On an identity by Chaundy and Bullard. II. More history

Tom H. Koornwinder and Michael J. Schlosser*

Abstract

An identity by Chaundy and Bullard writes $1/(1-x)^n$ $(n=1,2,\ldots)$ as a sum of two truncated binomial series. In a paper which appeared in 2008 in Indag. Math. the authors surveyed many aspects of this identity. In the present paper we discuss much earlier occurrences of this identity in works by Hering (1868), de Moivre (1738) and de Montmort (1713). A relationship with Krawtchouk polynomials in work by Greville (1966) is also discussed.

1 Introduction

In our paper [13] we surveyed the history of the often rediscovered formula

$$1=(1-x)^{n+1}\sum_{k=0}^{m}\binom{n+k}{k}x^k+x^{m+1}\sum_{k=0}^{n}\binom{m+k}{k}(1-x)^k. \tag{1}$$

We attributed the formula to Chaundy & Bullard [2, p.256] (1960). However, we later learnt that some giant steps back in time can be made to much earlier occurrences of this formula. Almost one century before Chaundy & Bullard the formula was given by Hering [10] (1868). Then, with a jump of more than one century, the formula was found in the work of de Moivre [14] (1738). Even 25 years earlier the formula was given in implicit form already by de Montmort [16] (1713).

The paper successively discusses these three early occurences of the formulas. Next a correspondence between Samuel Pepys and Isaac Newton, having some relation with identity (1), is briefly discussed. We conclude with a much more recent connection with Krawtchouk polynomials which is implicit in Greville [7] (1966).

Acknowledgements We thank Pieter de Jong for communicating us the occurrence of formula (1) in de Moivre [14], [15], for calling our attention to reference [3], and for sending us his manuscript [12]. We also thank the Mathematics Department of Lund University for sending us a copy of Hering's paper [10], which happened to be in the Small Boxes of the Gösta Mittag-Leffler Separate Collection, stored at this Department (see [3]).

*Partly supported by FWF Austrian Science Fund grant S9607.

2 Hering (1868)

In 1868 Hering [10, p.14, formula 1)] derived:

$$(1-x)_n^{-m} = (1-x)^{-m} - (1-x)^{-m}\, x^n\, (1-\overline{1-x})_m^{-n}. \tag{2}$$

Here $(1-x)_n^{-m}$ is the power series in x of $(1-x)^{-m}$ cut after the n-th term. Similarly, $(1-\overline{1-x})_m^{-n}$ is the power series in $1-x$ of $(1-(1-x))^{-n}$ cut after the m-th term. Thus Hering already had (1).

Hering's proof is different from any of the proofs given in (1). For generic non-integer m he writes for the left-hand side of (2):

$$(1-x)_n^{-m} = \sum_{k=0}^{n-1} \frac{(m)_k}{k!}\, x^k = \frac{(m)_{n-1}}{(n-1)!}\, x^{n-1}\, {}_2F_1\left(\begin{matrix} -n+1, 1 \\ -m-n+2 \end{matrix}; x^{-1}\right) = \frac{(m)_{n-1}}{(n-1)!}\, \frac{x^n}{x-1}\, {}_2F_1\left(\begin{matrix} -m+1, 1 \\ -m-n+2 \end{matrix}; \frac{1}{1-x}\right), \tag{3}$$

where inversion of the order of summation is used in the second equality and Pfaff's transformation formula in the third equality. If m tends to a positive integer, the last ${}_2F_1$ becomes

$$\sum_{k=0}^{m-1} \frac{(-m+1)_k}{(-m-n+2)_k}\, (1-x)^{-k} + \sum_{k=m+n-1}^{\infty} \frac{(-m-n+k+2)_{n-1}}{(-m-n+2)_{n-1}}\, (1-x)^{-k}. \tag{4}$$

(Here, although not emphasized by Hering, we should require for convergence that $|x-1|>1$. This can later be relaxed in (2) by analytic continuation.) After multiplication of (4) by $\binom{m+n-2}{n-1} x^n/(x-1)$ we can rewrite the first term (by inversion of the order of summation) as

$$-(1-x)^{-m}\, x^n \sum_{k=0}^{m-1} \frac{(n)_k}{k!}\, (1-x)^k = -(1-x)^{-m}\, x^n\, (1-\overline{1-x})_m^{-n},$$

and the second term as

$$(-1)^n (1-x)^{-m-n} x^n \sum_{k=0}^{\infty} \frac{(n)_k}{k!}\, (1-x)^{-k} = (1-x)^{-m}.$$

Thus by substitution in (3) Hering settled (2).

Formula (2) is just one of many formulas derived in [10]. Hering does not specially emphasize this particular result.

3 de Moivre (1738)

A much earlier reference was kindly communicated to us by Pieter de Jong and also mentioned in his manuscript [12]. In 1738 A. de Moivre [14, p.196] (see also the 1754 edition [15, p.224]) wrote:

> *But as there is a particular elegancy for the Sums of a finite number of Terms in those Series whose Coefficients are figurate numbers beginning at Unity, I shall Set down the Canon for those Sums.*

Let n denote the number of Terms whose Sum is to be found, and p the rank or order which those figurate numbers obtain, then the Sum will be

$$\frac{1-x^n}{(1-x)^p} - \frac{nx^n}{(1-x)^{p-1}} - \frac{n.n+1.x^n}{1.2.(1-x)^{p-2}} - \frac{n.n+1.n+2.x^n}{1.2.3.(1-x)^{p-3}} - \frac{n.n+1.n+2.n+3.}{1.2.3.4.(1-x)^{p-4}}, \text{ etc.}$$

which is to be continued till the number of Terms be = p.
[In the numerator of the last term above a factor x^n is missing. This may have been a printer's error.]

According to de Moivre [14, Corollary at end of p.195] the figurate numbers of order p are the successive coefficients in the power series of $\frac{1}{(1-x)^p}$. Thus, since he begins at unity, these are binomial coefficients $\binom{p+k-1}{k}$ $(k=0,1,2,\ldots)$. This is slightly different from the modern definition given by Dickson [5, p.7], who defines the k-th figurate number of order p as the binomial coefficient $\binom{p+k-1}{p}$. The difference is that de Moivre starts counting orders at 1 (for instance triangular numbers have order 3 for him), while Dickson starts counting them at 0, by which triangular numbers have order 2.

Thus by the above quotation de Moivre gives the identity

$$\sum_{k=0}^{n-1}\binom{p+k-1}{k}x^k = (1-x)^{-p} - (1-x)^{-p}x^n\sum_{k=0}^{p-1}\binom{n+k-1}{k}(1-x)^k.$$

Indeed, this shows that de Moivre had already (1) in 1738.

The section "Of the Summation of recurring Series" starting in de Moivre [14, p.193] gives some indication how he obtained his result. We will summarize this in modern terminology and we state everything at once for general p instead of stating it for $p=1,2,3$, etc.

First de Moivre discusses infinite power series $S=\sum_{k=0}^{\infty}c_kx^k$ in which the coefficients satisfy a recurrence relation $c_k = a_1c_{k-1} + a_2c_{k-2} + \cdots + a_pc_{k-p}$ with coefficients a_j independent of k. Then he observes that $S = q(x)/(1-a_1x-\cdots-a_px^p)$, where $q(x)$ is a polynomial of degree at most $p-1$ which can be explicitly computed. Next he observes that the figurate numbers

$$c_k = \binom{p+k-1}{k} = \frac{(k+1)_{p-1}}{(p-1)!},$$

being polynomials of degree $p-1$ in k which vanish for $k=-1,-2,\ldots,-p+1$, are annihilated by the p-th finite difference:

$$\sum_{l=0}^{p}(-1)^l\binom{p}{l}c_{k-l} = 0 \qquad (k=1,2,\ldots).$$

From this he derives that $S=(1-x)^{-p}$ in this case. Finally he applies the same method to terminating power series $S_n=\sum_{k=0}^{n-1}c_kx^k$. A recurrence relation $c_k = a_1c_{k-1} + a_2c_{k-2} + \cdots + a_pc_{k-p}$ will then yield $S_n = (q(x)+x^nr(x))/(1-a_1x-\cdots-a_px^p)$ for certain polynomials $q(x)$ and $r(x)$ of degree at most $p-1$. For c_k being the figurate numbers we get $S_n=(1+x^nr(x))/(1-x)^p$, from which $r(x)$ can be computed, in principle. However, de Moivre does not give an argument how he arrives at the nice

explicit expression for $r(x)$ expanded in powers of $1-x$. Probably, he found the expression for low values of p and then extrapolated.

De Moivre also made an important step by which he might have concluded the multi-variable generalization of (1) given in [13, (1.5)] and first obtained (as far as we know) by Damjanovic, Klamkin and Ruehr [19] in 1986. In fact, de Moivre [14, Problem LXIX, p.191], [15, pp.50,51] gave for any of the n summands in the outer sum in [13, (1.5)] a probabilistic interpretation coming from the problem of points with n players (see the case of 2 players below). Adding up these chances to 1 would have given him the multi-variable formula, just as was pointed out in [19] (in detail in a situation with three urns by Bosch and Steutel). We are puzzled why de Moivre missed this final step, and also why he did not give a probabilistic interpretation of (1).

4 de Montmort (1713)

In 1713 appeared the second edition of the *Essay d'analyse sur les jeux de hazard* [16] by Pierre Raymond de Montmort. It contained among others a new solution of the so-called *problem of points* for two players. This problem comes from a game of chance with two players Pierre and Paul who have chances p and $1-p$, respectively, of winning each round. The player who has first won a certain number of rounds (this number may be different for Pierre and Paul) will collect the entire prize. Suppose that the game is prematurely interrupted when Pierre has to win still n rounds and Paul m rounds. What is then a fair division of the stake? See Hald [9, §14.1] for a description how this problem was handled by de Montmort.

In the case of equal chances the problem was already solved by Pascal and Fermat in 1654. In the case of unequal chances Johann Bernoulli generalized their solution. Bernoulli gives his solution in a letter to de Montmort dated 17 March 1710. This letter is included in the second edition of de Montmort's book, see [16, pp.283–298], in particular p.295 (English translation available at [17]). De Montmort also gives this solution in his main text, see [16, pp. 244–245, §190] (English translation at [17]). Curiously, Bernoulli is not mentioned there by de Montmort. Neither he acknowledges this new result of Bernoulli in his polemical discussion of earlier work on the problem of points in the Avertissement of the second edition of his book. This discussion starts on p. xxxiv of [16] (English translation available at [17]). Bellhouse [1] gives an interesting discussion of the relationship between de Montmort and de Moivre.

Bernoulli's solution is as follows. Imagine Pierre and Paul still play $m+n-1$ rounds. Then there will certainly be a winner. Pierre will be the winner if he has won n or more of these rounds. His chance for this is

$$\sum_{k=0}^{m-1} \binom{m+n-1}{k} p^{m+n-k-1}(1-p)^k. \tag{5}$$

But if Pierre has won $n-1$ or less of these rounds then Paul will be the winner. The chance for this is

$$\sum_{\ell=0}^{n-1} \binom{m+n-1}{\ell} p^{\ell}(1-p)^{m+n-\ell-1}. \tag{6}$$

The two chances add up to 1, both by the probabilistic interpretation and by the binomial formula.

De Montmort [16, p.245, §191] continues to give other expressions for the chances for Pierre respectively Paul to win. Imagine they still play until there is a winner. Pierre will be the winner if he has won already $n-1$ rounds and Paul at most $m-1$ rounds, and if then the next round is won by Pierre. Thus the chance for Pierre to win is

$$p\sum_{k=0}^{m-1}\binom{n+k-1}{k}p^{n-1}(1-p)^k, \tag{7}$$

Similarly, for Paul the chance to win is

$$(1-p)\sum_{k=0}^{n-1}\binom{m+k-1}{k}p^k(1-p)^{m-1}. \tag{8}$$

These two chances necessarily add up to 1. Thus the resulting formula (not given by de Montmort) is

$$p^n\sum_{k=0}^{m-1}\binom{n+k-1}{k}(1-p)^k+(1-p)^m\sum_{k=0}^{n-1}\binom{m+k-1}{k}p^k=1.$$

by which formula (1) is proved in a probabilistic way. This is essentially the same proof as was quoted from much more recent literature in [13, end of Section 6].

Clearly the chances (5) and (7) are the same. This is not explicitly observed by de Montmort, but it is indicated in the example where $n=5$ and $m=3$. In the general case the resulting identity is [13, (2.7)]. There we referred to Guenther [8], who gave various proofs and references (but none older than 1933) for this identity, including the probabilistic proof we just observed.

5 Pepys and Newton (1693)

In 1693 Samuel Pepys wrote a letter to Isaac Newton with a question about a probabilistic problem coming from a question to Pepys by John Smith (see [2], [18]). The question was (in modern terms):

> Let $6k$ fair dice be tossed independently and suppose that at least k "6"'s appear. For which $k=1,2,3$ this has the greatest chance to happen?

Newton wrote back three times. He answered correctly that the case $k=1$ has the highest probability. He actually computed the probability for $k=1$ and 2. He also gave a theoretical argument about which Stigler [3], as late as 2006, observed that it was incorrect. Chaundy & Bullard [2] showed more generally:

> We work with fair dice with s faces. Let $g(sn,n)$ be the chance that a selected face turns up less than n times in sn throws. Then $g(sn,n)$ increases with n for fixed s.

They proved this statement by expressing $g(sn,n)$ in terms of (5) and then using that (5) is equal to (7). (In passing, in connection with their proof of this identity, they observed the identity (1).) Finally, by working with (7) they could prove their claim quoted above.

6 Greville (1966)

This last item does not push the history of identity (1) further back, but mentions an unexpected aspect of this identity which is offered by Greville [7, p.166]. A further description is given in [6, §4,6]. Greville considers the smoothing filter $f \to g$ given by

$$g(y) = \sum_{x=-N}^{N} f(y-x)\,\mathbf{K}_{2n}(x,0)\,w(x) \qquad (y \in \mathbb{Z}),$$

where $w(x) := \binom{2N}{N+x}$ and $\mathbf{K}_n$ is the Christoffel-Darboux kernel for the orthogonal polynomials p_n satisfying

$$\sum_{x=-N}^{N} p_n(x)\,p_m(x)\,w(x) = h_n \delta_{n,m} \qquad (n,m \in \{0,1,\dots,2N\}).$$

Then the polynomials p_n are special shifted Krawtchouk polynomials

$$p_n(x) = K_n(x+N;\tfrac{1}{2};2N),$$

but this is not explicitly mentioned by Greville. Then we also see that $h_n = 2^{2N}\binom{2N}{n}^{-1}$ and that

$$\mathbf{K}_{2n}(x,0) = \sum_{k=0}^{2n} \frac{p_k(x)p_k(0)}{h_k}.$$

Greville wants to compute the characteristic function (or transfer function) ϕ associated with this smoothing filter, given by

$$\phi(\omega) := \sum_{x=-N}^{N} \mathbf{K}_{2n}(x,0)e^{-i\omega x}.$$

Then he derives that

$$\phi(\omega) = 1 - (\sin^2(\omega/2))^{n+1} P(\sin^2(\omega/2)) = (\cos^2(\omega/2))^{N-n} Q(\sin^2(\omega/2)) \tag{9}$$

for certain polynomials P of degree $N-n-1$ and Q of degree n. Then, with the same argument as in [4, Section 6.1] and [13, Remark 2.2], Greville explicitly obtains P and Q. As a consequence, (9) takes the form of (1) with $m = N-n-1$. Greville also concludes from the explicit expression that ϕ is monotonically decreasing from 1 to 0 on $[0,\pi]$. Later Herrmann [11] independently computed (9) in a different way in order to arrive at this result of monotonical decrease of ϕ, which he called maximal flatness.

References

[1] D. Bellhouse. *Banishing Fortuna: Montmort and De Moivre*, J. History Ideas 69 (2008), 559–581.

[2] T. W. Chaundy and J. E. Bullard, *John Smith's problem*, Math. Gazette 44 (1960), 253–260.

[3] T. Claesson and J. Peetre, *Index of the Gösta Mittag-Leffler separate collection. Part I, Small Boxes*, Lund University, 1996; http://staff.science.uva.nl/~thk/art/2012/ChaundyBullard2/

[4] I. Daubechies, *Ten Lectures on Wavelets*, Regional Conference Series in Applied Math., vol. 61, SIAM, 1992.

[5] L. E. Dickson, *History of the theory of numbers. Vol. II: Diophantine analysis*, Carnegie Inst. of Washington, Washington, D.C., 1920; reprinted, Chelsea, New York, 1966.

[6] E. Diekema and T. H. Koornwinder, *Differentiation by integration using orthogonal polynomials, a survey*, J. Approx. Theory 164 (2012), 637–667; arXiv:1102.5219v2 [math.CA].

[7] T. N. E. Greville, *On stability of linear smoothing formulas*, SIAM J. Numer. Anal. 3 (1966), 157–170.

[8] W. C. Guenther, *Solution of Problem E 1829*, Amer. Math. Monthly 74 (1967), 1134–1135.

[9] A. Hald, *A history of probability and statistics and their applications before 1750*, Wiley, New York, 1990.

[10] A. G. Hering, *Summation der n ersten Glieder der binomischen Reihe mittelst der Theorie der hypergeometrischen Reihen*, Programm der Realschule in Chemnitz, 1868; http://staff.science.uva.nl/~thk/art/2012/ChaundyBullard2/; JFM 01.0089.04.

[11] O. Herrmann, *On the approximation problem in nonrecursive digital filter design*, IEEE Trans. Circuit Theory 18 (1971), 411–413.

[12] P. de Jong, *The arrangement of the arithmetical triangle*, manuscript, 2012.

[13] T. H. Koornwinder and M. J. Schlosser, *On an identity by Chaundy and Bullard. I*, Indag. Math. (N.S.) 19 (2008), 239–261; arXiv:0712.2125v3 [math.CA].

[14] A. de Moivre, *The doctrine of chances*, Second edition, Woodfall, London, 1738; reprinted, Frank Cass & Co., Ltd., London, 1967.

[15] A. de Moivre, *The doctrine of chances*, Third edition, A. Millar, London, 1756; reprinted, Chelsea, New York, 1967.

[16] P. R. de Montmort, *Essay d'analyse sur les jeux de hazard*, Seconde édition, Laurent le Conte, Paris, 1713; reprinted, Chelsea, New York, 1980.

[17] R. J. Pulskamp, *Pierre Raymond de Montmort*, http://www.cs.xu.edu/math/Sources/Montmort/montmort.html.

[18] S. M. Stigler, *Isaac Newton as a probabilist*, Statist. Sci. 21 (2006), 400–403.

[19] *Solution of Problem 85-10*, SIAM Rev. 28 (1986) 243–244.

T. H. Koornwinder, Korteweg-de Vries Institute, University of Amsterdam,
P.O. Box 94248, 1090 GE Amsterdam, The Netherlands;
email: T.H.Koornwinder@uva.nl

M. J. Schlosser, Fakultät für Mathematik, Universität Wien,
Nordbergstrasse 15, A-1090 Vienna, Austria;
email: michael.schlosser@univie.ac.at