The Catalan Number k-Fold Self-Convolution Identity: The Original Formulation

Peter J. Larcombe and David R. French (Retired)

School of Computing and Technology
University of Derby, Kedleston Road, Derby DE22 1GB, U.K.
{P.J.Larcombe@derby.ac.uk}

Abstract

A known convolution identity involving the Catalan numbers is presented and discussed. Catalan's original formulation, which is algebraically straightforward, is similar in style to one reported previously by the first author and the result has some interesting combinatorial aspects.

Introduction

In 1887, the identity presented here was published by Eugène Catalan in an Italian journal article [1] as part of the Introduction (see pp.194-195). The previous year it had appeared as the "Addition" (pp.62-64) to a paper [2] in a different journal (run under the auspices of the Société Royale des Sciences de Liège), with an attached date of April 1876—some ten years earlier.

In this offering we outline the derivation of the result according to Catalan, which involves the k-fold self-convolution of the Catalan sequence

$$\{c_0, c_1, c_2, c_3, c_4, \ldots\} = \{1, 1, 2, 5, 14, \ldots\} \tag{1}$$

whose $(n+1)$th term c_n is defined by

$$c_n = \frac{1}{n+1}\binom{2n}{n}, \qquad n = 0, 1, 2, \ldots \tag{2}$$

JCMCC **46** (2003), pp. 191-204

Verifying examples are included for the benefit of the reader. Our exposition is given in the same spirit that led the author P.J.L. to write on an obscure and seemingly forgotten convolution type identity of Catalan [3] (see also [4] for hypergeometric proofs) not too long ago. As in [3], the methodology adopted by Catalan is accessible at undergraduate level, although this does not diminish the formulation which demonstrates originality of thought at the time. Remarks concerning both combinatorial aspects of the identity and its appearances in the literature are also made.

Catalan's Formulation

We begin with an application of the powerful 1770 inversion formula of Lagrange; we state a result in which it can be found, taken for convenience from the treatise of Whittaker and Watson [5, p.133] (Catalan apparently used a version available in Bertrand's *Calcul Diffèrentiel*).

Theorem (Lagrange) Let $f(z)$ and $\phi(z)$ be functions of z analytic on and inside a contour C surrounding a point a, and let t be such that the inequality

$$|t\phi(z)| < |z-a|$$

is satisfied at all points z on the perimeter of C; then the equation

$$\zeta = a + t\phi(\zeta),$$

regarded as an equation in ζ, has one root in the interior of C; and further any function of ζ analytic on and inside C can be expanded as a power series in t by the formula

$$f(\zeta) = f(a) + \sum_{n=1}^{\infty} \frac{d^{n-1}}{da^{n-1}}[f'(a)\{\phi(a)\}^n]\frac{t^n}{n!}.$$

So as to be in line with Catalan's formulation we replace ζ, t with y, x, upon which, choosing as a special case $\phi(y) = 1/y$ and $f(y) = 1/y^k$ ($k > 0$ constant), the Theorem gives that the equation

$$y = a + \frac{x}{y} \tag{3}$$

in $y(x)$ yields a series form of $y^{-k}(x)$ as

$$\begin{aligned} y^{-k} &= a^{-k} - k\sum_{n=1}^{\infty} \frac{d^{n-1}}{da^{n-1}}[a^{-(k+n+1)}]\frac{x^n}{n!} \\ &= a^{-k} + k\sum_{n=1}^{\infty}(-1)^n a^{-(k+2n)} A(n;k)\frac{x^n}{n!} \end{aligned} \tag{4}$$

after a little work, $A(n;k)$ being the function

$$A(n;k) = (k+n+1)(k+n+2)\cdots(k+2n-1). \tag{5}$$

Although Catalan does not mention this, the $(n-1)$-term product $A(n;k)$ must be read as 1 for $n = 1$, but this becomes redundant if (4) is further written using factorials as

$$\begin{aligned} y^{-k} &= a^{-k} + k\sum_{n=1}^{\infty}(-1)^n a^{-(k+2n)}\frac{(k+2n-1)!}{n!(k+n)!}x^n \\ &= a^{-k} + \sum_{n=1}^{\infty}(-1)^n a^{-(k+2n)}B(n;k)x^n, \end{aligned} \tag{6}$$

where, to perhaps emphasise its integrality, he writes $B(n;k)$ as

$$B(n;k) = \binom{k+2n-1}{n} - \binom{k+2n-1}{n-1}. \tag{7}$$

Now, setting $a = 2$ and replacing x by $-4x$, we see from (3),(6) that associated with the particular equation

$$y = 2 - \frac{4x}{y} \tag{8}$$

is the kth inverse power of y

$$y^{-k} = 2^{-k}\left[1 + \sum_{n=1}^{\infty} B(n;k)x^n\right]. \tag{9}$$

On the other hand, however, the solution to (8) as a quadratic in y is $y(x) = 1 + \sqrt{1-4x}$ (taking necessarily the positive square root[1]), whence

$$\frac{1}{y} = \frac{1-\sqrt{1-4x}}{4x} \tag{10}$$

and in turn

$$y^{-k} = \frac{1}{2^k}G^k(x), \tag{11}$$

where

$$\begin{aligned} G(x) &= \frac{1-\sqrt{1-4x}}{2x} \\ &= c_0 + c_1x + c_2x^2 + \cdots \end{aligned} \tag{12}$$

[1]With $y(x) = 1+\sqrt{1-4x}$ then $y^{-1} \to 2^{-1}$ as $x \to 0$ (by (10)), so that, since integer $k > 0$, $y^{-k} = (y^{-1})^k \to (2^{-1})^k = 2^{-k}$ which is consistent with (9) in the limit; Catalan makes brief reference to this in the "Addition" to [2] as a footnote on the first page.

is the (ordinary) generating function for the Catalan numbers (known at the time to Catalan from the work of others). Reconciling (9) and (11), and multiplying both sides of the resulting equation by x^k, gives

$$\left[1+\sum_{n=1}^{\infty} B(n;k)x^n\right] x^k = (c_0 x + c_1 x^2 + c_2 x^3 + \cdots)^k, \tag{13}$$

from which the identity is immediate; writing a typical member of each of the k r.h.s. brackets as $c_{\alpha-1}x^{\alpha}, c_{\beta-1}x^{\beta}, c_{\gamma-1}x^{\gamma}, \ldots, c_{\lambda-1}x^{\lambda}$, then equating coefficients of x^{k+n} across (13) it follows that $(1 \le \alpha, \beta, \gamma, \ldots, \lambda \le n+1)$

$$\sum_{\alpha+\beta+\gamma+\cdots+\lambda=k+n} c_{\alpha-1}c_{\beta-1}c_{\gamma-1}\cdots c_{\lambda-1} = B(n;k), \tag{14}$$

with $B(n;k)$ (7) as defined. For specified $k, n \ge 1$ the sum contains $\binom{k+n-1}{n}$ terms, each a k-product of Catalan elements whose indices add to n. Note that what we know today as the Catalan numbers were then called "Segner numbers" in terms of which (14) was established in the aforesaid references [1,2] (thus explaining their titles) and in the "Addition" to [2].

Example 1: $n = 2,\ k = 4$

$$\begin{aligned}\sum_{\alpha+\beta+\gamma+\delta=6} & c_{\alpha-1}c_{\beta-1}c_{\gamma-1}c_{\delta-1} \\ &= c_2c_0c_0c_0 + c_0c_2c_0c_0 + c_0c_0c_2c_0 + c_0c_0c_0c_2 \\ &\qquad + c_1c_1c_0c_0 + c_1c_0c_1c_0 + c_1c_0c_0c_1 \\ &\qquad + c_0c_1c_1c_0 + c_0c_1c_0c_1 + c_0c_0c_1c_1 \\ &= 2+2+2+2+1+1+1+1+1+1 \\ &= 14 \\ &= \binom{7}{2} - \binom{7}{1} \\ &= B(2;4). \end{aligned} \tag{15}$$

Example 2: $n = 3,\ k = 4$

$$\begin{aligned}\sum_{\alpha+\beta+\gamma+\delta=7} & c_{\alpha-1}c_{\beta-1}c_{\gamma-1}c_{\delta-1} \\ &= c_3c_0c_0c_0 + c_0c_3c_0c_0 + c_0c_0c_3c_0 + c_0c_0c_0c_3 \\ &\qquad + c_2c_1c_0c_0 + c_2c_0c_1c_0 + c_2c_0c_0c_1 + c_0c_2c_1c_0 \\ &\qquad + c_0c_2c_0c_1 + c_1c_2c_0c_0 + c_1c_0c_2c_0 + c_1c_0c_0c_2 \end{aligned}$$

$$
\begin{aligned}
& + c_0c_1c_2c_0 + c_0c_1c_0c_2 + c_0c_0c_2c_1 + c_0c_0c_1c_2 \\
& \qquad + c_1c_1c_1c_0 + c_1c_1c_0c_1 + c_1c_0c_1c_1 + c_0c_1c_1c_1 \\
= \;& 5+5+5+5 \\
& +2+2+2+2+2+2+2+2+2+2+2+2 \\
& \qquad +1+1+1+1 \\
= \;& 48 \\
= \;& \binom{9}{3} - \binom{9}{2} \\
= \;& B(3;4). \qquad (16)
\end{aligned}
$$

Catalan himself illustrated his identity in the "Addition" to [2] for $n = 4$, $k = 3$.

<u>Remark</u> We make the remark that when $k = 1$ the identity merely confirms the definition of $c_n = \sum_{\alpha=n+1} c_{\alpha-1} = B(n;1) = \binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1}\binom{2n}{n}$, whilst the value $k = 2$ recovers the long established convolution

$$
\begin{aligned}
\sum_{\alpha+\beta=n+2} c_{\alpha-1}c_{\beta-1} &= c_0c_n + c_1c_{n-1} + \cdots + c_{n-1}c_1 + c_nc_0 \\
&= B(n;2) \\
&= \binom{2n+1}{n} - \binom{2n+1}{n-1} \\
&= \frac{1}{n+2}\binom{2(n+1)}{n+1} \\
&= c_{n+1}. \qquad (17)
\end{aligned}
$$

Another case perhaps worth validating is that for $n = 1$, where the identity correctly gives

$$
\begin{aligned}
& \sum_{\alpha+\beta+\gamma+\cdots+\lambda=k+1} c_{\alpha-1}c_{\beta-1}c_{\gamma-1}\cdots c_{\lambda-1} \\
= \;& c_1c_0c_0\cdots c_0 + c_0c_1c_0\cdots c_0 + c_0c_0c_1\cdots c_0 + \cdots + c_0c_0c_0\cdots c_1 \\
= \;& 1+1+1+\cdots+1 \\
= \;& k \\
= \;& \binom{k+1}{1} - \binom{k+1}{0} \\
= \;& B(1;k). \qquad (18)
\end{aligned}
$$

The analysis and examples above are consistent with Catalan's derivation. A possible oversight by him, however, is that $B(n;k)$ (7) takes the contracted form $B(n;k) = \frac{k}{2n+k}\binom{2n+k}{n}$ (familiar in combinatorial circles as a so called "ballot number" and available directly through a different cultivation of Lagrange inversion), *and we choose to state the result in final form thus, preserving Catalan's notation in the l.h.s. sum:* For $n \geq 0$, $k \geq 1$,

$$\sum_{\alpha+\beta+\gamma+\cdots+\lambda=k+n} c_{\alpha-1}c_{\beta-1}c_{\gamma-1}\cdots c_{\lambda-1} = \frac{k}{2n+k}\binom{2n+k}{n}. \tag{19}$$

Interpretations of $B(n;k)$

Before finishing with some concluding remarks we frame combinatorially the function $B(n;k)$ for completeness, selective values of which are shown below in a table whose columns have, in order (reading left to right), ordinary generating functions $G(x), G^2(x), G^3(x), G^4(x), \ldots$

k/n	1	2	3	4	5	6	7	8	
0	1	1	1	1	1	1	1	1	...
1	1	2	3	4	5	6	7	8	...
2	2	5	9	14	20	27	35	44	...
3	5	14	28	48	75	110	154	208	...
4	14	42	90	165	275	429	637	910	...
5	42	132	297	572	1001	1638	2548	3808	...
6	132	429	1001	2002	3640	6188	9996	15504	...
7	429	1430	3432	7072	13260	23256	38760	62016	...
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	

Table 1: Values of $B(n;k)$.

If columns $2, 3, 4, 5, \ldots$, are pushed down by the respective amount of places $1, 2, 3, 4, \ldots$ (and the gaps they leave replaced with zeros), then the modified array forms a particular Riordan matrix (introduced in the seminal 1991 paper by Shapiro *et al.*, *Disc. Appl. Math.*, **34**, pp.229-239) whose columns have generating functions $G(x), xG^2(x), x^2G^3(x), x^3G^4(x), \ldots$

Pleasing combinatorial views of $B(n;k)$ are to be found in considering walks in the 2D plane. Denoting, in the usual fashion, an upward step from general point (x, y) to $(x+1, y+1)$ as U, and likewise writing D for a step from (x, y) to $(x+1, y-1)$, then $B(n;k) = \frac{k}{2n+k}\binom{2n+k}{n}$ counts the number of paths from the origin $(0,0)$ to the point $(2n+k-1, k-1)$ staying on or above the x-axis via steps of type U and D; a formal proof of this result is not elementary (requiring the use of André's Reflection Principle), and is dealt with in the Appendix. Linked to this is the observation that the number of k concatenated Dyck paths[2] whose semi-length is n is again $B(n;k)$. It is perhaps helpful to provide an example for this last interpretation. Associated with the concatenation of, for instance, $k = 4$ Dyck paths is the generating function $G^4(x) = \sum_{n=0}^{\infty} B(n;4)x^n = 1 + 4x + 14x^2 + 48x^3 + \cdots$, whose (enumerating) coefficients are available from Table 1 and may be checked as follows (where e stands for the 'empty' Dyck path):

Paths of Semi-Length 0:

e	:	1 permutation of this type
Permutations Total	:	$1 = [x^0]\{G^4(x)\}$

Paths of Semi-Length 1:

UD, e, e, e	:	4 permutations of this type
Permutations Total	:	$4 = [x^1]\{G^4(x)\}$

Paths of Semi-Length 2:

UD, UD, e, e	:	6 permutations of this type
$UDUD, e, e, e$	:	4 permutations of this type
$UUDD, e, e, e$	:	4 permutations of this type
Permutations Total	:	$14 = [x^2]\{G^4(x)\}$

Paths of Semi-Length 3:

UD, UD, UD, e	:	4 permutations of this type
$UDUD, UD, e, e$	:	12 permutations of this type
$UUDD, UD, e, e$	:	12 permutations of this type

[2]A Dyck path comprised of steps U and D begins and ends at the same height (*i.e.*, y co-ordinate) and never, at any point, falls below that height. Its semi-length is one half of the difference in the x co-ordinates of the points at the beginning and end of the overall path.

$UUUDDD, e, e, e$	:	4 permutations of this type
$UDUDUD, e, e, e$	:	4 permutations of this type
$UDUUDD, e, e, e$	:	4 permutations of this type
$UUDDUD, e, e, e$	:	4 permutations of this type
$UUDUDD, e, e, e$	:	4 permutations of this type

Permutations Total : $48 = [x^3]\{G^4(x)\}$

This procedure can be continued as desired.

Further Remarks

We emphasise that it is the novel procedure of Catalan, in arriving at his result, which is of overriding concern here. A great deal of the theory now taken for granted by combinatorialists was undiscovered in the 19th century, and modern analysts have much more knowledge on which to draw. By way of illustration, a contrasting contemporary approach to (19) is made by Graham, Knuth and Patashnik [6] who define (see Section 7.5 therein) a power series $P(z)$ through the functional relation

$$P(z) = zP^m(z) + 1, \tag{20}$$

and show (using an observation of G.N. Raney regarding cyclic shifts) that $[z^n]\{P^l(z)\}$ is, for integer $l > 0$, the number of sequences of length $mn + l$ possessing the characteristics that (i) each term in the sequence is either 1 or $1 - m$, (ii) all partial sums are positive, and in particular (iii) the total sum is l. This number is found to be the (unique) combination of l sequences that each have the so called "m-Raney" property,[3] namely,

$$\sum_{n_1+n_2+\cdots+n_l=n} C_{n_1}^{(m)} C_{n_2}^{(m)} \cdots C_{n_l}^{(m)}, \tag{21}$$

where the "Fuss-Catalan" number of Graham *et al.* (sometimes referred to by people as the 'generalised' or 'higher' Catalan number) is[4]

$$C_n^{(m)} = \frac{1}{mn+1} \binom{mn+1}{n}. \tag{22}$$

[3]An m-Raney sequence $\{a_0, a_1, \ldots, a_{mn}\}$ of terms 1 or $1 - m$, defined after Raney, is one for which properties (i),(ii) hold and whose total sum is 1; the number of such sequences is $C_n^{(m)}$ (22).

[4]This entity arises naturally in considering polygon partitioning—a classic problem described in enumerative combinatorics with a long history. The interested reader is referred to the Appendix of an article [7] by Larcombe and Wilson for more details.

A simple counting argument (using a further result due to Raney) gives a closed form for the sum (21) as

$$\frac{l}{mn+l}\binom{mn+l}{n}. \tag{23}$$

For $m = 2$, $C_n^{(2)} = c_n$ (2) and $P(z) = G(z)$ (since (20) reads $P(z) = zP^2(z) + 1$, which the generating function $G(z)$ satisfies), and equating (21) with (23) reproduces (19) on changing l, z to k, x. Clearly, Catalan's identity relates in this situation to the composition of k 2-Raney sequences, each of individual length $2n+1$ and comprised of terms ± 1; it is an instructive interpretation, especially since (19)—although certainly known—is not overly discussed in the literature in this kind of way.

On this last point (and to conclude matters), we note that in [6], with the "generalised binomial series" $B_2(z)$ corresponding to $G(z)$ (see (5.68), p.203), the result (19) is listed in the equivalent form

$$G^k(x) = \sum_{n=0}^{\infty} \frac{k}{2n+k}\binom{2n+k}{n} x^n \tag{24}$$

as equation (5.70),[5] together with the related identity (5.72)

$$\frac{G^k(x)}{\sqrt{1-4x}} = \sum_{n=0}^{\infty} \binom{2n+k}{n} x^n. \tag{25}$$

These are also given in the useful paper by Deutsch and Shapiro [8] as (supplementary) Results I,II (p.33), and Wilf includes them as equations (2.5.16),(2.5.15) on p.54 of [9]. At the bottom of p.128 of [10] is a version of (24) (see also [9, p.170], where the inversion formula of Lagrange is applied to obtain it in yet a different guise), and doubtless there are other instances where it crops up in a particular context. Perhaps the most interesting technical formulations are the generalised ones of Gould [11] (the following results are also tabled as (1.121),(1.120) (p.15) in his well known work [12]) who, on defining

$$A_n(\alpha, \beta) = \frac{\alpha}{\alpha+\beta n}\binom{\alpha+\beta n}{n}, \tag{26}$$

derived the result (see p.85)

$$\sum_{n=0}^{\infty} A_n(\alpha, \beta) z^n = x^\alpha, \tag{27}$$

[5]The sum begins at $n = 0$ (where it gives the correct lead term in the expansion of $G^k(x)$ based on (12)), at which value, as it must, (19) holds with l.h.s. $c_0 c_0 c_0 \cdots c_0 = 1$.

where

$$z = \frac{x-1}{x^\beta} \tag{28}$$

and with convergence for $|z| < |(\beta-1)^{\beta-1}/\beta^\beta|$; setting $\alpha = k$ and $\beta = 2$ yields (24) from (26),(27) (emphasising that (28) implies $x(z) = G(z)$), and (25) is recovered in an analogous fashion (with a little more effort) by Gould's later identity (p.86)

$$\sum_{n=0}^{\infty} \binom{\alpha+\beta n}{n} z^n = \frac{x^{\alpha+1}}{(1-\beta)x+\beta}. \tag{29}$$

Other early appearances are in Riordan's 1968 text [13] (where obtaining (24) is set as part of Problem 2(a) (p.153), and showing (25) is likewise part of Problem 2(c) (p.154)[6]) and in the authoritative book by Pólya and Szegö [14] (see Gould's result (29) on p.302, the explicit identity (25) on p.303 and the references around them; also the form of (24) on p.301). Recently, Deutsch and Shapiro [15, p.246] mention that (24) can be derived via Lagrange inversion, and outline a neat inductive argument instead. In the same paper they make an appeal to (25) (which is written $[z^n]B(z)C^s(z) = \binom{2n+s}{n}$, with $B(z), C(z)$ appropriately defined) in the proof of Lemma 1 on p.256; a potted combinatorial derivation of the identity is offered in the accompanying footnote. Finally, we remark that hypergeometric function theory, applied by hand, produces without difficulty the corresponding series forms

$$G^k(x) = {}_2F_1\left(\begin{matrix}\frac{1}{2}k, \frac{1}{2}(k+1)\\ k+1\end{matrix}\,\middle|\, 4x\right) \tag{30}$$

and

$$\frac{G^k(x)}{\sqrt{1-4x}} = {}_2F_1\left(\begin{matrix}\frac{1}{2}(k+1), \frac{1}{2}(k+2)\\ k+1\end{matrix}\,\middle|\, 4x\right) \tag{31}$$

(written here in standard hypergeometric notation), which are generated directly (choosing $a = \frac{1}{2}k$, $z = 4x$) from the respective first and second equalities of the identity

$$\begin{aligned} {}_2F_1\left(\begin{matrix}a, \frac{1}{2}+a\\ 1+2a\end{matrix}\,\middle|\, z\right) &= 2^{2a}[1+\sqrt{1-z}]^{-2a} \\ &= \sqrt{1-z}\;{}_2F_1\left(\begin{matrix}1+a, \frac{1}{2}+a\\ 1+2a\end{matrix}\,\middle|\, z\right) \end{aligned} \tag{32}$$

[6] The lower entry of the r.h.s. binomial coefficient is incorrectly written k rather than (correctly) n, as is the summing index. Note that the start of Problem 2(c) re-states (24) with a minor variation of the function $B(n;k)$ (7) as the coefficient of x^n.

tucked away as (15.1.13) on p.556 of the celebrated handbook of Abramowitz and Stegun [16]. The reader may well have seen (24),(25) elsewhere. They will hopefully be included in the next edition of Volume 2 of R.P. Stanley's *Enumerative Combinatorics* (Cambridge University Press, Cambridge, U.K., 1999), and are currently to be found in the "Catalan Addendum"[7] to his oft cited problems involving the Catalan numbers in Chapter 6 of the book.

Summary

A convolution identity involving the Catalan numbers has been presented in a manner that is faithful to its original derivation by Catalan; it has also been framed combinatorially. Whilst the result is known, his personal formulation does not seem to have been discussed previously. It would appear that hypergeometric function theory cannot be applied to provide an alternative proof of it. We note that there is no evidence to suggest that Catalan had any interest in, or was at all seeking, a combinatorial interpretation of $B(n;k)$ or the identity (24), indeed we remind the reader that his work on which this article is based actually pre-dates by over a decade the publication by Désiré André of the elegant Reflection Principle in 1887 (invented by him to solve a now famous type of voting problem).

Appendix

Here we give a combinatorial interpretation, in the context of (integral) 2D lattice paths, for the expression $B(n;k) = \frac{k}{k+2n}\binom{k+2n}{n}$. Whilst the theory is standard, its inclusion is felt to be a useful one for the non-specialist reader.

Let P_1, P_2 have respective co-ordinates $(a_1, b_1), (a_2, b_2)$ in the x, y plane, where $0 \leq a_1 < a_2$, $b_1, b_2 \geq 0$. A straightforward path counting argument gives, via steps of type U, D, the number of *all* paths from P_1 to P_2 as

$$A(P_1, P_2) = \binom{a_2 - a_1}{\frac{1}{2}[a_2 - a_1 - (b_2 - b_1)]}, \tag{A1}$$

some of which may lie below the x axis at any point(s). Call a path *good* if it lies strictly above the x axis, otherwise it is *bad*. The well known Reflection Principle of André (see, for instance, Example 14.8 (pp.118-119)

[7]Located at the web site **http://www-math.mit.edu/~rstan/ec**; the two identities are set as problems, with succinct solutions provided (for the record, they were added in December 2001). The "Addendum" is continually being updated, and is worth a look for anyone genuinely interested in the Catalan numbers.

of van Lint and Wilson [17] or p.69 of the widely referenced discussion of paths in [18]) allows

$$B(P_1, P_2) = \binom{a_2 - a_1}{\frac{1}{2}[a_2 - a_1 - (b_2 + b_1)]} \tag{A2}$$

to be determined as the number of bad paths between P_1 and P_2, so that the good paths between them total

$$\begin{aligned} G(P_1, P_2) &= A(P_1, P_2) - B(P_1, P_2) \\ &= \binom{a_2 - a_1}{\frac{1}{2}[a_2 - a_1 - (b_2 - b_1)]} \\ &\qquad - \binom{a_2 - a_1}{\frac{1}{2}[a_2 - a_1 - (b_2 + b_1)]}; \end{aligned} \tag{A3}$$

as a small check, note that if either b_1 or b_2 is zero then there is no good path between P_1 and P_2 since one of them lies on the x axis, in which case (A3) correctly reads $G(P_1, P_2) = 0$. We now introduce the notion of an *almost good* path, which runs between two points and lies on or above (that is, not below) the horizontal. From P_1 to P_2 the set of such paths, numbering $AG(P_1, P_2)$, is by inspection equal to the set of good paths w.r.t. a new x axis translated downward by one unit from the original one at $y = 0$. In other words, we simply increase both b_1, b_2 by 1 in the r.h.s. of (A3) to give

$$\begin{aligned} AG(P_1, P_2) &= \binom{a_2 - a_1}{\frac{1}{2}[a_2 - a_1 - (b_2 - b_1)]} \\ &\qquad - \binom{a_2 - a_1}{\frac{1}{2}[a_2 - a_1 - (b_2 + b_1 + 2)]}. \end{aligned} \tag{A4}$$

Now, fix P_1 as the origin O and suppose a point P has co-ordinates (a, b), say $(a > 0,\ b \geq 0)$. Then (A4) contracts to

$$AG(O, P) = \binom{a}{\frac{1}{2}(a - b)} - \binom{a}{\frac{1}{2}(a - b - 2)}, \tag{A5}$$

which in turn simplifies to

$$AG(O, P) = \frac{2(b+1)}{a+b+2} \binom{a}{\frac{1}{2}(a - b)}. \tag{A6}$$

Setting $a = k + 2n - 1$, $b = k - 1$ as the particular co-ordinates of P, it is immediate that

$$B(n; k) = AG(O, P) = \binom{k + 2n - 1}{n} - \binom{k + 2n - 1}{n - 1} \tag{A7}$$

using (A5) (this is equation (7)), or, from (A6),

$$B(n;k) = \frac{k}{k+n}\binom{k+2n-1}{n} = \frac{k}{k+2n}\binom{k+2n}{n}. \qquad \text{(A8)}$$

Acknowledgements

Gratitude is expressed to an anonymous referee for some very helpful suggestions made to assist with the final preparation of the paper, and to Prof. E. Deutsch for attending so willingly to the many queries of the authors during this process; thanks also to Dr. D.G. Rogers who contributed one or two useful general comments about the work presented here.

References

[1] Catalan, E. (1887). Sur les nombres de Segner, *Rend. Circ. Mat. Pal.*, **1**, pp.190-201.

[2] Catalan, E. (1886). Sur les nombres de Segner, *Mél. Math.*, **2**, pp.54-61. [to this paper was attached the "Addition" (mentioned in the Introduction), and a further "Autre Addition" (pp.64-67) dated July 1885; it continued, under the same title, over pages 113-117 of *Mél. Math.*, **3** (1888)]

[3] Larcombe, P.J. (2000). A forgotten convolution type identity of Catalan, *Util. Math.*, **57**, pp.65-72.

[4] Larcombe, P.J. and Gessel, I.M. (2001). A forgotten convolution type identity of Catalan: two hypergeometric proofs, *Util. Math.*, **59**, pp.97-109.

[5] Whittaker, E.T. and Watson, G.N. (1950). A course of modern analysis, 4th Ed., Cambridge University Press, Cambridge, U.K.

[6] Graham, R.L., Knuth, D.E. and Patashnik, O. (1989). Concrete mathematics: a foundation for computer science, Addison-Wesley, Reading, U.S.A.

[7] Larcombe, P.J. and Wilson, P.D.C. (2001). On the generating function of the Catalan sequence: a historical perspective, *Cong. Num.*, **149**, pp.97-108.

[8] Deutsch, E. and Shapiro, L. (2001). Seventeen Catalan identities, *Bull. Inst. Comb. Appl.*, **31**, pp.31-38.

[9] Wilf, H.S. (1994). Generatingfunctionology, 2nd Ed., Academic Press, San Diego, U.S.A.

[10] Sedgewick, R. and Flajolet, P. (1996). An introduction to the analysis of algorithms, Addison-Wesley, Reading, U.S.A.

[11] Gould, H.W. (1956). Some generalizations of Vandermonde's convolution, *Amer. Math. Month.*, **63**, pp.84-91.

[12] Gould, H.W. (1972). Combinatorial identities, Rev. Ed., University of West Virginia, U.S.A.

[13] Riordan, J. (1968). Combinatorial identities, Wiley, New York, U.S.A.

[14] Pólya, G. and Szegö, G. (1954). Aufgaben und lehrsätze aus der analysis (Vol. 1), 2nd Ed., Springer-Verlag, Berlin, Germany.

[15] Deutsch, E. and Shapiro, L. (2001). A survey of the Fine numbers, *Disc. Math.*, **241**, pp.241-265.

[16] Abramowitz, M. and Stegun, I.A. (Eds.) (1970). Handbook of mathematical functions, 9th Printing, Dover Pubs., New York, U.S.A.

[17] van Lint, J.H. and Wilson, R.M. (1992). A course in combinatorics, Cambridge University Press, Cambridge, U.K.

[18] Hilton, P. and Pedersen, J. (1991). Catalan numbers, their generalization, and their uses, *The Math. Intell.*, **13**(2), pp.64-75.