# A BIJECTION OF DYCK PATHS AND MULTISETS OF BALANCED BINARY LYNDON WORDS

RICHARD J. MATHAR

Abstract. This is a scholarly explanation of how a multiset of Balanced Binary Lyndon Words is a standard factorization of one unique Balanced Binary Word which has a bijection to a Dyck Path of the same semilength.

## 1. Notations

1.1. **Lyndon Words.** A Lyndon $n$-word is a word with $n$ letters which is the lexicographic unique smallest of all words that could be deduced by the $n$ cyclic permutations (=rotations) of the letters. See for example [5] for an algorithm to generate these words given a length $n$.

**Definition 1.** *A* binary *Lyndon word is a word with an alphabet of 2 letters, which we will take from the set* $\{0, 1\}$.

**Definition 2.** *A* balanced *Lyndon word is a word where each letter occurs the same number of times in the word.*

**Definition 3.** *(BBL) A balanced binary Lyndon word (BBL) is a Lyndon word which is both binary and balanced.*

The binary Lyndon words are enumerated in sequence [7, A001037] of the Online Encyclopedia of Integer Sequences (OEIS). The binary balanced Lyndon words of semilength $n$ are the sequence [7, A022553]:
(1)

$$L(n) = \frac{1}{2n}\sum_{d|n}\mu(n/d)\binom{2d}{d} = 1, 1, 1, 3, 8, 25, 75, 245, 800, 2700, 9225, \quad (n \geq 0)$$

Table 1 shows the number of Binary Lyndon Words $L_e(n) = L_{-e}(n)$ with an excess of $e$, defined as the number of 0's minus the number of 1's. $L(n) = L_e(2n)$ is the sequence of entries in column $e = 0$ in that table.

**Remark 1.** *The non-empty balanced binary Lyndon words start with the letter 0. [Proof by contradiction: if the first letter were 1 one could shift the word until the first letter were a zero (known to exist) which would become lexicographically smaller.]*

**Remark 2.** *The non-empty balanced binary Lyndon words end with the letter 1. [Proof by contradiction: if the last letter were 0 one could shift the word cyclically once to the right with a new prefix 0 which would become lexicographically smaller.]*

---

*Date*: October 30, 2021.
1991 *Mathematics Subject Classification.* Primary 68R15; Secondary 06F05.

| $n\backslash e$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | | | | | | | | | | | | | |
| 1 | 0 | 1 | | | | | | | | | | | | |
| 2 | 1 | 0 | 0 | | | | | | | | | | | |
| 3 | 0 | 1 | 0 | 0 | | | | | | | | | | |
| 4 | 1 | 0 | 1 | 0 | 0 | | | | | | | | | |
| 5 | 0 | 2 | 0 | 1 | 0 | 0 | | | | | | | | |
| 6 | 3 | 0 | 2 | 0 | 1 | 0 | 0 | | | | | | | |
| 7 | 0 | 5 | 0 | 3 | 0 | 1 | 0 | 0 | | | | | | |
| 8 | 8 | 0 | 7 | 0 | 3 | 0 | 1 | 0 | 0 | | | | | |
| 9 | 0 | 14 | 0 | 9 | 0 | 4 | 0 | 1 | 0 | 0 | | | | |
| 10 | 25 | 0 | 20 | 0 | 12 | 0 | 4 | 0 | 1 | 0 | 0 | | | |
| 11 | 0 | 42 | 0 | 30 | 0 | 15 | 0 | 5 | 0 | 1 | 0 | 0 | | |
| 12 | 75 | 0 | 66 | 0 | 40 | 0 | 18 | 0 | 5 | 0 | 1 | 0 | 0 | |
| 13 | 0 | 132 | 0 | 99 | 0 | 55 | 0 | 22 | 0 | 6 | 0 | 1 | 0 | 0 |
| 14 | 245 | 0 | 212 | 0 | 143 | 0 | 70 | 0 | 26 | 0 | 6 | 0 | 1 | 0 |
| 15 | 0 | 429 | 0 | 333 | 0 | 200 | 0 | 91 | 0 | 30 | 0 | 7 | 0 | 1 |
| 16 | 800 | 0 | 715 | 0 | 497 | 0 | 273 | 0 | 112 | 0 | 35 | 0 | 7 | 0 |
| 17 | 0 | 1430 | 0 | 1144 | 0 | 728 | 0 | 364 | 0 | 140 | 0 | 40 | 0 | 8 |

TABLE 1. The number of Binary Lyndon Words $L_e(n)$ of length $n$ with imbalance $e$, i.e., which have $e$ more 0's than 1's [7, A051168,A092964,A185700].

1.2. **Dyck Paths.** A Dyck Path is a path on a square lattice with steps which are either u(p)-steps, (1,1)-steps, or d(own)-steps (1,-1), which start at the origin, and at the horizontal line, and never step below the horizontal line. The requirement of returning to the horizontal line implies that Dyck Paths exist only for even an even number of steps. The Dyck Paths of semilenght $n$ are counted by the Catalan numbers [7, A000108],

(2)

$$C(n) = \frac{1}{n+1}\binom{2n}{n} = 1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, \dots (n \geq 0)$$

**Remark 3.** *The Catalan numbers count the Binary Lyndon Words with imbalance 1: $C(n) = L_1(2n+1)$, column $e = 1$ in Table 1.*

## 2. BBL TO DYCK MAP

**Definition 4.** *(Natural Map) The Natural Map of a Lyndon Word to a lattice path is the translation $0 \to u$, $1 \to d$ in the alphabet. The Natural Map of Dyck Paths to words is the translation $u \to 0$, $d \to 1$.*

A BBL defines a Path under the Natural Map that returns to the horizontal line if started at the origin, because the *balanced* restriction of the words ensures that the last step returns to the horizontal line because the number of u-steps equals the number of d-steps.

We use a notation for BBL words which is

(1) a parenthetical representation in the letters 0 and 1 as a binary number,

| $n$ | |
|---|---|
| 1 | udN1 |
| 2 | uuddN3 ududN5r2 |
| 3 | uuudddN7 uududdN11 uuddudN13r2 uduuddN19r2 udududN21r3 |
| 4 | uuuuddddN15 uuududddN23 uuudduddN27 uuudddudN29r2 uuduudddN39 uudududdN43 uududdudN45r2 uudduuddN51r2 uuddududN53r3 uduuudddN71r2 uduududdN75r2 uduuddudN77r3 ududuuddN83r3 ududududN85r4 |
| 5 | uuuuudddddN31 uuuududdddN47 uuuuddudddN55 uuuuddduddN59 uuuuddddudN61r2 uuuduuddddN79 uuudududddN87 uuududduddN91 uuududddudN93r2 uuudduudddN103 uuuddududdN107 uuudduddudN109r2 uuuddduuddN115r2 uuudddududN117r3 uuduuuddddN143 uuduududddN151 uuduudduddN155 uuduudddudN157r2 uududuudddN167 uududududdN171 uudududdudN173r2 uududduuddN179r2 uududdududN181r3 uudduuudddN199r2 uudduududdN203r2 uudduuddudN205r3 uudduduuddN211r3 uuddudududN213r4 uduuuuddddN271r2 uduuududddN279r2 uduuudduddN283r2 uduuudddudN285r3 uduuduudddN295r2 uduudududdN299r2 uduududdudN301r3 uduudduuddN307r3 uduuddududN309r4 ududuuudddN327r3 ududuududdN331r3 ududuuddudN333r4 udududuuddN339r4 udududududN341r5 |

Table 2. All Dyck Paths of semilength $n \leq 5$.

(2) a capital `N` followed by a representation of the same binary number as an integer in base 10,
(3) if the Natural Map defines a Dyck Path with more than one return a lower-case `r` and the number of returns.
(4) If the Natural Map creates a path that dives below the horizontal line, a `r!` and the number of returns (including the returns from below).

If there is no `r`, the BBL matches a Dyck Path with a single (final) return.

**Remark 4.** *If the Lyndon Word does not represent a Dyck Path under the Natural Map, one can always find a word in the same equivalence class by using as many cyclic shifts as necessary to move the minimum point on the path to the origin. If there is more than one point with the same (negative) altitude, this number of shifts is not defined uniquely.*

*Vice versa, if a Dyck Path with one return is not a Lyndon Word under the Natural Map, one can define its unique associate BBL by cycling the word until the minimum word is found.*

The complete set of BBL up to semilength 5 is Table 3.

**Remark 5.** *The decimal representation would suffice to characterize a BBL, which means there are no two BBL which have the same decimal representation: the number of bits set in the integer (the Hamming weight) is the semilength, and for a given semilength no two BBL are the same (i.e., the number of leading zeros is unique).*

Statistics that is commonly applied to the Dyck Paths can be considered for Lyndon Words as well. We can refine the BBL by the number of returns to the

| $n$ | |
|---|---|
| 0 | empty word |
| 1 | (01)N1 |
| 2 | (0011)N3 |
| 3 | (000111)N7 (001011)N11 (001101)N13r2 |
| 4 | (00001111)N15 (00010111)N23 (00011011)N27 (00011101)N29r2 (00100111)N39 (00101011)N43 (00101101)N45r2 (00110101)N53r3 |
| 5 | (0000011111)N31 (0000101111)N47 (0000110111)N55 (0000111011)N59 (0000111101)N61r2 (0001001111)N79 (0001010111)N87 (0001011011)N91 (0001011101)N93r2 (0001100111)N103 (0001101011)N107 (0001101101)N109r2 (0001110011)N115r2 (0001110101)N117r3 (0001111001)N121r!3 (0010010111)N151 (0010011011)N155 (0010011101)N157r2 (0010100111)N167 (0010101011)N171 (0010101101)N173r2 (0010110011)N179r2 (0010110101)N181r3 (0011001101)N205r3 (0011010101)N213r4 |

TABLE 3. All BBL up to semilength $n \le 5$.

| $n\backslash r$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | | | | | | | |
| 2 | 1 | 1 | | | | | | |
| 3 | 2 | 2 | 1 | | | | | |
| 4 | 5 | 5 | 3 | 1 | | | | |
| 5 | 14 | 14 | 9 | 4 | 1 | | | |
| 6 | 42 | 42 | 28 | 14 | 5 | 1 | | |
| 7 | 132 | 132 | 90 | 48 | 20 | 6 | 1 | |
| 8 | 429 | 429 | 297 | 165 | 75 | 27 | 7 | 1 |

TABLE 4. The number of Dyck Paths of semilength $n$ which have $r$ returns to the horizontal line [7, A033184]. Row sums are the Catalan numbers (2).

vertical line as in Table 5, for example. For the paths that dive below the horizontal line the returns include returns from below. The last 1 in each row counts the words of semilength $n$ with $n-1$ returns which start `0011` and are repeatedly followed by 01.

**Conjecture 1.** *The number of BBL of semilength $n$ with $n-2$ returns in Table 5 is $2^{n-4} + [2n-3-(-1)^n]/4$ for $n \ge 4$.*

## 3. MULTISET COMPOSITION

3.1. **Bijections.** The Multiset Transform of the particular family of Binary Balanced Lyndon Words defined above is shown in Table 6, defined as usual [6]: If

$$L(x) = \sum_{n\ge 0} L(n)x^n; \quad C(x) = \sum_{n\ge 0} C(n)x^n = \frac{1-\sqrt{1-4x}}{2x}, \tag{3}$$

| $n\backslash r$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | | | | | | | | | | | | |
| 2 | 1 | 0 | | | | | | | | | | | |
| 3 | 2 | 1 | 0 | | | | | | | | | | |
| 4 | 5 | 2 | 1 | 0 | | | | | | | | | |
| 5 | 13 | 7 | 4 | 1 | 0 | | | | | | | | |
| 6 | 37 | 19 | 12 | 6 | 1 | 0 | | | | | | | |
| 7 | 107 | 60 | 41 | 25 | 11 | 1 | 0 | | | | | | |
| 8 | 325 | 181 | 133 | 90 | 51 | 19 | 1 | 0 | | | | | |
| 9 | 1009 | 576 | 439 | 324 | 210 | 105 | 36 | 1 | 0 | | | | |
| 10 | 3219 | 1839 | 1460 | 1137 | 801 | 480 | 220 | 68 | 1 | 0 | | | |
| 11 | 10465 | 6025 | 4913 | 3993 | 2996 | 1978 | 1098 | 463 | 133 | 1 | 0 | | |
| 12 | 34651 | 19966 | 16683 | 14001 | 10991 | 7791 | 4826 | 2483 | 978 | 261 | 1 | 0 | |
| 13 | 116376 | 67236 | 57255 | 49302 | 40097 | 29919 | 20027 | 11627 | 5595 | 2070 | 518 | 1 | 0 |
| 14 | 395944 | 228912 | 198211 | 174250 | 145726 | 113173 | 80138 | 50651 | 27688 | 12511 | 4378 | 1030 | 1 |

TABLE 5. The number of BBL of semilength $n$ which have $r$ returns to the horizontal line (under the Natural Map). Row sums are the Lyndon words (1).

| $n\backslash c$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 6 | 8 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | | | | | | | | | | | |
| 1 | 0 | 1 | | | | | | | | | | |
| 2 | 0 | 1 | 1 | | | | | | | | | |
| 3 | 0 | 3 | 1 | 1 | | | | | | | | |
| 4 | 0 | 8 | 4 | 1 | 1 | | | | | | | |
| 5 | 0 | 25 | 11 | 4 | 1 | 1 | | | | | | |
| 6 | 0 | 75 | 39 | 12 | 4 | 1 | 1 | | | | | |
| 7 | 0 | 245 | 124 | 42 | 12 | 4 | 1 | 1 | | | | |
| 8 | 0 | 800 | 431 | 138 | 43 | 12 | 4 | 1 | 1 | | | |
| 9 | 0 | 2700 | 1470 | 490 | 141 | 43 | 12 | 4 | 1 | 1 | | |
| 10 | 0 | 9225 | 5160 | 1704 | 504 | 142 | 43 | 12 | 4 | 1 | 1 | |
| 11 | 0 | 32065 | 18160 | 6088 | 1763 | 507 | 142 | 43 | 12 | 4 | 1 | 1 |

TABLE 6. Multiset transformation of the Binary Balanced Lyndon words of semilength $n$ [7, A289978]. The column $c = 1$ is given by (1). The row sums are $C(n) = \sum_{c\geq 0} D_c(n)$

.

are the generating functions of the BBL and Catalan sequences, then [1]

$$C(x) = \exp\left(\sum_{k\geq 1} \frac{L(x^k)}{k}\right) = \prod_{k\geq 1} \frac{1}{(1-x^k)^{L(k)}} \tag{4}$$

are a Euler Transform pair. Table 6 displays in row $n$ and column $c$ the number [6, Theor. I.1][8, (4.2.3)]

$$D_c(n) = \sum_{\substack{n=n_1+2n_2+\cdots+cn_c\\ n_i\ge 0}} \prod_{i=1}^{c} \binom{D_1(i)+n_i-1}{n_i}; \quad D_1(n)=L_n) \tag{5}$$

of multisets "containing" $c$ non-empty BBL words of arbitray indivdual length with a total of $n$ letters. The formula is a sum over the partitions of $n$ proposing to take $n_i$ words of length $i$; the product over $i$ indicates that the subsets of words of fixed length are independent, and the binomial factor is the number of multisets of size $n_i$ if the underlying set has $D_1(i)$ elements of size $i$ [2, 1.3.2.2].

The theme of this paper is to give a combinatorial interpretation of that relation of BBL multisets to Dyck Paths. What does "contain" actually mean?

The basic observation in this context is that a Multi*set* of Lyndon words of a fixed alphabet has a well-known property which translates the set into an ordered list: the standard lexicographic order relation of the words [5] defines one unique sequence of non-increasing words if larger words are placed to the left of smaller words; if there are repetitions, these words appear as neighbours in the sequence, creating "inversions" [9].

We shall use the symbol $\mathtt{o}$ to divide lists of words and to denote the (non-commutative) product/concatenation of words.

$D_c(n)$ is the number of BBL Multisets which have $c$ elements in their ordered list of elements. The elements left are a standard factorization of a binary balanced word of length $n$.

On the other hand, each Dyck Path can be uniquely split in $r$ unordered smaller Dyck Paths at the $r$ returns to the horizontal line. Each of the subpaths is a binary balanced word under the natural map, and—having no cycles due to the split—has a unique associate BBL according to Remark 4. A pair of two adjacent BBL in that unordered list can be concatenated (multiplied) and becomes a single, longer word if the left word is strictly smaller than the right word [4]. This process of multiplying BBL can be iterated; the merger is stopped by the presence of inversions. The number of BBL in the remaining list is the column $c$ in the table of the $D_c(n)$ which a Dyck Path of semilength $n$ is assigned to.

**Example 1.** *The Dyck Path uuddN3→(0011)N3 is already split and fully reduced to a single Lyndon word; it contributes to* $D_1(2)$.

*The Dyck PathududN3 is split to (ud)(ud)→* $(01)o(01)$ *and is already fully reduced (because the two factors are lexicographically equal). Two factors are left, and the Dyck Path contributes to* $D_2(2)$.

**Example 2.** *The Dyck Path uuudddN7→(000111)N7 is already split and fully reduced to a single Lyndon word; it contributes to* $D_1(3)$.

*The Dyck Path uududdN11→(001011)N11 is already split and fully reduced to a single Lyndon word; it contributes to* $D_1(3)$.

*The Dyck Path uuddudN13r2 is split to (uudd)(ud)→* $(0011)o(01)$ *and reduced to (001101)N13r2 as the left factor (0011) is smaller than the right factor (01). This leaves a single word and the Dyck path contributes to* $D_1(3)$.

*The Dyck Path uduuddN13r2 is split to (ud)(uudd)→* $(01)o(0011)$ *and irreducible because the left factor (01) is larger than the right factor (0011). This leaves a product of two words and the Dyck path contributes to* $D_2(3)$.

*The Dyck path ududud*N21r3 *is split to (ud)(ud)(ud)*$\rightarrow$ (01)$o$(01)$o$(01) *which is irreducible because all factors are equal. This leaves a product of three words and the Dyck path contributes to* $D_3(3)$.

The subsequent sections are illustrating the concept for $4 \leq n \leq 6$. For row number $n$ in Table 2 where the same representation (integer number after `N`) appears in row $n$ of Table 3, there is a direct map of the Dyck path to a single BBL and the Dyck path contributes to $D_1(n)$; only the remaining, nontrivial cases under the Natural Map are discussed below.

3.2. **n=4.** For semilength $n = 4$ Dyck paths that have no Lyndon Path are represented by the products

```
(00110011)N51r2 = (0011)N3 o (0011)N3
(01000111)N71r2 = (01)N1 o (000111)N7
(01001011)N75r2 = (01)N1 o (001011)N11
(01001101)N77r3 = (01)N1 o (001101)N13r2
(01010011)N83r3 = (01)N1 o (01)N1 o (0011)N3
(01010101)N85r4 = (01)N1 o (01)N1 o (01)N1 o (01)N1
```

Compatible with row $n = 4$ in Table 6, there are 4 Dyck Paths which need 2 Lyndon words, 1 which needs 3, and 1 which needs 4.

3.3. **n=5.** While mapping the Lyndon word N121r!3 one must map it following Remark 4 to get (0010001111)N143, one of the conjugates in the class, to match the Dyck Path N143. So N143 in row $n = 5$ of Table 2 contributes to $D_1(5)$.

**Remark 6.** *This also explains why the number of BBL with 1 return and* $n = 5$ *in table 5 is one less than the number of Dyck Paths with 1 return and* $n = 5$ *in table 4.*

The further interpretation of $D_{2\ldots5}(5)$ in Table 6 is: 11 Dyck paths are reduced to products of 2 BBL

```
(0011000111)N199r2 = (0011)N3 o (000111)N7
(0011001011)N203r2 = (0011)N3 o (001011)N11
(0011010011)N211r3 = (001101)N13r2 o (0011)N3
(0100001111)N271r2 = (01)N1 o (00001111)N15
(0100010111)N279r2 = (01)N1 o (00010111)N23
(0100011011)N283r2 = (01)N1 o (00011011)N27
(0100011101)N285r3 = (01)N1 o (00011101)N29r2
(0100100111)N295r2 = (01)N1 o (00100111)N39
(0100101011)N299r2 = (01)N1 o (00101011)N43
(0100101101)N301r3 = (01)N1 o (00101101)N45r2
(0100110101)N309r4 = (01)N1 o (00110101)N53r3
```

4 Dyck paths are reduced to products of 3 BBL

```
(0100110011)N307r3 = (01)N1 o (0011)N3 o (0011)N3
(0101000111)N327r3 = (01)N1 o (01)N1 o (000111)N7
(0101001011)N331r3 = (01)N1 o (01)N1 o (001011)N11
(0101001101)N333r4 = (01)N1 o (01)N1 o (001101)N13r2
```

1 Dyck path is reduced to a product of 4 BBL

```
(0101010011)N339r4 = (01)N1 o (01)N1 o (01)N1 o (0011)N3
```

1 Dyck path is reduced to a product of 5 BBL

```
(0101010101)N341r5 = (01)N1 o (01)N1 o (01)N1 o (01)N1 o (01)N1
```

3.4. **n=6.** Continuing with row $n = 6$ in Table 6, the following 5 BBL are mapped to Dyck paths with a single return:

```
(000011111001)N249r!3 = (001000011111)N543
(000101111001)N377r!3 = (001000101111)N559
(000110111001)N441r!3 = (001000110111)N567
(000111011001)N473r!4 = (001000111011)N571
(000111100101)N485r!4 = (001010001111)N655
```

These are the 5 Dyck Paths counted in the 42 in row $n = 6$ in Table 4 and are missing in the 37 in row $n = 6$ in Table 5. One BBL is mapped to a Dyck path with 2 returns:

```
(000111101001)N489r!4 = (010010001111)N1167r2
```

The 39 Dyck Paths of semilength 6 which need 2 BBL are

```
(000111000111)N455r2 = (000111)N7 o (000111)N7
(001011000111)N711r2 = (001011)N11 o (000111)N7
(001011001011)N715r2 = (001011)N11 o (001011)N11
(001100001111)N783r2 = (0011)N3 o (00001111)N15
(001100010111)N791r2 = (0011)N3 o (00010111)N23
(001100011011)N795r2 = (0011)N3 o (00011011)N27
(001100011101)N797r3 = (0011)N3 o (00011101)N29r2
(001100100111)N807r2 = (0011)N3 o (00100111)N39
(001100101011)N811r2 = (0011)N3 o (00101011)N43
(001100101101)N813r3 = (0011)N3 o (00101101)N45r2
(001101000111)N839r3 = (001101)N13r2 o (000111)N7
(001101001011)N843r3 = (001101)N13r2 o (001011)N11
(001101001101)N845r4 = (001101)N13r2 o (001101)N13r2
(001101010011)N851r4 = (00110101)N53r3 o (0011)N3
(010000011111)N1055r2 = (01)N1 o (0000011111)N31
(010000101111)N1071r2 = (01)N1 o (0000101111)N47
(010000110111)N1079r2 = (01)N1 o (0000110111)N55
(010000111011)N1083r2 = (01)N1 o (0000111011)N59
(010000111101)N1085r3 = (01)N1 o (0000111101)N61r2
(010001001111)N1103r2 = (01)N1 o (0001001111)N79
(010001010111)N1111r2 = (01)N1 o (0001010111)N87
(010001011011)N1115r2 = (01)N1 o (0001011011)N91
(010001011101)N1117r3 = (01)N1 o (0001011101)N93r2
(010001100111)N1127r2 = (01)N1 o (0001100111)N103
(010001101011)N1131r2 = (01)N1 o (0001101011)N107
(010001101101)N1133r3 = (01)N1 o (0001101101)N109r2
(010001110011)N1139r3 = (01)N1 o (0001110011)N115r2
(010001110101)N1141r4 = (01)N1 o (0001110101)N117r3
(010001111001)N1145r!4 = (01)N1 o (0001111001)N121r!3
(010010010111)N1175r2 = (01)N1 o (0010010111)N151
(010010011011)N1179r2 = (01)N1 o (0010011011)N155
(010010011101)N1181r3 = (01)N1 o (0010011101)N157r2
(010010100111)N1191r2 = (01)N1 o (0010100111)N167
(010010101011)N1195r2 = (01)N1 o (0010101011)N171
```

```
(010010101101)N1197r3 = (01)N1 o (0010101101)N173r2
(010010110011)N1203r3 = (01)N1 o (0010110011)N179r2
(010010110101)N1205r4 = (01)N1 o (0010110101)N181r3
(010011001101)N1229r4 = (01)N1 o (0011001101)N205r3
(010011010101)N1237r5 = (01)N1 o (0011010101)N213r4
```

The 12 Dyck Paths of semilength 6 which need 3 BBL are

```
(001100110011)N819r3 = (0011)N3 o (0011)N3 o (0011)N3
(010011000111)N1223r3 = (01)N1 o (0011)N3 o (000111)N7
(010011001011)N1227r3 = (01)N1 o (0011)N3 o (001011)N11
(010011010011)N1235r4 = (01)N1 o (001101)N13r2 o (0011)N3
(010100001111)N1295r3 = (01)N1 o (01)N1 o (00001111)N15
(010100010111)N1303r3 = (01)N1 o (01)N1 o (00010111)N23
(010100011011)N1307r3 = (01)N1 o (01)N1 o (00011011)N27
(010100011101)N1309r4 = (01)N1 o (01)N1 o (00011101)N29r2
(010100100111)N1319r3 = (01)N1 o (01)N1 o (00100111)N39
(010100101011)N1323r3 = (01)N1 o (01)N1 o (00101011)N43
(010100101101)N1325r4 = (01)N1 o (01)N1 o (00101101)N45r2
(010100110101)N1333r5 = (01)N1 o (01)N1 o (00110101)N53r3
```

The 4 Dyck Paths of semilength 6 which need 4 BBL are

```
(010100110011)N1331r4 = (01)N1 o (01)N1 o (0011)N3 o (0011)N3
(010101000111)N1351r4 = (01)N1 o (01)N1 o (01)N1 o (000111)N7
(010101001011)N1355r4 = (01)N1 o (01)N1 o (01)N1 o (001011)N11
(010101001101)N1357r5 = (01)N1 o (01)N1 o (01)N1 o (001101)N13r2
```

The 1 Dyck Path of semilength 6 which needs 5 BBL is

```
(010101010011)N1363r5 = (01)N1 o (01)N1 o (01)N1 o (01)N1 o (0011)N3
```

The 1 Dyck Path of semilength 6 which needs 6 BBL is

```
(010101010101)N1365r6 = (01)N1 o (01)N1 o (01)N1 o (01)N1 o (01)N1 o (01)N1
```

## 4. Summary

Mapping Dyck Paths to Binary Balanced Lyndon words with the prescription $\{u,d\} \leftrightarrow \{0,1\}$ (involving cyclic shifts where necessary) also maps multisets of Binary Balanced Lyndon words to Dyck Paths: Dyck Paths are sorted into the triangle of the multiset transformation according to the number of factors in their standard factorization; multisets are sorted according to the number of words (counted with multiplicity).

## References


1. Mira Bernstein and Neil J. A. Sloane, *Some canonical sequences of integers*, Lin. Alg. Applic. **226–228** (1995), 57–72, (E:) [3]. MR 1344554
2. Miklós Bóna, *Handbook of enumerative combinatorics*, CRC Press.
3. Richard A. Brualdi, *From the editor-in-chief*, Lin. Alg. Applic. **320** (2000), no. 1–3, 209–216. MR 1796542
4. L. J. Cummings and M. E. Mays, *Shuffled lyndon words*, Ars. Combin. **33** (1992), 47–56. MR 1174829
5. Jean Pierre Duval, *Génération d'une section des classes de conjugaison et arbre des mots de lyndon de longueuer bornée*, Theor. Comp. Sci. **60** (1988), no. 3, 255–283. MR 0979464
6. Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics*, Cambridge University Press, 2009. MR 2483235

7. O. E. I. S. Foundation Inc., *The On-Line Encyclopedia Of Integer Sequences*, (2021), https://oeis.org/. MR 3822822
8. Frank Harary and Edgar M. Palmer, *Graphical enumeration*, Academic Press, New York, London, 1973. MR 0357214
9. Guy Melançon and Christophe Reutenauer, *Lyndon words, free algebras and shuffles*, Canad. J. Math. **41** (1989), no. 4, 577–591. MR 1012617

*URL*: `https://www.mpia.de/~mathar`

Max-Planck Institute of Astronomy, Königstuhl 17, 69117 Heidelberg, Germany