International Journal of Pure and Applied Mathematics
Volume 43 No. 2 2008, 241-252

q-ANALOGUES OF A CONVOLUTION IDENTITY FOR CENTRAL BINOMIAL COEFFICIENTS

Ömer Eğecioğlu

Department of Computer Science
University of California Santa Barbara
Santa Barbara, CA 93106, USA
e-mail: omer@cs.ucsb.edu

Abstract: Let $A_n = \binom{2n}{n}$ for $n \geq 0$ denote the the n-th element in the axis of symmetry of the Pascal triangle. The generating function for A_n is $(1-4t)^{-1/2}$, from which it follows that $A_0A_n + A_1A_{n-1} + \cdots + A_nA_0 = 4^n$. There are also bijective proofs of this fact. Here we use a bijection between pairs of Catalan words and binary words of even length to construct q-analogues of this convolution identity.

AMS Subject Classification: 05A10, 05A15, 05A19
Key Words: convolution, Catalan word, Dyck word, central binomial coefficient, bijection, q-analogue

1. Introduction

We have the convolution identity

$$\sum_{k=0}^{n} A_k A_{n-k} = 4^n \,, \tag{1}$$

where $A_n = \binom{2n}{n}$ for $n \geq 0$ are the central binomial coefficients.

Since the generating function for A_n is $(1-4t)^{-1/2}$, this identity is immediate by algebraic means. There are also combinatorial, or "counting" proofs of (1); see for example [3], [1], [2]. Reference [2] contains natural bijective proofs using different combinatorial constructions relating various kinds of walks. Here we will make use of the bijective proof given in [1]. A short description

Received: December 20, 2007

is included to make this paper self-contained.

In one q-analogue of the convolution identity, the 4^n on the right-hand side of (1) becomes

$$\sum_{k=0}^{n} \binom{2n+1}{2k+1} q^k, \tag{2}$$

which q-counts binary words of length $2n$ by the number of peaks.

In another q-analogue, the 4^n on the right-hand side of (1) is replaced by

$$\sum_{k=0}^{2n} \binom{2n}{k} q^k = (1+q)^{2n}, \tag{3}$$

which q-counts binary words of length $2n$ by the number of 1's.

These q-analogues reflect properties of the bijection of [1], with the appropriate generating functions on the right, but the symmetry of the terms in the convolution sum on the left is lost. This is because the bijection q-counts A_k and A_{n-k} by related, but different statistics. Thus it is natural to look for the polynomials for more natural q-analogues of the central binomial coefficients to symmetrize the left-hand side. These are the coefficient polynomials obtained by expanding the square root of the generating function of the polynomials in for the right-hand sides in (2) and (3). For q-counting by peaks, we get an interesting set of candidate polynomials, whereas for the q-count by the number of occurrences of 1, then we only get an identity that is a direct algebraic consequence of the original convolution (1). We consider these issues of symmetrization after the construction of the two q-analogues.

The outline of this paper is as follows. In Section 2 we give the requisite definitions of the combinatorial objects that we use. This is followed by the description of the bijection for the proof of (1) in Section 3. q-analogues are constructed in Sections 4 and 5, followed by the discussion of the symmetric cases and conclusions in Section 6.

2. Preliminaries

Given a word w and a letter a, denote by $|w|$ the length of w, and by $|w|_a$ the number of occurrences of a in w. The empty word is denoted by ϵ. If $w = uv$, then u is a *prefix* and v is a *suffix* of w. A *Dyck word* w is a word over the binary alphabet $\Sigma = \{0, 1\}$ with the following properties:

1. For each prefix u of w, $|u|_1 \geq |u|_0$,

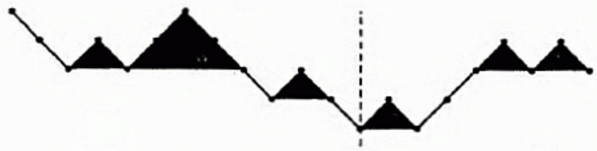

Figure 1: Catalan factorization: The lattice path and the factorization for $w = 00101100010010111010 \rightarrow w' = zz101100z10z10zz1010$, $r = 4$

2. $|w|_1 = |w|_0$.

Let $\mathcal{D}$ denote the language of Dyck words. If binary words are interpreted as lattice paths with diagonal steps ($1 = \nearrow, 0 = \searrow$), then $\mathcal{D}$ corresponds to those paths (Dyck paths) from the origin to some point $(2n, 0)$, that stay weakly above the x-axis. The number of Dyck paths of length $2n$ is the Catalan number $C_n = \binom{2n}{n}/(n+1)$ [4].

Each word w over Σ^{2n} with $|w|_0 = |w|_1 = n$ can be uniquely factored (using an extra letter z as marker) as $w' = d_0 z d_1 z d_2 \cdots d_{k-1} z d_k$ where k is even and each $d_i \in \mathcal{D}$. These are called *Catalan* words [5]. The first half of the z's in w' correspond to 0 in w, and the second half to 1. This is a bijection between words in Σ^{2n} with n occurrences of 1 and Catalan words of length $2n$. Consequently, there are $A_n = \binom{2n}{n}$ Catalan words of length $2n$ and identity (1) is equivalent to showing that each word $w \in \Sigma^{2n}$ corresponds bijectively to a pair of Catalan words (u, v) of total length $2n$.

An arbitrary word w over Σ^{2n} has a similar factorization as w', and can be unambiguously recovered from w' as soon as we specify the number r of z's from the left that correspond to 0 in w. The pair (w', r) is the *Catalan factorization* of w. This is best illustrated by shining a light source to the lattice path of w from the left and from the right. The steps corresponding to 0's that are lit from the left give the first r occurrences of z in w', those 1's that are lit from the right are the remaining z's in w'. This is illustrated in Figure 1 for $w = 00101100010010111010$. For a Catalan word w' the number $|w'|_z$ is even, and $r = \frac{1}{2}|w'|_z$, so this information is encoded in w' and need not be specified.

Given a Catalan word u of length $2k$, let $\phi_0(u)$ (resp. $\phi_1(u)$) denote the word in Σ^{2k} obtained by substituting 0 (resp. 1) for each z in u.

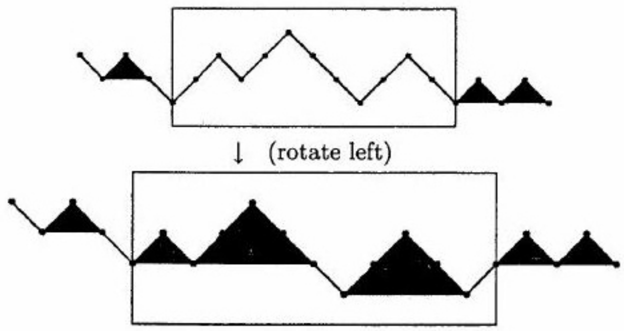

Figure 2: From (u, v) to w

3. The Bijection

The idea of the bijection is to break the Catalan factorization (w', r) of w into two pieces as u and v by splitting it immediately after the r-th z, as indicated by the vertical line in Figure 1. There is a difficulty when r is an odd number, in which case a type of transposition encodes this fact.

First start with a pair of Catalan words (u, v), where $|u| = 2k$ and $|v| = 2n - 2k$ for some k. From this pair, we construct $w \in \Sigma^{2n}$ as follows:

1. If $u = \epsilon$, then $w = \phi_1(v)$.
2. If $u = \alpha z$ then $w = \phi_0(u)\phi_1(v)$.
3. If $u = \alpha d$ for a Dyck word $d = 1\beta 0$ (where $\alpha = \epsilon$ or $\alpha = \gamma z$), then $w = \phi_0(\alpha\beta)\,(01)\,\phi_1(v)$.

Below are examples for $n = 10$ illustrating these three cases.

Example 1. Given pair $u = \epsilon$, $v = zz101100z10z10zz1010 \to w = 11101100110110111010$.

Example 2. Given pair $u = zz101100z10z$, $v = 10zz1010 \to w = (001011000100)(10111010)$ (this is the word whose path is displayed in Figure 1).

Example 3. Given pair $u = z10z110110001100$, $v = 1010$ (with $\alpha = z10z$, $d = 1\beta 0$, and $\beta = 1011000110$) $\to w = (01001011000110)(01)(1010)$. The indicated segment of the path of $\phi_0(u)\phi_1(v)$ on top portion of Figure 2 is transposed by a left circular rotation to obtain w as shown in the bottom.

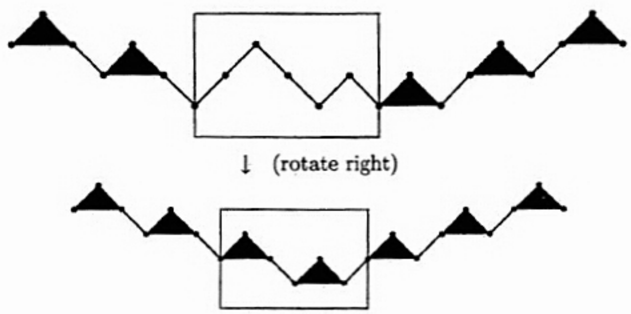

Figure 3: From w to (u, v)

To go in the other direction, consider the Catalan factorization (w', r) of an arbitrary $w \in \Sigma^{2n}$. Write $w' = u'v'$, where the last letter of u' is the r-th z from the left in w. If r is odd, then there is a factorization $w' = xd_1zd_2zy$, where d_1 and d_2 are the Dyck words immediately before and after the r-th z in w'.

1'. If $r = 0$, then $u = \epsilon$, $v = v'$.

2'. If $r > 0$ is even, then $u = u'$, $v = v'$.

3'. If r is odd and $w' = xd_1zd_2zy$, then $u = x1d_10d_2$, $v = y$.

The three examples below for $n = 10$ illustrate these cases.

Example 1'. Given word $w = 10110011110010101101$. Catalan factorization of w is $w' = 101100zz11001010z10z$, $r = 0$, which is mapped to $u = \epsilon$, $v = 101100zz11001010z10z$, with $(\phi_0(u), \phi_1(v)) = (\epsilon, 10110011110010101101)$.

Example 2'. Given word $w = 10010110001100101101$. Catalan factorization of w is $w' = 10z101100z110010z10z$, $r = 2$, which is mapped to $u = 10z101100z$, $v = 110010z10z$, with $(\phi_0(u), \phi_1(v)) = (1001011000, 1100101101)$.

Example 3'. Given word $w = 10010010010110110110$. Catalan factorization of w is $w' = 10z10z10z10z10z10z10$, $r = 3$. Locating the transposed portion of u as marked on top of Figure 3, we obtain the second path in the figure, which is then mapped to $u = 10z10z110010$, $v = 10z10z10$, with $(\phi_0(u), \phi_1(v)) = (100100110010, 10110110)$.

The table in Figure 4 gives the complete set of binary words of length 4 and the corresponding Catalan pairs (u, v) of total length 4 produced by the

w	$(u\ ,\ v)$	w	$(u\ ,\ v)$
1111	$(\epsilon\ ,\ zzzz)$	0111	$(10\ ,\ zz)$
1110	$(\epsilon\ ,\ zz10)$	0110	$(10\ ,\ 10)$
1101	$(\epsilon\ ,\ z10z)$	0101	$(1010\ ,\ \epsilon)$
1100	$(\epsilon\ ,\ 1100)$	0100	$(z10z\ ,\ \epsilon)$
1011	$(\epsilon\ ,\ 10zz)$	0011	$(zz\ ,\ zz)$
1010	$(\epsilon\ ,\ 1010)$	0010	$(zz\ ,\ 10)$
1001	$(1100\ ,\ \epsilon)$	0001	$(zz10\ ,\ \epsilon)$
1000	$(10zz\ ,\ \epsilon)$	0000	$(zzzz\ ,\ \epsilon)$

Figure 4: Binary words w of length 4 and the corresponding Catalan pairs (u, v)

bijection.

We now consider q-analogues of the convolution for the central binomial coefficients. These are obtained by making use of various properties of this bijection.

4. q-Counting by Peaks

We define two statistics on binary and Catalan words. $P(w)$ denotes the number of *peaks* in w, i.e. the number of occurrences of the word 10 in w. The other statistics $P'(w)$ is defined via the Catalan factorization of w. If the Catalan factorization of w is

$$w' = d_0 z d_1 z d_2 \cdots d_{k-1} z d_k$$

where each $d_i \in \mathcal{D}$, then we set

$$P'(w) = \begin{cases} P(w) & \text{if } \ d_k \in 10\mathcal{D}\,, \\ P(w) - 1 & \text{if } \ d_k \notin 10\mathcal{D}\,. \end{cases}$$

In other words, consider the longest suffix of w which is a Dyck word. If this Dyck word does not start with 10, then $P'(w)$ is the number of peaks in w as before. Otherwise $P'(w)$ is the number of peaks in w less one. We will give the weight $P'(u)$ to u and $P(v)$ to v in a Catalan pair (u, v).

The bijection for the convolution identity (1) has the following property: If $w \in \Sigma^{2n}$ is the word that corresponds to the Catalan pair (u, v) of total length $2n$, then in cases 1 and 2,

$$P(w) = P(u) + P(v).$$

In case 3, we lose a peak iff in the factorization $d = 1\beta 0$, β starts with 0. Therefore when the longest suffix of u which is a Dyck word starts with 10, we lose a peak fron the contribution of u to w. Thus in this case

$$P(w) = P'(u) + P(v).$$

Therefore summing

$$q^{P'(u)+P(v)}$$

over all pairs of Catalan words of total length $2n$ gives the generating function of the words $w \in \Sigma^{2n}$ by the number of peaks:

$$\sum_{(u,v)} q^{P'(u)+P(v)} = \sum_{w\in\Sigma^{2n}} q^{P(w)}. \tag{4}$$

We first rewrite the generating function on the right of (4).

Lemma 1.

$$\sum_{w\in\Sigma^{2n}} q^{P(w)} = \sum_{k=0}^{n} \binom{2n+1}{2k+1} q^k. \tag{5}$$

Proof. By unique factorization, a multiplicity-free regular expression for the language of $w \in \Sigma^*$ with exactly k peaks is

$$\mathbf{0^*1^*((10)0^*1^*)^k}.$$

It follows that the generating function of this language is

$$\frac{t^{2k}}{(1-t)^{2k+2}} = \sum_{n\geq 0} \binom{n+1}{2k+1} t^n .$$

Looking at even exponents of t, we get (5). □

Therefore, the bijection proves the following q-analogue of identity (1):

Theorem 2.

$$\sum_{k=0}^{n} \left[\sum_{|u|=2k} q^{P'(u)}\right] \cdot \left[\sum_{|v|=2n-2k} q^{P(v)}\right] = \sum_{k=0}^{n} \binom{2n+1}{2k+1} q^k, \tag{6}$$

where the inner summations are over Catalan words u and v.

5. q-Counting by the Number of 1's

Clearly, the expression in (3) q-counts the 4^n binary words of length $2n$ by the number of 1's they contain. On the other hand, the bijection is not additive with respect to the number of 1's, i.e. if $w \in \Sigma^{2n}$ corresponds to the pair of Catalan words (u, v), in general

$$|w|_1 \neq |u|_1 + |v|_1.$$

This is also the case if we start with the binary words that correspond to the Catalan words obtained by replacing the first half of z's by 0 and the second half by 1.

Next we consider how the number of 1's changes under the bijection. Suppose we start with a pair of Catalan words (u, v), where $|u| = 2k$ and $|v| = 2n - 2k$ for some k with $|u|_z = 2r$ and $|v|_z = 2s$. If $w \in \Sigma^{2n}$ is the word that corresponds to this pair, then we have

$$|w|_1 = n - r + s.$$

The reason for this is that

$$|w|_1 = |\phi_0(u)|_1 + |\phi_1(v)|_1$$

and

$$|\phi_0(u)|_1 = |u|_1 = k - r, \quad |\phi_1(v)|_1 = |v|_1 + 2s = n - k + s.$$

Since $|w| = |u| + |v| = 2n$, this means that

$$q^{-\frac{|u|_z}{2}} t^{|u|} q^{\frac{|v|_z}{2}} t^{|v|} = q^{|w|_1 - \frac{|w|}{2}} t^{|w|} . \tag{7}$$

Next we look at the generating function of Catalan words by the number of z's.

Lemma 3.

$$f(q,t) = \sum_u q^{|u|_z} t^{|u|} = \frac{2}{1 - q^2 + (1+q^2)\sqrt{1-4t^2}}, \tag{8}$$

where the summation is over all Catalan words u.

Proof. By unique factorization, a multiplicity-free expression for the language of Catalan words with $2r$ occurrences of z is

$$\mathcal{D}(\mathbf{z}\mathcal{D})^{2\mathbf{r}} . \tag{9}$$

Since the generating function of $\mathcal{D}$ by length is

$$\frac{1 - \sqrt{1-4t^2}}{2t^2},$$

the generating function of (9) is

$$q^{2r}t^{2r}\left[\frac{1-\sqrt{1-4t^2}}{2t^2}\right]^{2r+1}, \tag{10}$$

where the exponent of t is the total length and the exponent of q keeps track of the number of z's in the Catalan factorization. Summing (10) over r, we have a geometric series for the generating function $f(q,t)$. This simplifies to the expression in (8). □

Using (7)

$$f(q^{-\frac{1}{2}},t)f(q^{\frac{1}{2}},t) = \sum_{|w| \text{ even}} q^{|w|_1-\frac{|w|}{2}}t^{|w|}$$

$$= \sum_{n\geq 0} t^{2n}\sum_{w\in\Sigma^{2n}} q^{|w|_1-n} = \sum_{n\geq 0} t^{2n}q^{-n}\sum_{k=0}^{2n}\binom{2n}{k}q^k. \tag{11}$$

Let now

$$p_k(q) = \sum_u q^{|u|_z},$$

where the summation is over all Catalan words of length $2k$. The coefficient of t^{2n} on the left-hand side in (11) is

$$\sum_{k=0}^{n} p_k(q^{-\frac{1}{2}})p_{n-k}(q^{\frac{1}{2}}).$$

Therefore we obtain the q-analogue of the convolution identity (1) by the number of 1's in w as

$$\sum_{k=0}^{n} q^n p_k(q^{-\frac{1}{2}})p_{n-k}(q^{\frac{1}{2}}) = \sum_{k=0}^{2n}\binom{2n}{k}q^k.$$

We can state this as follows:

Theorem 4.

$$\sum_{k=0}^{n} q^n\left[\sum_{|u|=2k} q^{-\frac{|u|_z}{2}}\right]\left[\sum_{|v|=2n-2k} q^{\frac{|v|_z}{2}}\right] = (1+q)^{2n}, \tag{12}$$

where the inner summations are over Catalan words u and v.

As an example, we compute

$$\begin{aligned} p_0(q) &= 1, \\ p_1(q) &= 1+q^2, \end{aligned}$$

$$p_2(q) \;=\; 2+3q^2+q^4\,,$$

and for $n=2$,

$$q^2\left[p_2(q^{\frac{1}{2}})+p_1(q^{-\frac{1}{2}})p_1(q^{\frac{1}{2}})+p_2(q^{-\frac{1}{2}})\right]$$
$$=q^2\left(2+3q+q^2+(1+q^{-1})(1+q)+2+3q^{-1}+q^{-2}\right)$$
$$=q^2\left(q^{-2}+4q^{-1}+6+4q+q^2\right)=1+4q+6q^2+4q^3+q^4\,.$$

6. Symmetric Cases and Conclusions

Recall that the original convolution identity (1) is immediate by generating function methods since

$$\sum_{n\geq 0}4^n t^n=\frac{1}{1-4t} \quad \text{and} \quad \sum_{n\geq 0}\binom{2n}{n}t^n=\frac{1}{\sqrt{1-4t}}\,.$$

We try to work backwards to a more symmetric looking construction for a convolution q-analogue. If we take the square root of the generating function of the right-hand side expression, a symmetric interpretation would q-count Catalan words by some statistic which gives the coefficient polynomials $x_n(q)$, and which is symmetric, i.e. the q-count is such that the contribution from u is the same in pair (u,v) as it is in (v,u). We see that neither (6), nor (12) is symmetric in this sense.

First we consider the polynomials $x_n(q)$ in the case of q-counting by the number of 1's. The generating function of the right-hand side of (12) is

$$\sum_{n\geq 0}(1+q)^{2n}t^{2n}=\frac{1}{1-(1+q)^2t^2}\,.$$

Therefore

$$\frac{1}{\sqrt{1-(1+q)^2t^2}} \;=\; \sum_{n\geq 0}\frac{1}{4^n}\binom{2n}{n}(1+q)^{2n}t^{2n}$$
$$=\; 1+\frac{1}{2}(1+q)^2t^2+\frac{3}{8}(1+q)^4t^4+\frac{5}{16}(1+q)^6t^6+\cdots$$

In this case, the coefficient polynomials $x_n(q)$ are given by

$$x_n(q)=\frac{1}{4^n}\binom{2n}{n}(1+q)^{2n}.$$

But

$$\sum_{k=0}^{n}x_k(q)x_{n-k}(q)=\sum_{k=0}^{n}\frac{1}{4^n}(1+q)^{2n}A_kA_{n-k}=(1+q)^{2n}$$

because this is just identity (1) multiplied by $(1+q)^{2n}$.

For the case of q-counting by peaks, the right-hand side of (6) is

$$\sum_{k=0}^{n} \binom{2n+1}{2k+1} q^k .$$

Its generating function can be found as follows:

$$\begin{aligned}
\sum_{n\geq 0}\left[\sum_{k=0}^{n}\binom{2n+1}{2k+1}q^k\right]t^n &= \sum_{k\geq 0} q^k \left[\sum_{n\geq 0}\binom{2n+1}{2k+1}t^n\right] \\
&= \sum_{k\geq 0}\frac{q^k t^k}{2(1-t)^{2k+2}}\left[(1+\sqrt{t})^{2k+2}+(1-\sqrt{t})^{2k+2}\right] \\
&= \frac{1-(q-1)t}{1-2(q+1)t+(q-1)^2t^2} .
\end{aligned}$$

This means that a desirable q-analogue of $\binom{2n}{n}$ in this case should have as generating function, the square root of this. Call this q-analogue $x_n(q)$. Then

$$\frac{\sqrt{1-(q-1)t}}{\sqrt{1-2(q+1)t+(q-1)^2t^2}} = \sum_{n\geq 0} x_n(q)t^n .$$

A few of these polynomials are as follows:

$$\begin{aligned}
x_0(q) &= 1, \\
x_1(q) &= \frac{1}{2}(q+3), \\
x_2(q) &= \frac{1}{8}\left(3q^2+34q+11\right), \\
x_3(q) &= \frac{1}{16}\left(5q^3+125q^2+167q+23\right), \\
x_4(q) &= \frac{1}{128}\left(35q^4+1540q^3+4674q^2+2532q+179\right), \\
x_5(q) &= \frac{1}{256}\left(63q^5+4305q^4+23334q^3+28338q^2+8107q+365\right).
\end{aligned}$$

Calculating for $n=2$ as an example, we have

$$\begin{aligned}
x_0(q)x_2(q)+x_1(q)x_1(q)+x_2(q)x_0(q) &= \frac{1}{4}\left(3q^2+34q+11\right)+\frac{1}{4}(q+3)^2 \\
&= 5+10q+q^2 .
\end{aligned}$$

The polynomials $x_n(q)$ are q-analogues of $A_n = \binom{2n}{n}$, so that $x_n(1) = A_n$.

The leading coefficient of $x_n(q)$ is

$$\frac{\binom{2n}{n}}{4^n}.$$

What is required for a symmetric q-analogue of (1) by the number of peaks on the right-hand side is an interpretation of the $x_n(q)$ as q-counting Catalan words by a statistics, and a bijection which turns the sum of the statistics for u and v into the number of peaks of the resulting binary word w.

Acknowledgements

I would like to acknowledge the lectures of X. Viennot on bijective combinatorics given at UCSD in 1981 which contained a wealth of material on Catalan numbers, Dyck paths, and Catalan factorization.

References

[1] Ö. Eğecioğlu, A bijection for the convolution of central binomial coefficients, *Technical Report TRCS99-16*, Department of Computer Science, University of California at Santa Barbara (May 1999).

[2] Ö. Eğecioğlu, A. King, Random walks and Catalan factorization, *Congressus Numerantium*, **138** (1999), 129-140.

[3] D.J. Kleitman, A note on some subset identities, *Studies in Applied Math.*, **54** (1975), 289-293.

[4] J. Riordan, *An Introduction to Combinatorial Analysis*, Princeton University Press, Princeton, NJ (1978).

[5] X.G. Viennot, *Combinatorics Seminar Notes*, University of California at San Diego (1981).