Random Walks and Catalan Factorization

Ömer Eğecioğlu
Department of Computer Science
University of California Santa Barbara, CA 93106, U.S.A.
omer@cs.ucsb.edu

Alastair King,
Abdus Salam I.C.T.P.,
Strada Costiera 11, 34014 Trieste, Italy.
a.king@ictp.trieste.it

Abstract

In the theory of random walks, it is notable that the central binomial coefficients $\binom{2n}{n}$ count the number of walks of three different special types, which may be described as 'balanced', 'non-negative' and 'non-zero'. One of these coincidences is equivalent to the well-known convolution identity

$$\sum_{p+q=n} \binom{2p}{p}\binom{2q}{q} = 2^{2n}.$$

This article brings together several proofs of this 'ubiquity of central binomial coefficients' by presenting various relations between these classes of walks and combinatorial constructions that lead to the convolution identity. In particular, new natural bijections for the convolution identity based on the unifying idea of Catalan factorization are described.

Keywords: Central binomial coefficient, random walk, Dyck word, Catalan factorization, bijection, convolution.

1 The ubiquity of central binomial coefficients

A (random) walk of length N is a sequence $w = (\varepsilon_n)_{n=1}^{N}$ of elementary steps $\varepsilon \in \{+1, -1\}$, which we shall call, respectively, up-steps and down-steps. The lattice path corresponding to the walk is given by the partial sums $S_n(w) = \sum_{i=1}^n \varepsilon_i$ with diagonal steps as shown in Figure 1. The usual convention is that the initial value/level is $S_0 = 0$, although we shall occasionally speak of walks starting at levels other than 0. It is notationally convenient to encode a walk as a binary word, i.e. a sequence $w = (a_n)_{n=1}^N$ of letters $a \in \{1, 0\}$, by writing 1 for $+1$ and 0 for -1.

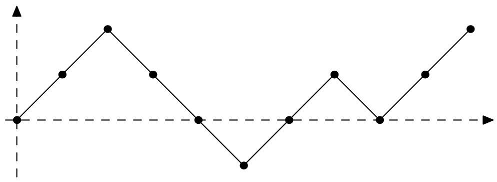

Figure 1: The lattice path of the walk $w = 1100011011$.

For the purposes of this article, we introduce the following terminology for certain special types of walks.

Definition 1 *A walk w of length N is*

(i) **balanced** *if $S_N(w) = 0$, i.e. it contains the same number of up-steps as down-steps. Hence, a balanced walk has even length.*

(ii) **non-negative** *if $S_n(w) \geq 0$, for $1 \leq n \leq N$, i.e. the level of the walk never falls below its initial level.*

(iii) **non-zero** *if $S_n(w) \neq 0$, for $1 \leq n \leq N$, i.e. the walk never returns to its initial level. A non-zero walk is either* positive *or* negative *depending on whether $S_n(w) > 0$ or $S_n(w) < 0$ for all $n > 1$. Reversing all the steps, i.e. reflecting the graph in the horizontal axis, provides a natural bijection between the positive and negative walks.*

Figure 2 shows all the balanced, non-negative and non-zero walks of length 4 as lattice paths and illustrates the following general result, well-known in the theory of random walks (cf. [1]).

Theorem 1 *The central binomial coefficient $\binom{2n}{n}$ counts the number of*

(i) balanced walks of length $2n$,

(ii) non-negative walks of length $2n$,

(iii) non-zero walks of length $2n$.

Proof Put $A_n = \binom{2n}{n}$. The number of balanced walks is clearly equal to A_n, because we must choose precisely n of the $2n$ steps as up-steps. A common way to prove the rest of the theorem is via the 'ballot problem' setting

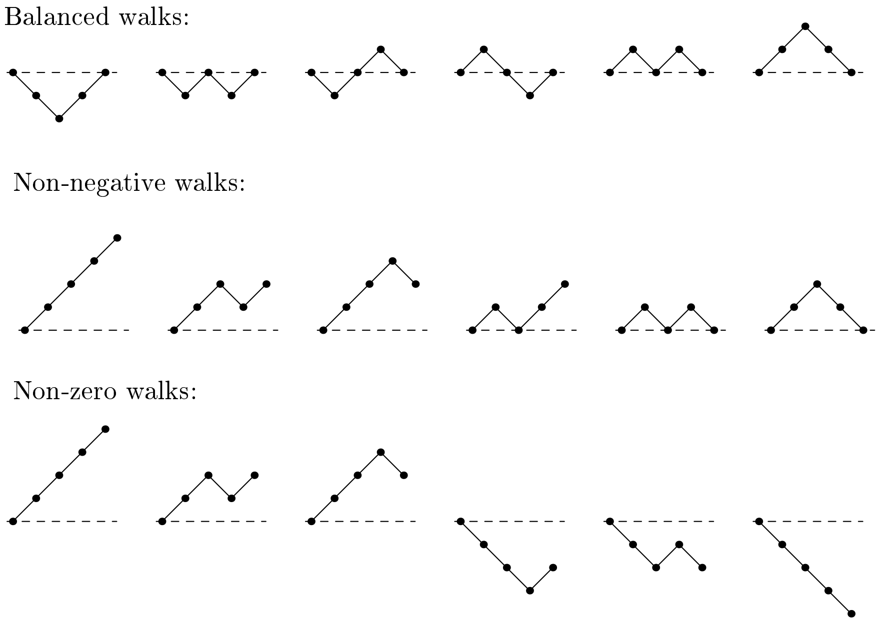

Figure 2: Three special types of walks of length 4.

([1]). By a standard reflection argument, the number of non-negative walks with m up-steps and k down-steps, where $k \le m$, is

$$B(m,k) = \binom{m+k}{k} - \binom{m+k}{k-1}. \tag{1}$$

We can then easily count all non-negative walks of length $2n$, obtaining

$$\sum_{k=0}^{n} B(2n-k,k) = A_n.$$

Furthermore a positive walk of length $2n$ consists of an up-step followed by an arbitrary non-negative walk of length $2n-1$. Therefore the number of positive walks of length $2n$ is equal to the number of non-negative walks of length $2n-1$, which is

$$\sum_{k=0}^{n-1} B(2n-1-k,k) = \binom{2n-1}{n} = \tfrac{1}{2}A_n.$$

Hence the number of non-zero walks of length $2n$ is also A_n. $\square$

The problem of counting non-negative walks also arises in the representation theory of the symmetric group. In this context, such walks are known as (two-letter) *Yamanouchi symbols* ([3] Chap. 7), and (1) gives the dimensions of certain irreducible representations of the symmetric group S_{n+m}.

A more combinatorial proof of Theorem 1, which constructs explicit bijections between the sets of walks, is sketched by Feller ([1] Problem III.10.7) and attributed to E. Nelson.[1]

Proof [Nelson's combinatorial proof of Theorem 1]

First, we describe a bijection between balanced walks and non-negative walks. Take the 'initial' segment of a balanced walk to be the part that ends at the first time it reaches its minimum value. Take the 'final' segment of a non-negative walk to be the part that starts from the last time it takes half its final value. The bijection takes the initial segment of a balanced walk, reverses the signs and order of the steps, and places it at the end of the walk. The inverse bijection takes the final segment of a non-negative walk, reverses the signs and order of the steps, and places it at the beginning of the walk.

Second, we describe a similar bijection between balanced walks and non-zero walks. Take the 'initial' segment of a balanced walk to be up to the first time it reaches either its minimum value, for walks that start with a down-step, or its maximum value, for walks that start with an up-step. Take the 'initial' segment of a non-zero walk to be up to the last time it reaches half its final value either with an up-step, for positive walks, or with a down-step, for negative walks. The bijection and its inverse reverse the signs and order of the steps in the initial segments. □

The bijections described above are constructed by factorizing one walk into two pieces and using the pieces to construct a new walk. In the process, significant global changes are made to the walks. The main goal of the present article is to describe are more subtle factorization of a walk called Catalan factorization, which may be used to construct bijections that only make local changes to the walks, i.e. only reverse the signs of certain critical steps.

2 Factorization and convolution identities

Before describing Catalan factorization, we discuss some other aspects of the relationship between factorization and enumeration of walks.

As a first example, observe that any walk w of length N has a unique factorization $w = uv$ into a balanced walk u of length $2k$ followed by a non-zero walk v of length $N - 2k$. This is done by finding the 'last return

[1]Strictly, only the first argument is sketched, but the two are sufficiently similar to reasonably attribute both to Nelson.

to 0', i.e. the last value of k for which $S_{2k}(w) = 0$. For example, for the walk w in Figure 1, we get $u = 11000110$ and $v = 11$. If $N = 2n$, we may use Theorem 1 to deduce the convolution identity

$$\sum_{k=0}^{n} A_k A_{n-k} = 2^{2n} \tag{2}$$

as follows. In the sum, the first factor counts the number of balanced walks of length $2k$, while the second counts the number of non-zero walks of length $2n - 2k$ and the total 2^{2n}, of course, counts all walks of length $2n$. A moment's thought shows that (2) is actually equivalent to the fact that A_{n-k} counts the number of non-zero walks of length $2n - 2k$, because whatever this number is, it is the unique correct second factor for this convolution identity.

Now (2) has an entirely independent proof as follows. If we introduce the generating function $A(t)$ for the central binomial coefficients

$$A(t) = \sum_{n=0}^{\infty} A_n t^n = 1 + 2t + 6t^2 + 20t^3 + \cdots,$$

then (2) is equivalent to the identity $A(t)^2 = (1 - 4t)^{-1}$, or

$$A(t) = (1 - 4t)^{-\frac{1}{2}}. \tag{3}$$

But now we may simply apply Newton's expansion formula, i.e. the binomial theorem with fractional powers, to the right-hand-side of (3) and see that the coefficient of t^n is A_n.

A second example is the use of factorization to count excursions.

Definition 2 *An* **excursion** *is a walk which is non-negative and balanced. These are also known as* **Dyck paths**. *The corresponding binary words are* **Dyck words**.

It is well-known, e.g. as a special case of (1), that the number of excursions of length $2n$ is the n-th Catalan number

$$C_n = \tfrac{1}{n+1} A_n. \tag{4}$$

We include here the standard derivation of the generating function for the Catalan numbers for completeness.

Note that excursions may be counted recursively, starting from the observation that every excursion w has a unique factorization $w = 1s0t$, where s and t are excursions of shorter length. In this case, the point at which t starts is the first time that the excursion returns to 0. An example of such a factorization is illustrated in Figure 3.

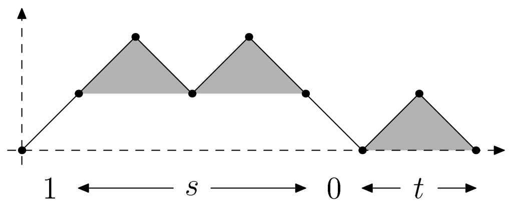

Figure 3: Factorizing an excursion into two shorter excursions.

Thus, if we chose to define C_n to be the number of excursions of length $2n$, we would have the convolution identity

$$\sum_{k=0}^{n} C_k C_{n-k} = C_{n+1} \quad \Longleftrightarrow \quad C(t)^2 = \frac{C(t)-1}{t} \tag{5}$$

where $C(t) = \sum_{n\geq 0} C_n t^n$ is the generating function. This equation for $C(t)$ is easily solved to give

$$C(t) = \frac{1-(1-4t)^{\frac{1}{2}}}{2t} \tag{6}$$

and Newton's expansion formula then recovers (4).

It is worth noting in passing that the factorization $w = 1s0t$ gives rise recursively to the well-known bijection between excursions of length $2n$ and binary trees with n internal nodes (cf. [2] and [5] 2.3.1 Exercise 6).

There are two other convolution identities which C_n and A_n and which follow from factorizations. Before describing them, note that reflecting the whole walk gives a bijection between the number of balanced walks that start with a down-step and those which start with an up-step. Hence the number of balanced walks of length $2n$ that start with a down-step is $\frac{1}{2}A_n$. The trivial walk of length 0 may be treated as a degenerate case, provided one is careful to interpret $\frac{1}{2}A_0$ as 1 and not $\frac{1}{2}$.

First, observe that every balanced walk has a unique factorization into an excursion followed by a balanced walk that starts with a down-step or has length 0. Thus

$$\sum_{k=0}^{n} C_k(\tfrac{1}{2}A_{n-k}) = A_n \iff C(t)\left(\frac{A(t)+1}{2}\right) = A(t). \tag{7}$$

Second, observe that every balanced walk w of positive length that starts with a down-step has a unique factorization $w = 0s1t$, where s is an arbitrary balanced walk and t is an excursion. Thus,

$$\sum_{k=0}^{n} A_k C_{n-k} = \tfrac{1}{2} A_{n+1} \iff A(t)C(t) = \frac{A(t)-1}{2t}. \qquad (8)$$

Note that from (7) and (8) it is possible to deduce (3) and (6). On the other hand, by repeatedly applying the two factorizations above, we may identify in any balanced walk, certain distinguished or critical steps, arising as the 0 and 1 is the second factorization, between which the walk is an excursion. It turns out that this equivalent to finding the Catalan factorization, which can actually be defined for any walk and which we describe next.

3 Catalan factorization

Definition 3 *For any walk w define a* **critical down-step** *to be the first step to each level less than the initial level, and a* **critical up-step** *to be the last step from each level less than the final level. The* **Catalan factorization** *of a walk is obtained by replacing the critical steps (up or down) by a neutral symbol z.*

The remaining sequences of 0's and 1's that occur between two consecutive z's are always excursions. An example of Catalan factorization is shown in Figure 4. Figuratively speaking, the critical steps are those illuminated when light is shone from the left below the initial level and from the right below the final level; the intermediate excursions are the parts of the walk that remain occluded.

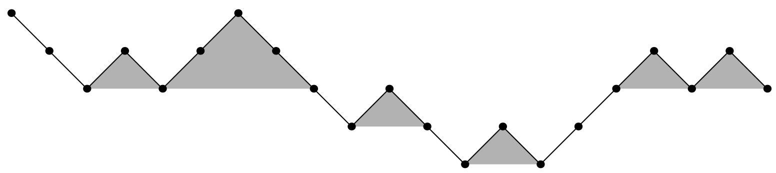

Figure 4: A walk with Catalan factorization $zz101100z10z10zz1010$.

The key feature of the Catalan factorization is that all the z's that represent down-steps in the original walk precede all the z's that represent up-steps. Therefore, to recover the walk from its Catalan factorization, it is necessary to know just one additional piece of information, namely, any one of the following quantities which we refer to as 'characteristic numbers':

(i) the number n_0 of critical down-steps, i.e. the difference between the initial and minimal levels,

(ii) the number n_1 of critical up-steps, i.e. the difference between the minimal and final levels,

(iii) the number $n_0 - n_1$, i.e. the difference between the initial and final levels.

These numbers satisfy the following constraints

(i) $n_0 \leq |w|_z$ (the total number of z's in w),

(ii) $n_1 \leq |w|_z$,

(iii) $|n_0 - n_1| \leq |w|_z$ and $n_0 - n_1 \equiv |w|_z \pmod 2$.

Definition 4 *A* **Catalan word** *is a word in the letters $\{0, 1, z\}$, for which each maximal segment in $\{0, 1\}$ is a Dyck word.*

A Catalan word is precisely the sort of word that may occur as the Catalan factorization of a walk. We may summarize the above discussion as follows.

Proposition 1 *Given any Catalan word and any value of one of the characteristic numbers n_0, n_1 or $n_0 - n_1$, which satisfies the corresponding constraint, there is a unique walk w with the given Catalan factorization and the given value of that characteristic number.*

Using this we may immediately find bijections between sets of Catalan words and sets of the various types of walks that we have considered earlier.

Proposition 2 *Catalan factorization provides natural bijections between*

(i) Catalan words of length $2n$,

(ii) balanced walks of length $2n$,

(iii) non-negative walks of length $2n$.

These restrict to natural bijections between

(i') Catalan words of length $2n$ that start with z,

(ii') balanced walks of length $2n$ that start with 0,

(iii') positive walks of length $2n$.

Proof For the first part, note that a walk is balanced if and only if $n_0 - n_1 = 0$, while a walk is non-negative if and only if $n_0 = 0$. Hence in both cases the Catalan factorization determines the walk by Proposition 1. The composite bijection between non-negative and balanced walks was described by Viennot in [6].

For the second part, note first that if a balanced (or any) walk starts with 0, then this will certainly be a critical down-step, while conversely in the Catalan factorization of a balanced walk the first z will always represent a down-step. On the other hand, a non-negative walk is positive, if and only if the first step is a critical up-step. $\square$

One other feature of Catalan words is that we may add a z at the beginning (or end) of the word and it remains a Catalan word. Conversely, given a Catalan word that begins (or ends) with z, we may remove this z and be left with a Catalan word. Thus we have the following.

Proposition 3 *The number of Catalan words of length N that begin (or end) with z is equal to the number of Catalan words of length $N-1$.*

Corollary 1 *The number of Catalan words of length $2n$ is A_n, while the number of length $2n-1$ is $\frac{1}{2}A_n$.*

Proof The first part is immediate from the first part of Proposition 2, while the second follows from Proposition 3 and the second part of Proposition 2, since the number of balanced walks that start with 0 is precisely half the total number of balanced walks, by reflection. $\square$

Proposition 2 essentially provides the promised combinatorial proof of Theorem 1 via Catalan factorization, because we may easily extend the bijection between positive walks and balanced walks that start with 0 to a bijection between non-zero walks and all balanced walks as follows. Starting with a negative walk, first reflect it, then take its Catalan factorization; after reinterpreting it as a balanced walk, reflect again to obtain a balanced walk that starts with 1. For example, the ordering of the walks in Figure 2 precisely reflects the bijections constructed via Catalan factorization.

An alternative strategy, which may be used to construct the same bijection between non-zero and balanced walks in a seemingly more symmetric way, was given by Kleitman [4].[2] This strategy compares two copies of a non-zero walk, the first with initial value zero and the second translated vertically so that the final value becomes zero. Working backwards from the end, find the first step that takes the second walk further away from zero than the first walk is at the same time. Define a new second walk by reversing the sign of this step and repeat the process. The process stops when the new second walk is balanced. An example is shown in Figure 5; the bold line marks the step to be reversed at each stage.

[2]Strictly, Kleitman constructs a slightly different bijection using the same strategy.

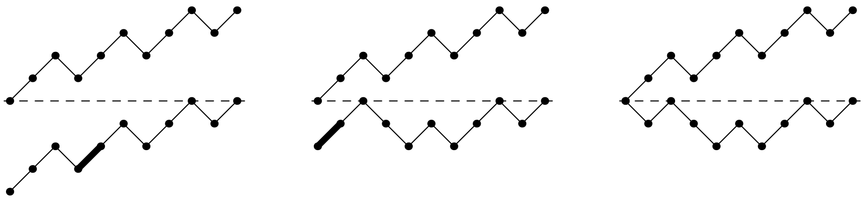

Figure 5: Kleitman's strategy applied to $w = 1101101101$.

4 Further uses of Catalan factorization

Following Viennot [6], we may prove (4) itself combinatorially using Catalan factorization, by constructing a bijection between Catalan words of length $2n$ and pairs consisting of a Dyck word of length $2n$ and an integer r, $0 \leq r \leq n$. Suppose we start with a Catalan word w of length $2n$. If there are $2k$ occurrences of z, we split w immediately after the k-th z to obtain two Catalan words. Replace the z in the first one of these words by 1 to obtain u, and replace the z's in the the second one by 0 to obtain a word v. Then uv is a Dyck word. We let r denote the number of occurrences of 1 in u. This is a bijection, as r identifies where the split has to be made in the Dyck word to go back to the original Catalan word by means of the juxtaposition of the Catalan factorizations of u and v.

We conclude with three direct combinatorial proofs of the convolution identity (2), all based roughly on the idea of dividing the Catalan factorization of a walk into initial and final segments for which the z's are to be interpreted as, respectively, all down-steps and all up-steps.

Proof [Proof 1] Consider the Catalan factorization of a walk of length $2n$ and suppose that it has the form $w = uzv$, where the given z is the last z which represents a down-step. Then u and v are both Catalan words, and all the z's in u are down-steps, while all the z's in v are up-steps. Hence knowing u and v determines w. Such a factorization is possible unless all the z's in w represent up-steps, in which case w itself determines the walk. Writing a_k for the number of Catalan words of length k, this counting procedure yields the formula

$$2^N = a_N + \sum_{k=1}^{N} a_{k-1}a_{N-k} \tag{9}$$

When $N = 2n$, Corollary 1 can be used to convert this to (2). The case $N = 2n - 1$ yields a slightly more complicated formula, which proves (2) recursively. □

This proof has the disadvantage that the counting method does not well reflect the convolution identity (2).

Proof [Proof 2] Consider again the Catalan factorization w of a walk of length $2n$. If the number of critical down-steps is even, then write $w = u_1u_2$, where the last letter of u_1 is the z representing the last critical down-step, or u_1 is trivial if there are no critical down-steps. In this case, (u_1, u_2) is a pair of Catalan words of even length and u_1, if non-trivial, ends with z. The original walk can be recovered from (u_1, u_2) because all z's in u_1 represent down-steps, while all z's in u_2 represent up-steps.

On the other hand, if the number of critical down-steps is odd, then write $w = v_1v_2$, where the last letter of v_1 is the z representing the first critical up-step. In this case, v_1 represents a walk $\alpha s0t1$, where s and t are excursions and α is either empty, or has even length and ends with a critical down-step. If we replace v_1 by the Catalan word v_1' that represents the rearranged walk $\alpha 1s0t$, then (v_1', v_2) is a pair of Catalan words of even length and v_1' ends with the non-trivial excursion $1s0t$. Thus we obtain precisely the pairs of even Catalan words that were not obtained in the first case. Note that the pair (v_1', v_2) determines the original walk, because the factorization $1s0t$ of the final excursion in v_1' is uniquely determined, as observed in Section 2. An example of this second case is shown in Figure 6; the additional modification takes place within the box, and is obtained by applying a single circular rotation to this portion of the path.

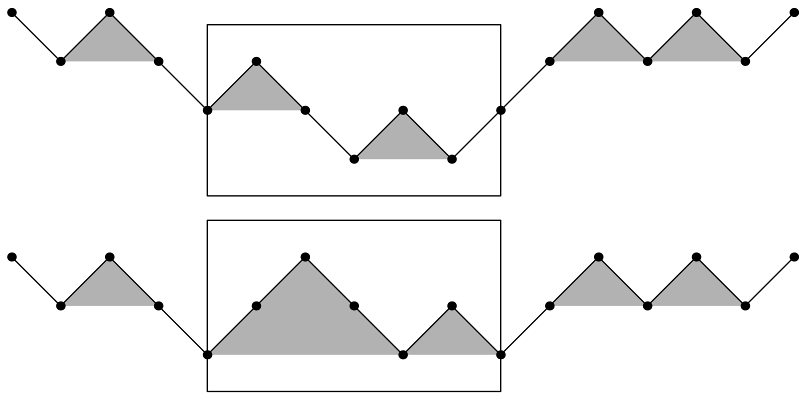

Figure 6: $0100100101110101 \mapsto (z10z110010, z1010z)$.

Thus we have constructed a one-one correspondence between arbitrary walks of length $2n$ and pairs of even Catalan words whose lengths sum to $2n$, thereby giving a very natural combinatorial proof of (2). $\square$

Proof [Proof 3] Consider a walk of length $2n+1$ containing an even (or odd) number of down-steps. Note that the number of such walks is 2^{2n}. The Catalan factorization of such a walk has an odd number of z's and there is therefore a unique $z = \hat{z}$ with the property that there are an even number of z's before $\hat{z}$ which all represent 0, and there are an even number of z's after $\hat{z}$ which all represent 1. Marking the position of $\hat{z}$ in the Catalan factorization determines the original walk, because the parity of the number of down-steps determines whether $\hat{z}$ represents 0 or 1. Thus, the original walk of length $2n+1$ is determined by two Catalan words of even length with $\hat{z}$ between them, thereby providing yet another combinatorial proof of (2). $\square$

Acknowledgement

The first author gratefully acknowledges 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] W. Feller, An Introduction to Probability Theory and its Applications, 3rd Edition, Wiley, New York, 1957.

[2] M. Gardner, *Catalan numbers: an integer sequence that materializes in unexpected places*, Scientific American **34** (1976) No. 6, 120-125.

[3] M. Hamermesh, Group Theory and Its Applications to Physical Problems, Dover Publications reprint, 1990.

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

[5] D.E. Knuth, The Art of Computer Programming, Volume 1 (Fundamental Algorithms), 2nd Edition, Addison-Wesley, 1973.

[6] X. Viennot, *Combinatorics Lectures Seminar Notes*, (unpublished), University of California at San Diego, Spring 1981.