

Catalan Multijections

Garth Isaak* Larry Langley†

1 Introduction

The Catalan numbers $C_n = \frac{1}{n+1} \binom{2n}{n}$ are ubiquitous. There are thousands of papers and at least a half dozen books devoted to these numbers and their generalizations. There are hundreds of families of objects counted by the Catalan numbers $C_n = \frac{1}{n+1} \binom{2n}{n}$. We do not make any attempt to reference the vast literature on Catalan numbers. We note, for example, that there are a number of books or portions of books devoted to Catalan numbers. Four such are [8], [13], [14] and [19]. There are also, currently, several websites devoted to these numbers, their generalizations and their history.

Our focus here will be on one category of proofs related to these numbers and an exploration of when these proofs are, in a certain sense, different or are translations with the same proof structure but using different terminology.

There are many different approaches to proving that a family of objects is counted by the Catalan numbers. One approach is to find a bijection to another family counted by Catalan numbers. For direct proofs that a family is counted by the Catalan numbers there are several standard approaches including: using recursion and generating functions; using what is known as the reflection principle to count non Catalan sequences; using induction on a slightly more general family of objects and using uniform partitions. We will look only at versions of proofs of the last type.

To illustrate we start with two families counted by Catalan numbers.

Definition 1 *Ballot lists are strings of n 0's and n 1's where each initial substring has at least as many 1's as 0's. We will refer to the larger family of $\binom{2n}{n}$ binary strings with n 0's and n 1's as balanced binary strings.*

Definition 2 *We will call Catalan multisets, the size n multisets from $\{0, 1, \dots, n\}$ dominated by $\{0, 1, 2, \dots, n\}$ when listing its elements in nondecreasing order $a_1 \leq a_2 \leq a_n$. That*

*Department of Mathematics, Lehigh University, gisaak@lehigh.edu

†Department of Mathematics, University of the Pacific, llangley@pacific.edu

is, we have $a_i \leq i - 1$ for all i . For notational simplicity we drop parentheses and commas and write the multiset as a nondecreasing string. We will refer to the larger family of the $\binom{n+1+n-1}{n} = \binom{2n}{n}$ size n multisets from $\{0, 1, \dots, n\}$ simply as multisets since the size and ground set will remain the same in our examples.

When $n = 3$ we have Catalan number $C_3 = \frac{1}{3+1} \binom{6}{3} = \frac{1}{4} \cdot 20 = 5$. Also when $n = 3$ there are $\binom{6}{3} = 20$ balanced binary strings and $\binom{6}{3} = 20$ size 3 multisets from $\{0, 1, 2, 3\}$.

We display the 20 balanced binary strings, grouped into 5 classes of size 4:

111000	110100	110010	101100	101010
110001	101001	100101	011001	010101
100011	010011	001011	100110	010110
000111	001110	011100	001101	011010

Table 1: Cycle Lemma Partition $n = 3$

The ballot sequences are exactly the top entry in each group. This grouping is produced by the cycle lemma, which maps an arbitrary balanced binary string to a ballot list. The groupings are the equivalence classes formed by the preimages of this map.

We display the 20 size 3 multisets from $\{0, 1, 2, 3\}$, grouped into 5 classes of size 4:

000	001	002	011	012
111	112	113	122	123
222	223	022	233	023
333	033	133	003	013

Table 2: Chung-Feller-MacMahon Partition $n = 3$

The Catalan multisets are exactly the top entry in each group. This grouping is produced by one proof of the Chung-Feller-MacMahon Theorem, which maps an arbitrary multiset to a Catalan multiset. The groupings are the equivalence classes formed by the preimages of this map. We note that the Chung-Feller-MacMahon Theorem is often presented using somewhat different terminology. We make use of a version from [14] which in turn references an earlier paper of Feller.

We can then ask if these groupings are the same. For this we need to refer to some bijection between balanced binary strings and multisets.

One common bijection that maps a binary string to a multiset by recording the number of 0's preceding each 1. We will call this the 0 count bijection. This maps ballot lists to Catalan multisets. Using this bijection we reproduce the binary string table from the cycle lemma written as multisets:

000	001	002	011	012
003	013	023	113	123
033	133	233	022	122
333	222	111	223	112

Table 3: Cycle Lemma Partition $n = 3$, written as multisets using 0 count bijection

We have arranged the tables so locations of the images under the bijection are the same. That is 111000 maps to 000, 110100 maps to 001 etc. In particular the top rows are an example of the bijection between ballot lists and Catalan multisets. It appears that the two approaches yield different groupings. That is, Tables 2 and 3 have different groupings. If we map the multiset table back to balanced binary strings using the inverse of the 0 count bijection we get:

111000	110100	110010	101100	101010
011100	011010	011001	010110	010101
001110	001101	100110	001011	100101
000111	100011	010011	110001	101001

Table 4: Chung-Feller-MacMahon Partition $n = 3$, written as binary strings

Since this is the same bijection we expect that we again get different groupings. But, closer inspection reveals a pattern. If we take the complement of each string, swapping 0's and 1's, we do get the same partition. In particular the complement of the first and fifth groups under the cycle lemma are the first and fifth groups respectively under Chung-Feller-Macmahon. Similarly the complement of the second group is the fourth group, the complement of the third group is the second group and the complement of the fourth group is the third group.

We could pick a different bijection between balanced binary strings and multisets suggested by the observation of complementation above. Call the 1 count bijection the one that maps binary strings to multisets by recording the number of 1's preceding each 0.

Using this bijection we reproduce the binary string table from the cycle lemma written as multisets:

333	233	223	133	123
222	122	112	022	012
111	011	001	113	013
000	003	033	002	023

Table 5: Cycle Lemma Partition $n = 3$, written as multisets using 1 count bijection

Here we have again kept the locations of the objects the same. As expected from the observation about complementation of the binary strings, we can in this notation readily see that the groupings are the same, although rearranged. In particular, the 1 count

bijection does not map ballot lists to Catalan multisets. Had we swapped the role of 0 and 1 and defined ballot lists as those where each initial substring has at least as many 0's as 1's then the connection would have been immediate and this new version of ballot lists would map to Catalan multisets.

So to explore when uniform partition proofs are 'the same' or 'different' in terms of the partitions they produce, we will of course need to have a definition that depends on a bijection between the objects being considered. We will do this and then see what happens with the two examples described so far and a third uniform partition proof that appears to produce different partitions. This framework can also lead us to find translations of standard proofs into different terminology that might be of interest.

2 Definitions and History

2.1 Multijections

The proofs hinted at above make use of groupings into equivalence classes which can be defined via preimages of a function. While, for the examples above, we can also directly define the equivalence relation, it will be useful to think of these in terms of function preimages. We will introduce a shorthand terminology to make easy reference to this approach.

One of the first counting arguments students encounter in a combinatorics class determines the number $\binom{n}{k}$ of k element subsets of an n set by associating with each k subset its $k!$ permutations. These collections partition the $\frac{n!}{(n-k)!}$ length k permutations of an n set into parts of size $k!$ establishing that $\binom{n}{k} = \frac{1}{k!} \cdot \frac{n!}{(n-k)!}$. This sort of proof gets presented in various ways. One can use the product rule and show that the number, $\binom{n}{k}$ of k subsets times the number, $k!$ of permutations of a k set equals the number, $\frac{n!}{(n-k)!}$ of k permutations of an n set. Equating yields the formula. Alternatively, we partition the $\frac{n!}{(n-k)!}$ length k permutations of an n set into equivalence classes of size $k!$, two such permutations being equivalent if they have the same underlying set. Since each of the $\frac{1}{k!} \cdot \frac{n!}{(n-k)!}$ equivalence classes corresponds to a size k subset, we get the formula for subsets. We can think of the equivalence classes as being defined by the preimages of a function which maps k permutations to subsets by mapping to the underlying set of elements.

For easy reference we make the following definition.

Definition 3 *Assume that A and $\mathcal{A}$ are finite sets. We will say that a map α from A to $\mathcal{A}$ is a t uniform multijection if the preimages $\alpha^{-1}(a)$ all have size t . For a given function, the relation xRy if $f(x) = f(y)$ is trivially an equivalence relation so the preimages partition the domain A . Then, since all have the same size, t , we get $|\mathcal{A}| = \frac{1}{t}|A|$ for a t uniform multijection.*

In the subset example α maps length k permutations to the underlying set and α^{-1} associates a set with the $k!$ permutations of the set. We will drop t uniform from multijection as all of our examples will be $n + 1$ uniform.

2.2 Different vs Translation

In order to discuss whether two different multijjective proofs are ‘different’ in some sense or essentially ‘translations’ of the same partition into different notation we use the following.

Definition 4 *Let α be a t -uniform multijection from A to $\mathcal{A}$ and β a t -uniform multijection from B to $\mathcal{B}$ with $|A| = |B|$ and hence also $|\mathcal{A}| = |\mathcal{B}|$. Let σ be a bijection from A to B*

We will say that α and β are translations of each other with respect to σ if, under σ the preimages induce the same partition. That is, if $\alpha(a_1) = \alpha(a_2)$ implies $\beta(\sigma(a_1)) = \beta(\sigma(a_2))$ and hence also $\beta(b_1) = \beta(b_2)$ implies $\alpha(\sigma^{-1}(b_1)) = \alpha(\sigma^{-1}(b_2))$.

So, formally we cannot say that two multijections are different or translations without reference to a bijection σ . Informally, if we drop the explicit reference to the bijection, we would tend to think of two multijections as translations of each other if they are translations under some ‘natural’ or ‘standard’ choice of σ .

Since $|\mathcal{A}| = |\mathcal{B}|$ we can also think about bijections between these families. Indeed, for our examples these are families of objects counted by Catalan numbers and exploring such bijections very common in Catalan numbers. To include this more explicitly we expand our definition.

If we have t -uniform multijections that are translations under σ then a bijection between $\mathcal{A}$ and $\mathcal{B}$ is implied as there is a bijection between the groupings in the partition. This is π so that $\pi(\alpha(a)) = \beta(\sigma(a))$ for all $a \in A$.

With the arrows to the right in the following, the diagram commutes. We include arrows in both directions as a reminder that σ and π are bijections.

$$\begin{array}{ccc} & A & \xleftrightarrow{\sigma} B \\ \alpha \downarrow & & \downarrow \beta \\ & \mathcal{A} & \xleftrightarrow{\pi} \mathcal{B} \end{array}$$

This allows us to explore various questions related to uniform partition proofs for Catalan numbers. Indeed, if any three of the functions are specified we can ask about the implied fourth and if there is a nice description of it. In the case of σ , since what we need is to map equivalence classes to equivalence classes there are multiple possibilities.

For example, if we have two known uniform partition proofs, that is α, β , our question of whether they are ‘different’ or ‘translations’ is asking if there is some ‘nice’ σ and hence an implied π which makes the diagram commute. Of course the sizes tell us that such a σ exists. The question is whether there is a natural description. If we have such a σ we can ask what π is and whether or not it has a ‘nice’ description. We will discuss this throughout and in particular in Sections 3.4 and 4.1 and 4.2. Alternatively if we have a known uniform partition proof α and natural choices of σ and π , we can ask if we have a ‘nice’ description of the implied β . In other words a translation of a known proof to a new setting. We will see this in Section 4.3.

2.3 Catalan Families

There are hundreds of families of objects counted by the Catalan numbers $C_n = \frac{1}{n+1} \binom{2n}{n}$. The book [19] includes descriptions of bijections between many of these families. These would be π in our diagram. Our examples here will primarily make reference to only three Catalan families.

We have already defined ballot lists and Catalan multisets. We will refer to one further family.

Definition 5 *A string of n left (and n right) is called balanced. It is well formed if we can pair off the parentheses as open-closed pairs (). That is we can find a bijection between the left (and the right) so that the left (always appears first in the string. Equivalently, each initial segment has at least as many left parentheses as right.*

In [10] these are called CG-arrangements after John Conway and Richard Guy. We will use the terms well formed CG-arrangements and will refer to the larger family of all strings of n left (and n right) as general CG-arrangements.

Although well formed sequences of parentheses do not seem to appear explicitly in [19] there is a trivial bijection to ballot lists by identifying 1’s with (and 0’s with). Viewing these as parentheses provides a natural pairing of the parentheses that will be useful.

We note that bracketings which appear in [19], which also involve left and right parentheses, are the number of ways to use parentheses to multiply $n + 1$ numbers. These are different from CG-arrangements, which use parentheses but no other symbols. While both are counted by Catalan numbers, the underlying parentheses strings for well formed CG-arrangements and bracketings are different. Indeed quoting Guy [7]

[7] Before we say goodbye to the Catalan numbers, here’s an observation which may not be widely known. It originated from a discussion with John Conway only six months ago. What *is* well known is that Catalan numbers

are associated with parenthesization. By that most people mean the numbers of possible orders of n nonassociative operations, usually indicated by $n - 1$ pairs of parentheses: ... But they are also the numbers of ways of arranging n pairs of parentheses as a pattern, just for their own sake:... An examination of the symmetries in the two cases makes it unlikely that you'll find a direct combinatorial comparison.

See [10] and [15] for further discussion about connections (i.e., 'direct combinatorial comparisons') between these two families. In a broad sense what we will be looking at here is an examination of when certain proofs related to Catalan numbers have a 'direct comparison' in some 'natural' way.

2.4 Multijjective Proofs

Two standard multijjective approaches to Catalan numbers are frequently called the cycle lemma of Dvorsky and Motzkin and the Chung-Feller-MacMahon Theorem. See for example, [9], [16] or any of the books mentioned previously. We will refer to these as CDM and CFM respectively. Each of these goes well beyond the basic counting problem. We will omit details of the additional structure as well as formal proofs as they can be found in many standard references. We will provide only brief descriptions and examples of these approaches in order to understand our notion of 'different' multijjective proofs.

We will also describe a multijection from general CG-arrangements of parentheses to well formed CG-arrangements of parentheses and refer to this as CGP. We will see that this does seem to produce different groupings than CDM and CFM.

The Dvorsky-Motzkin Cycle Lemma and the Chung-Feller-Macmahon Theorem are typically phrased in terms of equivalence classes or uniform partitions, each part of which contains a Catalan object. In these cases the collection of Catalan objects is a subset of a larger set of size $\binom{2n}{n}$. We prefer the multijection terminology since it fits better with the possibility of proofs where the Catalan object has no such natural embedding. For example, binary trees are Catalan objects but there is no obvious larger class of $\binom{2n}{n}$ trees for which the binary trees are a special case. However, multijjective proofs using binary trees on n vertices as images to partition a set of size $\binom{2n}{n}$, for example balanced binary strings might also be interesting to examine. See [2], [4] for examples multijection examples related to plane trees.

3 Some Catalan Multijection Proofs

We now describe, mostly through examples, the three approaches to multijjective proofs for Catalan numbers that we consider here.

3.1 CDM

For our purposes, the cycle lemma of Devorsky and Motzkin, CDM, can be briefly described as follows. Given a balanced binary string append an additional 1 at the beginning, write down all cyclic shifts which start with a 1 and then record these strings, omitting the initial 1. The content of the cycle lemma is that exactly one of these will be a ballot list. We omit the proof as many nice proofs can be easily found in the literature. A nice geometric interpretation of the cycle lemma is in [6]. Some examples of CDM proofs and generalizations include [3], [11], [18].

As a multijection, α , the image of a balanced binary string is the ballot list from the process described above. The preimage of a ballot list under α is the set of all cyclic shift from the process applied to the ballot list.

To illustrate CDM as as described above we find the preimage of 110010 by first appending a 1 at the front to obtain 1110010 and then write down all cyclic shifts of this string which start with a 1.

```

1 1 1 0 0 1 0
  1 1 0 0 1 0 1
    1 0 0 1 0 1 1
      1 0 1 1 1 0 0

```

Removing the lead 1 from these we get the preimage of 110010 to be 110010, 100101, 001011, 011100.

When $n = 3$, in a similar manner we associate a set of 4 binary strings with three 1's and three 0's with each ballot list. This partitions all $\binom{6}{3} = 20$ such strings into 5 preimages each of size 4 each of which contains exactly one ballot list. We write down the ballot lists at the top of each preimage. This is Table 1 repeated here for convenience. The example above is the third column.

111000	110100	110010	101100	101010
110001	101001	100101	011001	010101
100011	010011	001011	100110	010110
000111	001110	011100	001101	011010

3.2 CFM

For our purposes the version we present of the Chung-Feller-MacMahon proof, CFM, can be briefly described as follows. Given a multiset written in nondecreasing order consider the lists obtained by adding k to each term, reducing modulo $n + 1$ and rearranging to get a multiset listed in nondecreasing order. The content of the Chung-Feller-Macmahon Theorem is that exactly one of these will be a Catalan multiset.

This will suffice for our counting purposes even though it does not capture the full generality

of CFM. The version we have given is found, for example, in [14], however there are many different proofs of the full CFM, frequently given in terms of other combinatorial objects.

To illustrate CFM as described above we find the preimage of 002 by first adding 0, 1, 2, 3 modulo 4, respectively, to get the lists 002, 113, 220, 331. We then rearrange these in non-decreasing order to write these multisets as 002, 113, 022, 133.

When $n = 3$, in a similar manner we associate a set of 4 size 3 multisets from $\{0, 1, 2, 3\}$ with each Catalan multiset. This partitions all $\binom{4+3-1}{3} = \binom{6}{3} = 20$ such multisets into 5 preimages, each of size 4, each of which contains exactly one Catalan multiset. We write down the Catalan multisets at the top of each preimage. This is Table 2 repeated here for convenience. The example above is the third column.

000	001	002	011	012
111	112	113	122	123
222	223	022	233	023
333	033	133	003	013

3.3 CGP

We now proceed to give some examples to illustrate a multijective proof involving CG-arrangements. This appears to first be due to Rubenstein [17]. Because of notational differences, on the surface his construction of preimages seems different, making use of objects called inner and outer nests. However, a careful examination shows that the multijection we will describe is indeed the same as Rubenstein's. Our approach to finding the preimages seems to be a bit more straightforward. We do follow the approach of [17] in describing the other direction, the multijection taking arbitrary sequences to well formed sequences.

Consider an arbitrary CG-arrangement. We will use an example with $n = 5$: $)()((())($. We first identify adjacent open-closed pairs $()$. In the following there are 2 such pairs, indicated with large parentheses. Removing these we get the string on the right, which has one new open-closed pair.

$$)(\big())((\big())(\Rightarrow))()(.$$

We iteratively remove this pair, indicated as large parentheses on the left below, to get a new sequence with no open-closed pair. That is, a string of left $)$ followed by right $)$.

$$))(\big())(\Rightarrow))(.$$

Finally, we reverse, in the original string, those parentheses that remain. Those that remain and get swapped are identified as small on the left below, with the parentheses that are

paired shown as larger. We reverse the small, non paired parentheses on the right below to get a string that is well formed:

$$)\bigl(\bigr)\bigl(\bigl(\bigr)\bigr)\Rightarrow (\bigl(\bigr)\bigl(\bigl(\bigr)\bigr)\bigr)$$

This procedure converts an arbitrary arrangement of parentheses into a well formed arrangement, where each open parenthesis is later followed by its corresponding closed parenthesis. If the string is well formed initially the procedure results in an empty string and we make no changes.

We can reverse this procedure as well, by creating the collection of sequences that would result in the same single well formed original. That is, the preimages under the map. Consider the well formed arrangement of parentheses $((((())))$ from the example above. For each of the $n = 5$ paired $()$ take the pair and all other pairs containing it and swap $()$ to $)()$ to get a new string of 5 left and 5 right parentheses. The 5 new arrangements are shown below. We indicate from left to right on each line, using larger sizes to indicate the selected pair first; then the selected pair and pairs containing it; and finally the result after the swap. Note that here the reversed pairs are now large and those that are paired are small.

$$\begin{aligned} & \bigl(\bigr)((()))) \Rightarrow \bigl(\bigr)((()))) \Rightarrow \bigr)\bigl(\bigr)((()))) \\ & ()\bigl(\bigr)((()))) \Rightarrow ()\bigl(\bigr)((()))) \Rightarrow ()\bigr)\bigl(\bigr)((())))\bigl(\\ & ()\bigl(\bigr)((()))) \Rightarrow ()\bigl(\bigr)((()))) \Rightarrow ()\bigr)\bigr)\bigl(\bigr)((())))\bigl(\\ & ()\bigl(\bigr)((()))) \Rightarrow ()\bigl(\bigr)((()))) \Rightarrow ()\bigr)\bigr)\bigr)\bigl(\bigr)((())))\bigl(\\ & ()\bigl(\bigr)((()))) \Rightarrow ()\bigl(\bigr)((()))) \Rightarrow ()\bigr)\bigr)\bigr)\bigr)\bigl(\bigr)((())))\bigl(\end{aligned}$$

For one further example we find the preimage of $((()))$. We identify the paired $()$. For each such pair we get a new arrangement reversing $()$ to $)()$ for the pair as well as for all pairs enclosing it. In the display, the left term has as large parentheses the selected pair, the middle also makes large the pairs containing it and the right reverses all of the large parentheses.

$$\begin{aligned} & \bigl(\bigr)() \Rightarrow \bigl(\bigr)() \Rightarrow \bigr)()\bigl(\\ & \bigl(\bigr)() \Rightarrow \bigl(\bigr)() \Rightarrow \bigr)\bigr)\bigl(\bigr)() \\ & \bigl(\bigr)() \Rightarrow \bigl(\bigr)() \Rightarrow \bigl(\bigr)\bigr)\bigl(\end{aligned}$$

Along with the initial set, corresponding to picking no parenthesis, we get the following preimage: $((\))(,),\)((\),\))((\),\ ((\))(.$

When $n = 3$, in a similar manner we associate a set of 4 general CG-arrangements with each well formed CG-arrangement. This partitions all $\binom{6}{3} = 20$ CG-arrangements into 5 preimages each of size 4 each of which contains exactly one well formed CG-arrangement. We write down the well formed CG-arrangements at the top of each preimage:

$((\))$	$((\))$	$((\))$	$((\))$	$((\))$
$)((\))$	$)((\))$	$)((\))$	$)((\))$	$)((\))$
$)((\))$	$)((\))$	$)((\))$	$)((\))$	$)((\))$
$)((\))$	$)((\))$	$)((\))$	$)((\))$	$)((\))$
$)((\))$	$)((\))$	$)((\))$	$)((\))$	$)((\))$

Table 6: CGP Partition, $n = 3$

For comparison we write down the corresponding balanced binary strings using replacing (with 1 and) with 0 and then also as multisets using the 0 count bijection:

111000	110100	110010	101100	101010
011001	010101	010110	011100	011010
001011	001101	001110	100101	100110
000111	010011	110001	100011	101001

Table 7: CGP Partition $n = 3$, written as binary strings

000	001	002	011	012
113	123	122	111	112
233	223	222	023	022
333	133	003	033	013

Table 8: CGP Partition $n = 3$, written as multisets using 0 count bijection

For this example with $n = 3$ the groupings for CGP appear different from CDM and CFM under obvious bijections. We will note this more generally in Section 4.2.

We observe here that initial part of the process described for CGP can be used to describe a different map of arbitrary CG-arrangements to formed CG-arrangements. After identifying matched pairs, we have a string of right) followed by left (. Taking the cyclic shift of the original starting with the first left (produces a well formed CG-arrangement. Our first example with matched pairs larger was $)((\))((\))((\))(.$ Removing these leaves the smaller $)((\))(.$ Taking a cyclic shift starting with the first left (yields the well formed CG-arrangement $((\))((\))((\))(.$ However, it is easy to check that this map is not uniform. The preimages do not form a uniform partition so this cannot be directly used for a proof of the Catalan number formula. This provides a hint as to why the cycle lemma process makes use of an added leading 1 rather than just finding a cyclic shift of the initial sequence.

3.4 Comparing the Partitions

As we have already observed, at least for the $n = 3$ case, CDM and CFM produce different partitions if σ records the number of 0's before each 1 and produce the same partition if σ records the number of 1's before each 0. In the second case, π is not σ , ballot lists do not map to Catalan multisets. We could ask what π is to make the diagram commute. The implied version of π is: given a ballot list, find the balanced binary string in the preimage which has a complement that is a ballot list. This is not particularly insightful. We do not see any 'nice' shorter description of this π .

In our examples for $n = 3$, CGP appears to be different from CDM and CFM under obvious choices for the bijections. We are not aware of any simple direct connection between the CGP partitions and either CDM or CFM partitions. Perhaps the only way to describe such a sigma is via the steps in the diagram. However, it would be interesting if there was a 'nicer' description of such a σ . That is, find a 'nice' description of a σ that makes CGP and CDM (or CFM) translations with respect to σ . The case $n = 3$ being between Tables 7 and 1 or between Tables 8 and 2.

4 Proof Sketches

We briefly sketch the idea for proofs of the observations we have been making and note an example where our approach leads us to a different perspective on a proof.

4.1 Connection between CFM and CDM

As observed in the examples when $n = 3$ it seems that CFM and CDM are translations when σ maps balanced binary strings to multisets by recording the number of 1's preceding each 0. We called this the 1 count bijection.

Note that while this bijection preserves the partition, it does not map ballot lists to Catalan multisets. That is, π to make the diagram commute is not σ even though σ could be applied to the Catalan sets. For example, consider the third part 110010, 100101, 001011, 011100 from the CDM partition in Table 1. Under σ the corresponding multisets are 223, 112, 001, 330, with 330 written as 033. This is the fourth part in the CFM partition of Table 2. However, the image of the ballot list 110010 is 223 which is not a Catalan multiset. Instead π is as described above in Section 3.4.

In general, as described in Section 3.1 the preimage of a balanced binary string is obtained by appending a lead 1, taking all cyclic shifts starting with a 1 and for each shift deleting the lead 1. So, if a_1 and a_2 are in the preimage of the same ballot list a we can write a , with added 1 at the start as $1x1y1z$, where x, y, z may be empty strings and write a_1 and a_2 as

$y1z1x$ and $z1x1y$. Let $\#x, \#y, \#z$ be the number of 1's in the corresponding substrings. For a 0 in substring z or x the number of preceeding 1's in a_2 is $\#y + 1$ more than in a_1 . For a 0 in substring y the number of preceeding 1's in a_2 is $2 + \#x + \#z$ less than in a_2 . Since, modulo $n + 1$, $-(2 - \#x - \#z) \equiv \#y + 1$ in multisets we obtain $\sigma'(a_2)$ from $\sigma'(a_1)$ by adding $\#y + 1$ modulo $n + 1$. This is exactly the process for CFM described in Section 3.2. When one of a_1 or a_2 is the ballot list, the analysis is similar.

4.2 CGP different than CFM and CDM

We noted in the case $n = 3$ the apparent difference between CGP and CDM and CFM in Tables 7 and 8. For completeness we briefly note that this difference holds in general under the bijection from general CGP-arrangments to CDM, balanced binary strings by replacing left (with 1 and right) with 0 and then using the 0 count bijection to multisets. A similar analysis for the 1 count bijection is straightforward. Since we have seen that CDM and CFM are translations under the right perspective we get the difference of CGP with both CDM and CFM. We are not aware of a 'nice' description for a bijection σ that makes CGP a translation of either CFM or CDM. The partitions appear to be different under natural bijections.

For general n the well formed arrangement consisting of n left (followed by n right) corresponds to the binary string with n ones followed by n zeroes and the multiset of all 0's. The preimage of this under CDM consists of balanced binary strings where all of the 0's are consecutive. As multisets these are $00 \dots 0nn \dots n$. That is $a_1 = a_2 = \dots a_k = 0$ and $a_{k+1} = \dots = a_n = n$ for $k = 0, 1, 2, \dots, n$ with all n 's when $k = 0$ and all 0's when $k = n$.

Under CGP the preimages of $((\dots))$ swap, for $k = 0, 1, 2, \dots, n$ the first k left (for right) and the last k right) for left (. As a binary string we have $0 \dots 01 \dots 10 \dots 01 \dots 1$ where there are k 0's followed by $n - k$ 1's followed by $n - k$ 0's followed by k 1's. So unlike CDM, the 0's are not consecutive. As multisets these are $0, 0, \dots, 0, n - k, n - k, \dots, n - k$, with k 0's followed by $(n - k)$ $(n - k)$'s. That is $a_1 = \dots = a_k = 0$ and $a_{k+1} = \dots = a_n = n - k$ with all n 's when $k = 0$ and all 0's when $k = n$. So again we see that these these do not match the part containing $00 \dots 0$ under CDM viewed as multisets.

4.3 Translations to New Perspective

We can also use the description of translations to rephrase multijections in other terminology. As one example consider CDM and CFM, keep both σ and π as the bijection that records the number of 0's preceding each 1. We have already noted that this σ maps ballot lists to Catalan multisets so can be used for π .

We can then ask what β on multisets would be to make the diagram commute and hence

have a translation. This allows us to ‘discover’ a different multijective perspective on the cycle lemma in the language of multisets. To be clear, if written out in detail this would not be a new proof, but an old proof written with different terminology. We assume that the alternate approach described here is not new but are not aware of a first explicit reference.

For example when $n = 3$ consider Table 3, the image of Table 1 under σ . One grouping using CDM is 110010, 100101, 001011, 011100. The corresponding grouping under σ is 002, 023, 233, 111. In particular the ballot list 110010 maps to the Catalan multiset 002. So we look to describe how to generate such preimage groupings under CDM using the terminology of multisets rather than the terminology of binary strings.

We describe this briefly by example, without proof. Filling in the details is straightforward.

Given a Catalan multiset from $\{0, 1, \dots, n\}$, $0 \leq a_1 \leq a_2 \leq \dots \leq a_n \leq n$ create a new string $b_1 \leq b_2 \leq \dots \leq b_n$ by setting $b_i = a_{i+1} - a_1$ for $i = 1, 2, \dots, n-1$ and $b_n = n - a_1$. That is, we consider $a_2 \leq a_3 \leq \dots \leq a_n \leq n$ and subtract a_1 from each term, including n . This yields a new string in the equivalence class. Repeat this n times, and include the original multiset to get the preimage. Observe that this same process works for any n multiset so applying it to any n multiset and repeating $(n+1)$ times returns to the original. We get a set of $n+1$ multisets containing exactly one Catalan multiset.

For example consider 002 when $n = 3$. The first additional multiset in the preimage is $0-0, 2-0, 3-0 = 023$, the second is $2-0, 3-0, 3-0 = 233$, the third is $3-2, 3-2, 3-2 = 111$. If we were to iterate one more time we get $1-1, 1-1, 3-1 = 002$, the starting Catalan multiset. Note that this is indeed the family we obtained from 110010 using CDM translated to multisets via σ . So we have reframed, using the terminology of multisets, the process for finding preimages of balanced binary strings using CDM. This is different from the CFM preimages defined by adding some j to each term, reducing modulo $n+1$ and rearranging.

We can directly describe the string $0 \leq c_1 \leq c_2 \leq \dots \leq c_n \leq n$ obtained after j applications of the process in the previous paragraph to $0 \leq a_1 \leq a_2 \leq \dots \leq a_n \leq n$. This is $c_i = a_{i+1} - a_j$ for $i = 1, 2, \dots, n-j$ and $c_{n-j+1} = n - a_j$ and $c_{n-j+k} = n - a_j + a_k$ for $x = 2, 3, \dots, j$. To illustrate with a slightly larger example, consider $n = 5$ and the multiset 11244. We append $n = 5$ followed by a copy of the string with $n = 5$ added to each term to get 11244566788. To find, for example, the multiset obtained by applying the process $j = 3$ times we look at the terms 44566 following $a_3 = 2$ and subtract $a_3 = 2$ from each to get 22344. Note that we are simply describing a translation to multisets of the shifting that occurs in the cycle lemma.

4.4 Other CDM and CFM proofs

We note that the cycle lemma is also sometimes viewed with the appended 1 as part of the family of objects and we consider all cyclic shifts, not just those starting with a 1. That is, we consider strings with $n+1$ ones and n zeros and consider all $2n+1$ cyclic shifts, with

exactly one shift corresponding to a Catalan object. Thus we view the Catalan numbers as $C_n = \frac{1}{2n+1} \binom{2n+1}{n} = \frac{1}{n+1} \binom{2n}{n}$. We leave as an exercise the translation, similar to that in the previous section, into size $n+1$ multisets from $\{0, 1, \dots, n\}$, where the preimage families now have size $2n+1$.

Proofs of Chung-Feller-MacMahon Theorem are frequently given in terms of other objects, in particular Dyck paths and prove more than what we have described here. The broader context creates a different partition of a family of size $\binom{2n}{n}$, into $n+1$ parts each of size C_n , with exactly one of these parts consisting of the Catalan family. Each part has a common value of some statistic. For the example of CFM with strings represented as multisets $0 \leq a_1 \leq \dots \leq a_n \leq n$ with Catalan strings those for which $a_i \leq i-1$ for $i = 1, 2, \dots, n$, the exceedance is the number of violations of $a_i \leq i-1$, that is the number of $a_i \geq i$.

The Catalan lists are those with exceedance 0. The general CFM proofs show that for each $j = 0, 1, \dots, n$ the number of lists with exceedance j is the same, providing a uniform partition with Catalan lists being exactly one part. The partitions we have described earlier for CFM, containing exactly one Catalan list in each part, are an orthogonal partition in the sense that each part contains exactly one list with each value of the exceedance. There are many variations on such statistics in the context of Catalan numbers and generalizations. See [1] for a description of such a statistic on the CGP families. See, for example, [5], [12], [20] for examples of other CFM proofs and generalizations. We can ask which of these are translations and which are different under obvious choices of σ and ask what happens what can be gleaned about the statistics under the various translations.

For our purposes we note that different proofs of CFM in fact describe different preimage partitions. For example, it is not difficult to check with small examples that under natural bijections σ and π the CFM preimage families described in [12], [16], [20] are different from the one we have described. Indeed [12] explicitly creates a different preimage partition so that strings in a part differ only in two locations. It might be interesting to determine if there is a somewhat natural σ under which these different proofs are translations of each other. It might also be interesting to describe the translation under natural bijections to other objects, as we did in Section 4.3 for multisets and the version of CFM used here above.

5 Conclusion

Our goal here was to examine different versions of what we call multijjective proofs for Catalan numbers. The framework sets up a large number of possible exercises: describing how known multijjective proofs translate to other families of Catalan objects, describing bijections that make two seemingly different multijjective proofs into translations, and finding new multijjective proofs that are different under natural choices for the bijections.

References

- [1] Callan, D., Two Uniformly Distributed Parameters Defining Catalan Numbers, *Amer. Math. Monthly* **106** (1999), 948–949.
- [2] Chen, W.Y.C, Nelson, Y.L., and Shapiro, L.W., The Butterfly Decomposition of Plane Trees, *Discrete Appl. Math.*, **155** (2007), 2187–2201.
- [3] Dershowitz, N., and Zaks, S., The Cycle Lemma and Some Applications, *European J. Combin.*, **1** (1990), 35–40.
- [4] Feldman, D., Counting Plane Trees, unpublished manuscript, 1992. Linked under A006082 and A006936 in On-Line Encyclopedia of Integer Sequences.
- [5] Huq, A., Generalized Chung-Feller Theorems for Lattice Paths, Ph.D. Dissertation, Brandeis Univ. 2009.
- [6] Graham, R., Knuth, D. and Patashnik, O., *Concrete Mathematics*, Addison-Wesley, 1989.
- [7] Guy, R.K., The second strong law of small numbers, *Math. Mag.*, **63** (1990), 3–20.
- [8] Koshy, T., *Catalan Numbers with Applications*, Oxford Univ. Press, 2009.
- [9] Krattenthaler, C., Lattice Path Enumeration, Chapter 10 in *Handbook of Enumerative Combinatorics*, CRC press, 2015. 589–678.
- [10] Krusemeyer, M., A Parenthetical Note (to a Paper of Guy), *Math. Mag.*, **69** (1996), 257–260.
- [11] Ma, J., Shen, H., and Yeh, Y., Rooted Cyclic Permutations of Lattice Paths and Uniform partitions, *Disc. Math.*, **338** (2015), 1111–1125.
- [12] Mütze, T., Standke, C. and Wiechert, V., A Minimum-Change Version of the Chung-Feller Theorem for Dyck Paths, *European J. Combin.*, **69** (2018), 260–275.
- [13] Mohanty, S.G., *Lattice Path Counting and Applications*, Academic press, 1979.
- [14] Narayana, *Lattice Path Combinatorics, with Statistical Applications*, Univ. Toronto press, 1979.
- [15] Penrice, S., Stacks, Bracketings, and CG-Arrangements, *Math. Mag.*, **72** (1999), 321–324.
- [16] Renault, M., Four proofs of the Ballot Theorem, *Math. Mag.*, **80** (2007), 345–352.

- [17] Rubenstein, D., Catalan Numbers revisited, J. Combin. Theory Ser. A, **68** (1994), 486–490.
- [18] Snevily, H.S. and West, D.B., The Bricklayer Problem and the Strong Cycle Lemma, Amer. Math. Monthly **105** (1998), 131–143.
- [19] Stanley, R.C., Catalan Numbers, Cambridge Univ. Press, 2015.
- [20] Woan, W., Uniform Partitions of Lattice Paths and Chung-Feller Generalizations, Amer. Math. Monthly **108** (2001), 556–559.