma^2}{2} - \gamma \log x + \frac{\log^2 x}{2} - \frac{1}{x} F\left( [1, 1, 1], [2, 2, 2], \frac{-1}{x} \right) \right), \quad (44)$$

due to the integral (43), which can be evaluated explicitly at x = 1.

Let us denote

$$[k]_m = [k, k, \dots, k], \quad (m \text{ times}), \quad k = 1, 2.$$

**Proposition 6.** The GFs  $Y_m(x)$ ,  $m \in \mathbb{N}$  have explicit forms

$$Y_m(x) = \mathbf{e}^{1/x} \left( C_m + R_m(\log x) + \frac{(-1)^{m-1}}{x} F\left( [1]_{m+1}, [2]_{m+1}, \frac{-1}{x} \right) \right), \quad (45)$$

where  $C_m$  are uniquely defined constants expressed in quadratures, and  $R_m(x)$  are polynomials of degree m without constant terms.

**Proof** is easily obtained by induction.

Since the function  $\exp(1/x)$  is flat on the wrong side of the origin, we face a paradoxical situation, i.e., an exact formula is not necessarily a good thing. Namely, we cannot compute the functions  $Y_m(x)$  too close to the origin even if we knew the constants  $C_m$  exactly (and, in a sense, we do know them). This is because irrational numbers cannot be known exactly from the computational viewpoint. Thus a small error for small x > 0 is multiplied by  $\exp(1/x)$  and ruins the result. So the original divergent PS (40) or partial FS for them are better for computation for small x.

For example, we take  $Y_2(x)$  and x = 0.1. The exact result to 16 decimal places (all correct) is  $Y_2(0.1) = 0.7770348214281297 \times 10^{-2}$ . Summation of PS (40) to the smallest term truncates at n = 10 and gives the error less than  $5.2 \times 10^{-5}$ . If fact the PS for  $Y_2(x)$  is alternating, so the series can be accelerated. Factorial transformation does just that and produces rapidly convergent FS. Ten terms of FS give an error less than  $2 \times 10^{-6}$ , and 20 terms give an error less than  $1.1 \times 10^{-9}$ .

If we take x = 0.01 and use ordinary double float arithmetic ( $\approx 16$  decimal places), then both PS (40) and FS for it give  $Y_2(0.01) = 0.97105257566526 \times 10^{-4}$  (all digits correct), but in exact formula (44), all digits are lost.

Unfortunately, Stirling transforms of sequences  $\{(-1)^n S_1(n,m), n \in \mathbb{N}\}, m \in \mathbb{N}$  do not appear to have a hypergeometric pattern. Thus FS for PS (40) are most likely not summable in a closed form. For example, if we take the sequence  $\{(-1)^n S_1(n,1), n \in \mathbb{N}\}$ , then (thanks to [8])

$$\log(1 - \log(1 - x)) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \left( \sum_{m=1}^n (-1)^m S_1(n, m) S_1(m, 1) \right) x^n,$$

and since

$$\log(1 - \log(1 - x)) = x F([1, 1], [2], x) F([1, 1], [2], -x F([1, 1], [2], x)),$$

it can hardly be expected to simplify to a hypergeometric form.

An alternative form of FS (with  $(-x)^n$  instead of  $x^n$ ) does not produce anything new except

$$\log(1 + \log(1 + x)) = \sum_{n=1}^{\infty} \frac{1}{n!} \left( \sum_{m=1}^{n} S_1(n, m) S_1(m, 1) \right) x^n.$$

But there appears to be another transformation entirely different from FS.

We cannot fail to notice a certain symmetry between two Stirling transforms on the one hand and the functions  $Y_n(x)$  and Q(x, n) on the other. Thus we have

**Theorem 3.** For any formal PS, both series

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} Y_n(x) \sum_{m=0}^n S_2(n,m) a_m,$$
(46)

have the same asymptotic expansion as  $x \to 0$ ,  $\operatorname{Re}(x) > 0$ .

**Proof** is exactly the same as for the Theorem 2, since the most difficult part of that proof was to establish the last equality in (23). And here we started this section with its counterpart (40). The rest is the same, since Stirling numbers are mutually orthogonal.  $\blacksquare$ 

The usefulness of this "anti-factorial" transform is highly questionable. But we will demonstrate in the next section that it works.

#### § 7. Borel integral summation

This paper would be incomplete without considering Borel summation, which is known as (almost) the most powerful and useful summation technique on the market (see [15, p. 182]).

Borel summation is basically a combination of Borel and Laplace transforms. Borel transform produces an EGF from the original GF to be summed. This

EGF is expected to be an analytical function such that a simple form of Laplace transform is applicable. This can be summarized as follows

$$\sum_{n=0}^{\infty} a_n x^n = \int_0^{\infty} \exp(-t) \left( \sum_{n=0}^{\infty} \frac{a_n}{n!} (x t)^n \right) dt,$$
(47)

where rhs(47) defines lhs(47) as a summed up formal PS in an integral form (since both sides of (47) have the same asymptotic expansion).

Thus Borel summation is especially useful when we already have an EGF, as is the case with most combinatorial sequences.

In this section, we examine what Borel summation can do for combinatorial sequences considered in this paper.

It turns out that Borel summation indeed can sum up almost all of them in an integral form. Thus Borel summation is at least as strong as factorial summation (i.e.,  $PS \rightarrow FS$ ). However, Borel summation produces quadratures, which are certainly analytical functions, but not as good as functions expressed in a closed form in known functions.

This observation becomes transparent when we try to plot a function or find some of its properties.

Thus factorial summation, when applicable on a par with Borel summation, gives much more than the latter and basically evaluates the quadratures that Borel summation produces in closed forms.

First, we consider the function G(x) for Bernoulli numbers (see (27)). Using EGF (2) with  $x \to -x$ , we obtain

$$G(x) = x \int_{0}^{\infty} \frac{t \exp(-t)}{1 - \exp(-xt)} dt = \frac{1}{x} \Psi(1, \frac{1}{x}).$$

This is, actually, a unique quadrature that was evaluated by CAS independantly without the help from factorial summation.

The functional Eq (1) has the solution

$$\beta_1(x) = x^2 \int_0^\infty \frac{t \exp(-t)}{\exp(x t) - 1} dt = \Psi(1, 1 + \frac{1}{x}).$$

Combining the formulas (23) and (33), we produce an integral representation of a key function in factorial summation, i.e.,

$$Q(x,n) = \frac{(-1)^n}{n!} \int_0^\infty \exp(-t) \left(1 - \exp(-tx)\right)^n dt.$$

This formula immediately gives Borel sum for GF W(x) for Bell numbers (see Sect. 4)

$$W(x) = \int_{0}^{\infty} \exp(-t) \, \exp(\exp(-t\,x) - 1) dt,$$

(see also (35)).

It is now clear why FS U(x) was not summable. Borel summation gives

$$U(x) = \int_{0}^{\infty} \exp(-t) \, \exp(\exp(t \, x) - 1) dt,$$

which does not exist for  $\operatorname{Re}(x) > 0$ .

Euler and Tangent GFs give especially interesting examples of evaluation of integrals in a closed form. We have

$$E_g(x) = \int_0^\infty \frac{\exp(-t)}{\cosh(t\,x)} dt, \quad T_g(x) = \int_0^\infty \exp(-t)\left(1 + \tanh(t\,x)\right) dt$$

Since these integrals can be evaluated at special values of x explicitly, we have some unusual hypergeometric evaluations. For example,  $E_g(1) = \log 2$ , and

$$\log 2 = 1 - \frac{1}{6} F\left(\left[\frac{3}{4}, 1\right], \left[\frac{7}{4}\right], \frac{-1}{4}\right) - \frac{1}{8} F\left(\left[1, 1\right], \left[2\right], \frac{-1}{4}\right) - \frac{1}{20} F\left(\left[1, \frac{5}{4}\right], \left[\frac{9}{4}\right], \frac{-1}{4}\right).$$
Also,  $T_g(1) = \pi/2$ , and
$$\frac{\pi}{2} = 1 + \frac{1}{2} F\left(\left[\frac{1}{2}, 1\right], \left[\frac{3}{2}\right], \frac{-1}{4}\right) + \frac{1}{6} F\left(\left[\frac{3}{4}, 1\right], \left[\frac{7}{4}\right], \frac{-1}{4}\right) - \frac{1}{20} F\left(\left[1, \frac{5}{4}\right], \left[\frac{9}{4}\right], \frac{-1}{4}\right).$$
The functional identities (39) can also be considered as integral transformations for these GFs. They can be applied recursively. We give just one

$$\int_{0}^{\infty} \exp(-t) \left(1 + \tanh(t\,x)\right) dt = 2 - \frac{1}{1+2\,x} \int_{0}^{\infty} \exp(-t) \left(1 + \tanh\left(\frac{t\,x}{1+2\,x}\right)\right) dt,$$

which is far from being obvious.

Finally, we use Borel summation for GF for Stirling numbers  $S_1(n, m)$  and obtain another representation for the functions  $Y_m(x)$ , i.e.,

$$Y_m(x) = \frac{1}{m!} \int_0^\infty \exp(-t) \log^m (1+t\,x) dt, \quad m \in \mathbb{N}_0.$$
(48)

It is easy to verify that integration by parts gives  $Y_1(x)$  and  $Y_2(x)$  explicitly as before, but for m > 2, this does not produce anything meaningful.

Explicit forms (48) for  $Y_m(x)$  allow to demonstrate that the transform for asymptotic PS in Theorem 3 works. For example, consider the sequence  $\{a(n) = (-1)^n n!, n \in \mathbb{N}_0\}$ , This is a shifted sequence for the Stirling numbers  $S_1(n, 1)$ ; thus the GF for  $\{a(n)\}$  is  $Y_1(x)/x$ . Since the inverse Stirling transform in Theorem 3 for  $\{a(n)\}$  produces  $\{(-1)^n\}$ , we can sum up the integrals there and obtain

$$\frac{1}{x} \exp\left(\frac{1}{x}\right) \operatorname{Ei}\left(1, \frac{1}{x}\right) = \int_{0}^{\infty} \frac{\exp(-t)}{1+t \, x} \, dt.$$

Of course, this can also be obtained directly by Borel summation.

In conclusion, we remark that divergent PS have numerous applications from combinatorics (see [1]) to quantum physics (see [16]). In a recent review [17], we read that "the relevance of asymptotic series in physical problems is hard to overestimate", and that "perturbation series for quantum mechanical systems are almost always divergent, and define instead asymptotic series for the perturbed energy eigenvalues".

As a reflection of such an importance of divergent PS, an article in Wikipedia on divergent series counts more than 15 summation techniques. The factorial summation is noticeably absent.

Hopefully, we have demonstrated that factorial transform can be at least as useful for summation of divergent PS as the Borel summation.

# References

- S. Lando, Lectures on Generating Functions (Student Math. Lib., Amer. Math. Soc., 27, 2003).
- [2] T. Arakawa, T. Ibukiyama, M. Kaneko, Bernoulli Numbers and Zeta Functions (Springer, Japan, 2014).
- [3] J. Frame, "The Hankel power sum matrix inverse and the Bernoulli continued fraction", Math. Comp., 33, (146), 815-826, (1979).

- [4] L. Euler, "De seriebus divergentibus", Novi Comment. Acad. Sci. Petropolitanae, 5, 205-237, (1754/55). reprint: "Opera omnia", Ser. I, 14, Teubner, Leipzig, 585-617, (1925).
- [5] N. Nielsen, *Die Gammafunktion* (Teubner, Leipzig, Berlin, 1906) = (Chelsea, New York, 1965).
- [6] E. J. Weniger, "Summation of divergent power series by means of factorial series", [arXiv:1005.0466v1] (2010). (http://arxiv.org/abs/1005.0466v1).
- [7] K. Knopp, *Theory and applications of infinite series* (Blackie & Son, London, 1946).
- [8] "Sloane online encyclopedia of integer sequences", (http://oeis.org).
- [9] S. Khrushchev, Orthogonal Polynomials and Continued Fractions (Encycl. of Math. and its Aappl. 122, Cambridge Univ. Press, 2008).
- [10] S. R. Finch, *Mathematical Constants* (Encycl. of Math. and its Aappl. 94, Cambridge Univ. Press, 2003).
- [11] B. Candelpergher, Ramanujan Summation of Divergent Series (Lecture Notes in Math., Springer, 2017).
- [12] G. N. Watson, "The transformation of an asymptotic series into a convergent series of inverse factorials", Rend. Circ. Mat. Palermo, 34, 41-88, (1912).
- [13] M. Petkovšek, H.S. Wilf, D. Zeilberger, A=B (Taylor & Francis, 1996).
- [14] M. Bernstein, N.J.A. Sloane, "Some Canonical Sequences of Integers", Lin. Algebra and its Appl., 226-228, 57-72 (1995).
- [15] G.H. Hardy, *Divergent Series* (New York, Chelsea, 1949, 1992).
- [16] J. Glimm, A. Jaffe, Quantum physics (2nd ed.) (Berlin, New York: Springer, 1987).
- [17] C. Bender, C. Heissenberg, "Convergent and Divergent Series in Physics", [arXiv:1703.05164v2] (2016). (https://arxiv.org/abs/1703.05164v2).