ENUMERATION OF SMIRNOV WORDS OVER THREE-LETTER AND FOUR-LETTER ALPHABETS

ANDREY O. MATVEEV

1. Smirnov Words

Smirnov words are defined to be words without consecutive equal letters. The importance of these words can easily be explained [1, p. 205]:

> Start from a Smirnov word and substitute for any letter a_j that appears in it an arbitrary nonempty sequence of letters a_j. When this operation is done at all places of a Smirnov word, it gives rise to an unconstrained word. Conversely, any word can be associated to a unique Smirnov word by collapsing into single letters maximal groups of contiguous equal letters.

2. Ternary Smirnov Words

Remark 1 [3, Rem. A.1]: *In the theoretical framework of the breakthrough article* [2]*, the numbers* $\mathfrak{T}(\theta,\mathfrak{s};k,i,j)$ *of distinct ternary Smirnov words that start with the letter* θ*, end with a letter* $\mathfrak{s}\in\{\theta,\beta\}$*, and contain* k *letters* θ*,* i *letters* α*, and* j *letters* β*, can be calculated in the following way:*

(i)

$$\mathfrak{T}(\theta,\theta;k,i,j)=\begin{cases}\binom{k-1}{\frac{k+i-j-1}{2}}\binom{\frac{k+i+j-3}{2}}{k-2}\ , & \textit{if } k+i+j \ \textit{odd}\ ,\\[2ex] (k+i-j)\cdot\binom{k-1}{\frac{k+i-j}{2}}\binom{\frac{k+i+j}{2}-2}{k-2}\ , & \textit{if } k+i+j \ \textit{even}\ .\end{cases} \tag{2.1}$$

(ii)

$$\begin{aligned}&\mathfrak{T}(\theta,\beta;k,i,j)=\\ &\qquad\begin{cases}(k+j-i)\cdot\binom{k-1}{\frac{k+j-i-1}{2}}\binom{\frac{k+i+j-3}{2}}{k-1}\ , \quad \textit{if } k+i+j \ \textit{odd}\ ,\\[2ex] \binom{k-1}{\frac{k+j-i}{2}-1}\binom{\frac{k+i+j}{2}-1}{k-1}+(k+j-i)\cdot\binom{k-1}{\frac{k+j-i}{2}}\binom{\frac{k+i+j}{2}-2}{k-1}\ ,\\ \qquad\qquad\qquad\qquad\qquad\qquad \textit{if } k+i+j \ \textit{even}\end{cases}\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad =\mathfrak{T}(\beta,\theta;k,i,j)\ .\end{aligned} \tag{2.2}$$

3. Smirnov Words over a Four-letter Alphabet

Remark 2 [3, Rem. A.2]: *In the framework of* [2], *the numbers* $\mathfrak{F}(\theta,\mathfrak{s};k,i,j,h)$ *of distinct Smirnov words, over the four-letter alphabet* $(\theta,\alpha,\beta,\gamma)$ *and with the Parikh vector* (k,i,j,h), *that start with the letter* θ *and end with a letter* $\mathfrak{s}\in\{\theta,\alpha\}$, *can be calculated in the following way:*

(i)

$$\begin{aligned}\mathfrak{F}(\theta,\theta;k,i,j,h) = {} & \\ & \sum_{\substack{0\le \mathrm{p}\le k-1\,,\\ 0\le \mathrm{r}\le \lfloor\frac{1}{2}(i+j+h-k+1)\rfloor}} \ \sum_{\substack{\mathrm{p}\le \mathrm{s}\le k-1\,,\\ \mathrm{r}\le \mathrm{t}\le \lfloor\frac{1}{2}(i+j+h-k+1)\rfloor}} \binom{k+\mathrm{t}-1}{k-\mathrm{s}-1,\ \ \mathrm{p},\ \ \mathrm{s}-\mathrm{p},\ \ \mathrm{r},\ \ \mathrm{t}-\mathrm{r}} \\ & \times 2^{-i-j-h+3k-2\mathrm{s}-2\mathrm{r}+4\mathrm{t}-3}\cdot 3^{i+j+h-2k+\mathrm{s}+\mathrm{r}-3\mathrm{t}+2} \\ & \times \Bigg(\binom{k-\mathrm{s}-1}{-i-j-h+3k-2\mathrm{s}-\mathrm{r}+3\mathrm{t}-3}\binom{-i-j-h+3k-2\mathrm{s}+3\mathrm{t}-3}{-i+k+\mathrm{p}-\mathrm{s}+\mathrm{t}-1}\binom{-j-h+2k-2\mathrm{p}+2\mathrm{t}-2}{-j+k-\mathrm{p}+\mathrm{t}-1} \\ & -\frac{4}{9}\cdot\binom{k-\mathrm{s}-1}{-i-j-h+3k-2\mathrm{s}-\mathrm{r}+3\mathrm{t}-1}\binom{-i-j-h+3k-2\mathrm{s}+3\mathrm{t}}{-i+k+\mathrm{p}-\mathrm{s}+\mathrm{t}}\binom{-j-h+2k-2\mathrm{p}+2\mathrm{t}}{-j+k-\mathrm{p}+\mathrm{t}} \Bigg)\ . \end{aligned} \tag{3.1}$$

(ii)

$$\begin{aligned}\mathfrak{F}(\theta,\alpha;k,i,j,h) = {} & \\ & \sum_{\substack{0\le \mathrm{p}\le k-1\,,\\ 0\le \mathrm{r}\le \lfloor\frac{1}{2}(i+j+h-k)\rfloor}} \ \sum_{\substack{\mathrm{p}\le \mathrm{s}\le k-1\,,\\ \mathrm{r}\le \mathrm{t}\le \lfloor\frac{1}{2}(i+j+h-k)\rfloor}} \binom{k+\mathrm{t}-1}{k-\mathrm{s}-1,\ \ \mathrm{p},\ \ \mathrm{s}-\mathrm{p},\ \ \mathrm{r},\ \ \mathrm{t}-\mathrm{r}} \\ & \times 2^{-i-j-h+3k-2\mathrm{s}-2\mathrm{r}+4\mathrm{t}-2}\cdot 3^{i+j+h-2k+\mathrm{s}+\mathrm{r}-3\mathrm{t}+1} \\ & \times \Bigg(\binom{k-\mathrm{s}-1}{-i-j-h+3k-2\mathrm{s}-\mathrm{r}+3\mathrm{t}-2}\binom{-i-j-h+3k-2\mathrm{s}+3\mathrm{t}-2}{-i+k+\mathrm{p}-\mathrm{s}+\mathrm{t}}\binom{-j-h+2k-2\mathrm{p}+2\mathrm{t}-2}{-j+k-\mathrm{p}+\mathrm{t}-1} \\ & + \frac{2}{3}\cdot\Bigg(\binom{k-\mathrm{s}-1}{-i-j-h+3k-2\mathrm{s}-\mathrm{r}+3\mathrm{t}-1}\binom{-i-j-h+3k-2\mathrm{s}+3\mathrm{t}-1}{-i+k+\mathrm{p}-\mathrm{s}+\mathrm{t}} \\ & + \frac{2}{3}\cdot\binom{k-\mathrm{s}-1}{-i-j-h+3k-2\mathrm{s}-\mathrm{r}+3\mathrm{t}}\binom{-i-j-h+3k-2\mathrm{s}+3\mathrm{t}}{-i+k+\mathrm{p}-\mathrm{s}+\mathrm{t}} \Bigg)\cdot\binom{-j-h+2k-2\mathrm{p}+2\mathrm{t}}{-j+k-\mathrm{p}+\mathrm{t}} \Bigg) \\ & = \mathfrak{F}(\alpha,\theta;k,i,j,h)\ . \end{aligned} \tag{3.2}$$

References

[1] *Flajolet P.*, *Sedgewick R.* Analytic combinatorics. Cambridge: Cambridge University Press, 2009.

[2] *Prodinger H.* Ternary Smirnov words and generating functions. Integers, 2018, 18, Paper A69.

[3] *Matveev A.O.* Symmetric cycles: A 2D perspective on higher dimensional discrete hypercubes, the power sets of finite sets, and set families. Leanpub: 2022.