Two New Bijections on Lattice Paths

Glenn Hurlbert*
Vikram Kamat†

Department of Mathematics and Statistics
Arizona State University
Tempe, Arizona 85287-1804

July 9, 2007

Abstract

Suppose $2n$ voters vote sequentially for one of two candidates. For how many such sequences does one candidate have strictly more votes than the other at each stage of the voting? The answer is $\binom{2n}{n}$ and, while easy enough to prove using generating functions, for example, only three combinatorial proofs exist, due to Kleitman, Gessel, and Callan. In this paper we present two new bijective proofs.

Key words. lattice path, bijection, ballot problem

*hurlbert@asu.edu, corresponding author
†vikram.kamat@asu.edu

1 Introduction

Suppose A and B are candidates and there are $2n$ voters, voting sequentially. In how many ways can A get n votes and B get n votes such that A is always ahead of or tied with B? This is the famous ballot problem, the solution of which is counted by the Catalan number C_n. A more general version of this problem is stated in terms of probability: let candidate A receive a votes and B receive b votes and compute the probability that A never falls behind B. This is known to be $\frac{a-b}{a+b}$, first proved by André [1] using his reflection principle (this also appears in [4] and again in [6]). A q-binomial variation of the problem appears in [8], and a weighted variation is found in [3]. Some have also considered an n-dimensional version by generalizing André's reflection proof to the many candidate ballot problem – one of these approaches can be found in [13]. Others [5, 9, 10] have considered the situation in which the number of votes of the two candidates remains close, while [11] discusses a number of variations in the context of generating functions. In this paper, we concern ourselves with another variation of the two candidate ballot problem discounting all instances when the two candidates are tied. It can be more formally expressed in terms of plus minus sequences.

Let $\mathcal{S}_n = \{s_1 \dots s_{2n} \mid s_i \in \{-1,+1\}\}$. For $S \in \mathcal{S}_n$ let $\sigma(S) = (\sigma_1, \dots, \sigma_{2n})$, where $\sigma_i = \sum_{j=1}^{i} s_i$. We write that $\sigma \neq 0$ (resp, $\sigma > 0$, $\sigma < 0$) when each $\sigma_i \neq 0$ (resp. $\sigma_i > 0$, $\sigma_i < 0$), and call S *zero-free* (resp. *positive*, *negative*) if $\sigma(S) \neq 0$ (resp. $\sigma(S) > 0$, $\sigma(S) < 0$). The set of zero-free (resp. positive, negative) sequences of $\mathcal{S}_n$ is denoted by $\mathcal{F}_n$ (resp. $\mathcal{P}_n$, $\mathcal{N}_n$). We will find it useful to denote σ_{2n} by $\sum S$. A sequence $S \in \mathcal{S}_n$ is *balanced* if $\sum S = 0$, and we use $\mathcal{B}_n$ to denote all balanced sequences in $\mathcal{S}_n$. We also use the notations $\mathcal{S}_n^+, \mathcal{F}_n^+, \mathcal{B}_n^+$ to denote those sequences that start with $+1$ (with the obvious analogous definitions for $\mathcal{S}_n^-$, etc.). Note that $\mathcal{F}_n^+ = \mathcal{P}_n$, $\mathcal{F}_n^- = \mathcal{N}_n$, and $|\mathcal{P}_n| = |\mathcal{N}_n|$.

It is known that $|\mathcal{F}_n| = |\mathcal{B}_n|$ for every n and, to our knowledge, only two bijections have appeared in print, due to Kleitman [7] and Gessel (see [12]). Another, due to Callan and similar to our second, appears on his website (see [2]).

Here we give two new bijections for this result, one indirect (Section 2.1) and one direct (Section 2.2). Of course, it is enough to show these for sequences in $\mathcal{S}_n^+$. That our direct bijection g (see Section 2.2) differs from that of Kleitman's bijection k can be seen for virtually any sequence $P \in P_n$ for any n, such as the following.

$$\begin{aligned}
P &= ++-+++---+++--++++-+--+++++- \\
\mathsf{g}(P) &= +-+---++++--++++--+-++------+ \\
\mathsf{k}(P) &= ---+++---+++----++-+--+++++-
\end{aligned}$$

In addition, ours is considerably simpler to navigate. Like Kleitman's, Gessel's and Callan's bijections do not preserve the first coordinate. However, one might notice that ours is essentially the negative of Callan's.

As is well known, one of the interesting applications of this result is the derivation of the generating function $F(x)$ for the sequence $\{\binom{2n}{n}\}$. Indeed, the factoring of a sequence $S \in \mathcal{S}_n$ into its maximum length balanced initial subsequence and corresponding terminal zero-free subsequence results in the relation $|\mathcal{S}_n| = \sum_k |\mathcal{B}_k||\mathcal{F}_{n-k}|$, from which the bijection yields

$$4^n = \sum_k \binom{2k}{k}\binom{2n-2k}{n-k}, \tag{1}$$

one short convolutional step from proving that $F(x) = (1-4x)^{-1/2}$. Marta Sved [12] recounts the history of Identity (1) and notes some combinatorial proofs of it that were submitted by readers.

Most satisfyingly, this paper is the result of the first author challenging the students in his graduate combinatorics course to find such a bijection.

In this article we use the notations $[s] = \{1, 2, \ldots, s\}$ and $(r, s) = [r+1, s-1] = \{r+1, \ldots, s-1\}$.

2 Proofs

2.1 Indirect Bijection

Let $\mathcal{S}^+_{n,k}$ be the set of all $S \in \mathcal{S}^+_n$ with $\sum S = 2k$. Such an S contains $(n-k)$ '-1's and $(n+k)$ '1's. Therefore,

$$|\mathcal{S}^+_{n,k}| = \binom{2n-1}{n-k},$$

since each such S begins with a '1'. For each $1 \le k \le n$, let $\mathcal{P}_{n,k}$ be the set of all $P \in \mathcal{P}_n$ having $\sum P = 2k$. Then $|\mathcal{P}_n| = \sum_{k=1}^{n} |\mathcal{P}_{n,k}|$. Define $\mathcal{T}^+_{n,k} = \mathcal{S}^+_{n,k} - \mathcal{P}_{n,k}$, so that

$$|\mathcal{P}_{n,k}| = |\mathcal{S}^+_{n,k}| - |\mathcal{T}^+_{n,k}| = \binom{2n-1}{n-k} - |\mathcal{T}^+_{n,k}|.$$

Finally, let $\mathcal{S}^-_{n,k}$ be all sets $S \in \mathcal{S}^-_n$ for which $\sum S = 2k$. Because such an S begins with a '-1', we have

$$|\mathcal{S}^-_{n,k}| = \binom{2n-1}{n-(k+1)}.$$

If $|\mathcal{T}^+_{n,k}| = |\mathcal{S}^-_{n,k}|$, then

$$\begin{aligned}\sum_{k=1}^{n} |\mathcal{P}_{n,k}| &= \sum_{k=1}^{n}\left[\binom{2n-1}{n-k} - \binom{2n-1}{n-(k+1)}\right] \\ &= \binom{2n-1}{n-1},\end{aligned}$$

by telescoping. Thus, we obtain $|\mathcal{F}_n| = \binom{2n}{n} = |\mathcal{B}_n|$ if $|\mathcal{T}^+_{n,k}| = |\mathcal{S}^-_{n,k}|$. We show that this indeed holds by demonstrating a bijection between the two sets.

From $\mathcal{T}^+_{n,k}$ to $\mathcal{S}^-_{n,k}$.

Let $T \in \mathcal{T}^+_{n,k}$. We can factor T as follows:

$$T = T_1 T_2$$

where T_1 is the smallest balanced initial subsequence, say of length m. (In lattice path language, m is the first step that hits the diagonal $y = x$.) By definition every sequence in $\mathcal{T}^+_{n,k}$ can be represented this way. Let $\overline{T_1}$ be the subsequence of length m obtained by negating every term of T_1. Now, define a function $f\colon \mathcal{T}^+_{n,k} \to \mathcal{S}^-_{n,k}$ as follows:

$$f(T) = f(T_1 T_2) = \overline{T_1} T_2 \ .$$

Consider an example. Let

$$T \quad = \quad ++-++-+--\boxed{-}++-+++++++-+ \ .$$

For each sequence, we define a path which is obtained by starting from the origin and taking a step $(1,1)$ for every $+$ and a step $(1,-1)$ for every $-$. Paths which never go below the x-axis are called Dyck paths. It can be seen that $\mathcal{P}_n$ counts all Dyck paths which touch the x-axis only at the origin. The path for T is given below:

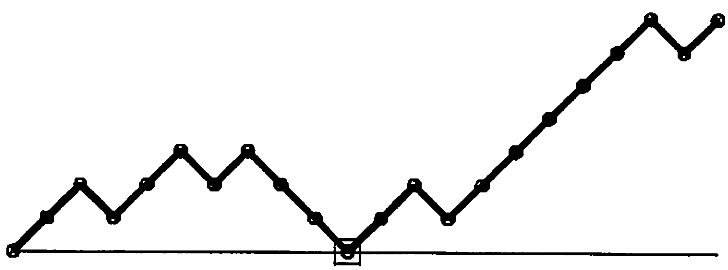

Figure 1: Path for T

Here, we see that $T \in \mathcal{T}^+_{11,4}$ and $m = 10$ (with the 10th position boxed). Also,

$$f(T) \quad = \quad --+--+-++\boxed{+}++-+++++++-+ \ .$$

Figure 2 gives the corresponding path for $f(T)$.

Clearly, $f(T)$ belongs to $\mathcal{S}^-_{11,4}$. The reason why this will hold in general is because only a balanced subsequence is negated and so σ is unchanged. Moreover, the smallest balanced subsequence always contains the first element(which is a 1) and thus, $f(T)$ will always have -1 as its first element.

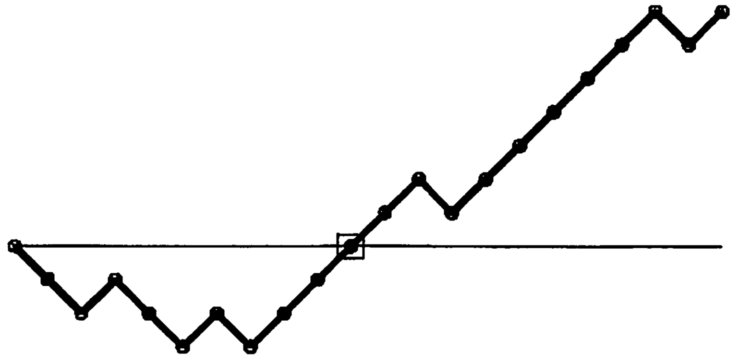

Figure 2: Path for $f(T)$

From $\mathcal{S}^-_{n,k}$ to $\mathcal{T}^+_{n,k}$.

Analogously, define a function $g\colon \mathcal{S}^-_{n,k} \to \mathcal{T}^+_{n,k}$ as follows: for any $S \in \mathcal{S}^-_{n,k}$,

$$g(S) = g(S_1S_2) = \overline{S_1}S_2 \ ,$$

where $S = S_1S_2$ with S_1 its smallest balanced initial subsequence, and $\overline{S_1}$ the negation of S_1.

In the example above, we have

$$S \quad = \quad --+--+-++\boxed{+}++-+++++++-+ \ \in \ \mathcal{S}_{11,4}$$

and

$$g(S) \quad = \quad ++-++-+--\boxed{-}++-++++++++-+ \ .$$

Clearly, $g(S) \in \mathcal{T}^+_{11,4}$. As above, this holds in general since σ is unchanged and $g(S)$ begins with a '1' because S begins with a '-1'.

Theorem 1 *The functions $f\colon \mathcal{T}^+_{n,k} \to \mathcal{S}^-_{n,k}$ and $g\colon \mathcal{S}^-_{n,k} \to \mathcal{T}^+_{n,k}$ are inverses of each other.*

Proof. Let $T \in \mathcal{T}^+_{n,k}$ and let $S = f(T)$. We will show that $g(S) = T$. We have $T = T_1T_2$ where $|T_1| = m$, as defined above, and $S = \overline{T_1}T_2$. Now, when we apply function g to S, since T_1 is the smallest balanced initial subsequence in T, $\overline{T_1}$ is the smallest balanced initial subsequence in S. This gives $g(S) = T$.

A similar argument proves that, for any $S \in \mathcal{S}^-_{n,k}$, $f(T) = S$ when $g(S) = T$. □

2.2 Direct Bijection

As previously mentioned, we will give a bijection between $\mathcal{B}^+_n$ and $\mathcal{P}_n$. The obvious bijection between $\mathcal{B}^-_n$ and $\mathcal{N}_n$, and hence $\mathcal{B}_n$ and $\mathcal{F}_n$ follows.

From $\mathcal{B}_n^+$ to $\mathcal{P}_n$. For $B \in \mathcal{B}_n^+$ we define a set $\pi(B)$ of *peaks* of B as follows. Set $t = \max\{\sigma_i(B)\}$ and for $1 \le k \le t$ let π_k be the index of the left-most occurrence of k in $\sigma = \sigma(B)$: $\pi_k = \min\{i \mid \sigma_i = k\}$. For example, if

$$B = \boxed{+} - + - - - + + + \boxed{+} - - + + \boxed{+}\boxed{+} - - + - + + - - - - - +$$

then

$$\sigma(B) = (\boxed{1}, 0, 1, 0, -1, -2, -1, 0, 1, \boxed{2}, 1, 0, 1, 2, \boxed{3}, \boxed{4}, 3, 2, 3, 2, 3, 4, 3, 2, 1, 0, -1, 0)$$

and

$$\pi(B) = \{1, 10, 15, 16\} .$$

We have boxed in the peak locations as shown.

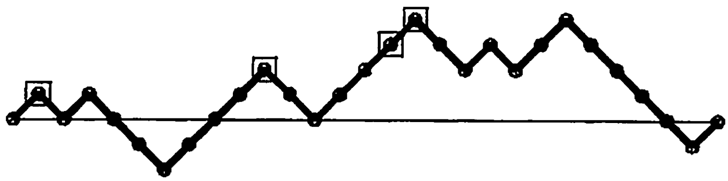

Figure 3: Path for B

Next we define the set of intervals $I_k = (\pi_k, \pi_{k+1})$, with $I = \cup I_k = [2n] - \pi(B)$. (Artificially, we set $\pi_{t+1} = 2n + 1$ in order to define I_t; in this case we have $\pi_5 = 29$ and $I_4 = (16, 29) = [17, 28]$.) The key property here is that $\sigma_i \le k$ for every $i \in I_k$.

Finally we define $\mathsf{f} = \mathsf{f}(B)$ by $\mathsf{f}_i = B_i$ for all $i \in \pi(B)$ and $\mathsf{f}_i = -B_i$ otherwise (for all $i \in I$). For this example, we obtain

$$\mathsf{f} = \boxed{+} + - + + + - - - \boxed{+} + + - - \boxed{+}\boxed{+} + + - + - - + + + + + -$$

and

$$\sigma(\mathsf{f}) = (\boxed{1}, 2, 1, 2, 3, 4, 3, 2, 1, \boxed{2}, 3, 4, 3, 2, \boxed{3}, \boxed{4}, 5, 6, 5, 6, 5, 4, 5, 6, 7, 8, 9, 8) .$$

(Figure 4 shows the corresponding path). Note that we have $\sigma(\mathsf{f}) > 0$, so that

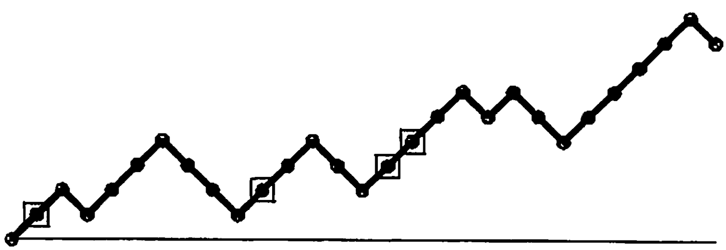

Figure 4: Path for f

$f(B) \in \mathcal{P}_n$. This holds in general for the following reason. By the definition of π_k we have $\sum B_{\pi_{k+1}-1} = \sum B_{\pi_k}$. This means that B is balanced on each interval I_k with $k < t$. Hence each $\sum f_{\pi_k} = \sum B_{\pi_k}$ and thus $\sigma(f)_i \geq k$ for every $i \in I_k$. In particular, $f_i \geq 1$ for all i.

From $\mathcal{P}_n$ to $\mathcal{B}_n^+$.

For $P \in \mathcal{P}_n$ we define a set $\Pi(P)$ of *pivots* of P as follows. Set $T = \frac{1}{2}\sum P$, let $\Pi_1 = 1$, and for $1 < k \leq T$ let Π_k be one more than the index of the rightmost occurrence of $k-1$ in $\sigma = \sigma(P)$: $\Pi_k = 1 + \max\{j \mid \sigma_j = k-1\}$. For example, if

$$P = \boxed{+} + - + + + - - - \boxed{+} + + - - \boxed{+}\boxed{+} + + - + - - + + + + + -$$

then

$$\sigma(P) = (\boxed{1}, 2, 1, 2, 3, 4, 3, 2, 1, \boxed{2}, 3, 4, 3, 2, \boxed{3}, \boxed{4}, 5, 6, 5, 6, 5, 4, 5, 6, 7, 8, 9, 8)$$

and

$$\Pi(P) = \{1, 10, 15, 16\} .$$

We have boxed in the pivot locations as shown.

Next we define the set of intervals $J_k = (\Pi_k, \Pi_{k+1})$, with $J = \cup J_k = [2n] - \Pi(P)$. (Artificially, we set $\Pi_{T+1} = 2n+1$ in order to define J_T; in this case we have $\Pi_5 = 29$ and $J_4 = (16, 29) = [17, 28]$.) The key property here is that $\sigma_j \geq k$ for every $j \in J_k$. Finally we define $g = g(P)$ by $g_j = P_j$ for all $j \in \Pi(P)$ and $g_j = -P_j$ otherwise (for all $j \in J$). For this example, we obtain

$$g = \boxed{+} - + - - - + + + \boxed{+} - - + + \boxed{+}\boxed{+} - - + - + + - - - - - +$$

and

$$\sigma(g) = (\boxed{1}, 0, 1, 0, -1, -2, -1, 0, 1, \boxed{2}, 1, 0, 1, 2, \boxed{3}, \boxed{4}, 3, 2, 3, 2, 3, 4, 3, 2, 1, 0, -1, 0) .$$

Note that we have $\sum g = 0$, so that $g(P) \in \mathcal{B}_n^+$. This holds in general for the following reason. By the definition of Π_k we have $\sum P_{\Pi_{k+1}-1} = \sum P_{\Pi_k}$. This means that P is balanced on each interval J_k with $k < T$. Hence each $\sum g_{\Pi_k} = \sum P_{\Pi_k}$ and thus $\sigma(g)_j \leq k$ for every $j \in J_k$. In particular, by the definition of T we have $\sum_{j \in J_T} P_j = T$, so that $\sum_{j \in J_T} g_j = -T$, and hence $\sum g = 0$.

Theorem 2 *The functions $f : \mathcal{B}_n^+ \to \mathcal{P}_n$ and $g : \mathcal{P}_n \to \mathcal{B}_n^+$ are bijections and, in fact, inverses of each other.*

Proof. The arguments above show that f and g are well-defined. That they are bijections will follow from their inverse relationship.

We suppose first that $f(B) = P$ for $B \in \mathcal{B}_n^+$, and show that $g(P) = B$. This will follow from showing inductively that $\Pi(P) = \pi(B)$. Of course, $\Pi_1 = 1 = \pi_1$, so assume that $\Pi_k = \pi_k$. Because we know for all k that $\sigma(B)_i \leq k$ for all $i \in I_k$ and that $\sigma(B)_{\pi_{k+1}} = k+1$, we therefore know for all k that $\sigma(P)_i \geq k$ for all

$i \in I_k$ and that $\sigma(P)_{\pi_{k+1}} = k+1$. In particular, the right-most occurrence of k in $\sigma(P)$ occurs with index $\pi_{k+1} - 1$; i.e. $\Pi_{k+1} = \pi_{k+1}$.

Next we suppose that $\mathsf{g}(P) = B$ for $P \in \mathcal{P}_n$, and show that $\mathsf{f}(B) = P$. As above, we show that $\pi(B) = \Pi(P)$ by induction. Again, $\pi_1 = 1 = \Pi_1$, so we assume that $\pi_k = \Pi_k$. We know for all k that $\sigma(P)_j \geq k$ for all $j \in J_k$ and that $\sigma(P)_{\Pi_{k+1}} = k+1$, and thus we know for all k that $\sigma(B)_j \leq k$ for all $j \in J_k$ and that $\sigma(B)_{\Pi_{k+1}} = k+1$. In particular, the left-most occurrence of $k+1$ in $\sigma(B)$ occurs with index Π_{k+1}; i.e. $\pi_{k+1} = \Pi_{k+1}$. $\square$

Acknowledgement

We thank Ira Gessel and David Callan for several helpful comments.

References

[1] D. André, Solution directe du probléme résola par M. Bertnand, C. R. Acad. Sci. Paris **105** (1887), 436–437.

[2] D. Callan, Bijections for the identity $4^n = \sum_{k=0}^{n} \binom{2k}{k}\binom{2(n-k)}{n-k}$, `http://www.stat.wisc.edu/~callan/papersother/`.

[3] S.P. Eu, T.S. Fu and Y.N. Yeh, Refined Chung-Feller theorems for lattice paths, J. Combin. Theory, Ser. A **112** (2005), 143–162.

[4] W. Feller, An introduction to probability theory and its applications, Vol. I, third ed., Wiley, New York, 1968.

[5] R.D. Fray and D.P. Roselle, Weighted lattice paths, Pacific J. Math. **37** (1971), no. 1, 85–96.

[6] P. Hilton and J. Pedersen, The ballot problem and Catalan numbers, Nieuw Arch. Wisk. (4) **8** (1990), no. 2, 209–216.

[7] D. Kleitman, A note on some subset identities, Studies in Appl. Math. **54** (1975), no. 4, 289–292.

[8] C. Krattenthaler and S.G. Mohanty, q-generalization of the ballot problem, Discrete Math. **126** (1994), no. 1-3, 195–208.

[9] N. Loehr, Note on André's reflection principle, Disc. Math. **280** (2004), 233–236.

[10] R.C. Lyness, Al Capone and the death ray, Math. Gaz. **25** (1941), 283–287.

[11] L. Shapiro and W-J. Woan, Some generating function proofs of old and new results in probability, Congr. Numer. **143** (2000), 193–205.

[12] M. Sved, Counting and recounting: the aftermath, Math. Intelligencer **6** (1984), no. 2, 44–45.

[13] D. Zeilberger, André's reflection proof generalized to the many-candidate ballot problem, Discrete Math. **44** (1983), no. 3, 325–326.