J. Math. Biology (1982) 15: 351 – 357

Journal of
Mathematical
Biology

A Markov Chain Occurring in Enzyme Kinetics

Louis W. Shapiro[1] and Doron Zeilberger[2]

[1] Department of Mathematics, Howard University, Washington, D.C. 20059, USA
[2] Department of Theoretical Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel

Abstract. A certain Markov chain which was encountered by T. L. Hill in the study of the kinetics of a linear array of enzymes is studied. An explicit formula for the steady state probabilities is given and some conjectures raised by T. L. Hill are proved.

Key words: Markov chain – Steady state probabilities – Catalan numbers – Sequence enumeration – Enzyme kinetics

1. Introduction and Results

Consider the following *continuous time* Markov chain. The set of states is $\{0, 1\}^M$, namely all the $2^M(0-1)$ vectors with M components. The transitions are $0\alpha \to 1\alpha$, $\alpha 10\beta \to \alpha 01\beta$, $\beta 1 \to \beta 0$. *All the transition rates* are equal; α and β are (possibly empty) strings of 0's and 1's which make the above vectors have M components. For example, if $M = 9$, 010110101 may become one of the following: 110110101, 001110101, 010101101, 010110011, 010110100.

This Markov chain was considered by T. L. Hill [2], [3] (Ch. 7) as a model for the kinetics of a linear array of enzymes where 0 means "oxidized" and 1 means "reduced". Hill ([2], p. 551) observed that this also represents a model for the diffusion of a ligand across a membrane, from one bath to another, by jumping from site to site along a row of M sites. In this model 0 means 'empty' and 1 means 'occupied'. We refer the reader to [2] and [3] for a detailed discussion of the science behind the model and will go on to treat the mathematical problem of finding the steady state probabilities.

The 2^M steady state probabilities $P(s)$ $(s \in \{0, 1\}^M)$ satisfy the following system of 2^M homogeneous equations

$$k(s)P(s) = \sum_{s' \to s} P(s'), \qquad \forall s \in \{0, 1\}^M. \tag{1}$$

Here $k(s)$ is the outdegree of s, meaning the number of s'' such that $s \to s''$. Hill [2] solved the system (1) for $M = 1$ through 7 and conjectured that

$$P\text{ (1st component of } s \text{ is } 0) = (M + 2)/2(2M + 1), \tag{2}$$

which tends to $\frac{1}{4}$ as $M \to \infty$. Hill also conjectured expressions for $P(s_r = 0)$ for every

0303 – 6812/82/0015/0351/$01.40

r. We are going to give complete proofs of Hill's conjectures (Theorem 2) as well as an explicit formula for the steady state probabilities (Theorem 1).

Theorem 1. *The steady state probability of the state* $s = (s_1, \ldots, s_M)$ *is given by*

$$P(s) = \det\left[\binom{\lambda_i + 1}{j - i + 1}\right]_{k \times k} \Big/ \left[\frac{1}{M+2}\binom{2M+2}{M+1}\right], \tag{3}$$

where k is the number of 0's *in* s $(k = M - \sum_{i=1}^{M} s_i)$ *and* λ_i $(i = 1, \ldots, k)$ *is the number of* 1's *to the left of the i-th zero.*

$$\binom{a}{b} = \frac{a!}{b!(a-b)!}$$

is the binomial coefficient which equals zero if $b < 0$ *or* $b > a$.

Examples. (See [2], p. 535.) (i) If $s = (0,0,0,0,0)$, $M = 5$, $\lambda = (0,0,0,0,0)$; the determinant is 1 and $P(00000) = [\frac{1}{7}\binom{12}{6}]^{-1} = 1/132$.

(ii) If $s = (10100)$, then $M = 5$, $\lambda = (1,2,2)$ and

$$\det\left[\binom{\lambda_i + 1}{j - i + 1}\right]_{3 \times 3} = \det\begin{pmatrix} \binom{2}{1} & \binom{2}{2} & \binom{2}{3} \\ \binom{3}{0} & \binom{3}{1} & \binom{3}{2} \\ \binom{3}{-1} & \binom{3}{0} & \binom{3}{1} \end{pmatrix} = \det\begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 3 \\ 0 & 1 & 3 \end{pmatrix} = 9.$$

Thus $P(10100) = 9/132$.

Theorem 2. ([2], (14) in p. 540.) *Let*

$$C_k = \binom{2k}{k} \Big/ (k+1) \qquad (k \text{ integer}).$$

For $r = 1, \ldots, M$, *the probability that the r-th component of s is zero is given by*

$$P(s_r = 0) = \sum_{i=1}^{r} C_i C_{M+1-i} / C_{M+1}.$$

In particular $P(s_1 = 0) = C_M / C_{M+1} = (M+2)/2(2M+1)$.

The numbers

$$C_k = \binom{2k}{k} \Big/ (k+1)$$

are called the Catalan numbers and occur in many areas of Mathematics and Computer Science. We will see that the reason that the Catalan numbers come up in the present context is strongly related to the fact that the Catalan numbers enumerate "ballot sequences" (Mohanty [4], p. 2).

2. Proofs

Lemma 3. *For* $s \in \{0,1\}^M$ *let*

$$W(s) = \left\{ t = (t_1, \dots, t_M) \in \{0,1\}^M,\ \sum_{i=1}^{k} t_i \leqslant \sum_{i=1}^{k} s_i,\right.$$

$$\left. k = 1, \dots, M-1,\ \sum_{i=1}^{M} t_i = \sum_{i=1}^{M} s_i \right\},$$

$w(s) = |W(s)|$ ($|A|$ *denotes the number of elements of a set* A). *Then* $\{w(s); s \in \{0,1\}^M\}$ *satisfy* (1), *namely*

$$k(s)w(s) = \sum_{s' \to s} w(s'), \qquad \forall s \in \{0,1\}^M. \tag{4}$$

Proof. Case I: $s_1 = 0$, $s_M = 1$. Let $R \subset \{1, \dots, M\}$ be the set of indices r such that $s_{r-1} = 0$ and $s_r = 1$, and for $r \in R$ let

$$s^r = (s_1, \dots, s_{r-2}, 1, 0, s_{r+1}, \dots, s_M).$$

It is easily seen that since $s_1 = 0$ and $s_M = 1$, the set $\{s'; s' \to s\}$ equals the set $\{s^r; r \in R\}$ and $k(s) = |R| + 1$. Thus we have to prove

$$(|R| + 1)w(s) = \sum_{r \in R} w(s^r). \tag{5}$$

Let $\lambda = (\lambda_1, \dots, \lambda_k)$ be the partition ($\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_k$) corresponding to s as in the statement of Theorem 1. Conversely, given $\lambda_1 \leqslant \cdots \leqslant \lambda_k$ we associate to it

$$s = (0, \overbrace{1, 1, 1, \dots, 1}^{\lambda_1}, 0, \overbrace{1, \dots, 1}^{\lambda_2 - \lambda_1}, 0, \dots, \overbrace{1, 1, 1, \dots, 1}^{\lambda_k - \lambda_{k-1}}, 0, 1).$$

Note that $w(s) = F(\lambda)$ where $F(\lambda)$ is the cardinality of $\{(\mu_1, \dots, \mu_k);\ \mu_1 \leqslant \mu_2 \leqslant \cdots \leqslant \mu_k;\ 0 \leqslant \mu_i \leqslant \lambda_i,\ i = 1, \dots, k\}$. Since for every μ in the above set we have either $\mu_1 = 0$ or $\mu_1 > 0$, we derive the recurrence

$$F(\lambda_1, \dots, \lambda_k) = F(0, \lambda_2, \dots, \lambda_k) + F(\lambda_1 - 1, \dots, \lambda_k - 1). \tag{6}$$

Making the convention that $F(a_1, \dots, a_k) = 0$ if we do *not* have $a_1 \leqslant \cdots \leqslant a_k$, (5) can be rewritten

$$(|R| + 1)F(\lambda_1, \dots, \lambda_k) = F(1, \lambda_1, \dots, \lambda_k) + \sum_{i=1}^{k} F(\lambda_1, \dots, \lambda_i + 1, \dots, \lambda_k). \tag{7}$$

We are going to prove (7) by induction on $\lambda_1 + \cdots + \lambda_k$; when $\lambda = (0)$, $s = (0, 1)$, (7) says that $2F(0) = F(1)$, which is certainly true.

Unfortunately we have to divide the proof into subcases:

Case Ia: $\lambda_1 > 1$. Here the $|R|$ of $(0, \lambda_2, \dots, \lambda_k)$ is $|R| - 1$, and the $|R|$ of $(\lambda_1 - 1, \lambda_2 - 1, \dots, \lambda_k - 1)$ is $|R|$. We have

$$
\begin{aligned}
&(|R| + 1)F(\lambda_1, \ldots, \lambda_k) \\
&\quad \overset{(6)}{=} (|R| + 1)[F(0, \lambda_2, \ldots, \lambda_k) + F(\lambda_1 - 1, \ldots, \lambda_k - 1)] \\
&\quad = F(0, \lambda_2, \ldots, \lambda_k) + |R|F(0, \lambda_2, \ldots, \lambda_k) \\
&\qquad + (|R| + 1)F(\lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_k - 1) \\
&\quad \underset{\text{hypothesis}}{\overset{\text{inductive}}{=}} F(\lambda_2, \ldots, \lambda_k) + F(1, \lambda_2, \ldots, \lambda_k) + \sum_{i=2}^{k} F(\lambda_2, \ldots, \lambda_i + 1, \ldots, \lambda_k) \\
&\qquad + F(1, \lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_k - 1) \\
&\qquad + \sum_{i=1}^{k} F(\lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_i, \ldots, \lambda_k - 1) \\
&\quad = [F(\lambda_2, \ldots, \lambda_k) + F(\lambda_1, \lambda_2 - 1, \ldots, \lambda_k - 1)] \\
&\qquad + \sum_{i=2}^{k} [F(\lambda_2, \ldots, \lambda_i + 1, \ldots, \lambda_k) + F(\lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_i, \ldots, \lambda_k - 1)] \\
&\qquad + F(1, \lambda_2, \ldots, \lambda_k) + F(1, \lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_k - 1) \\
&\quad \overset{(6)}{=} F(\lambda_1 + 1, \lambda_2, \ldots, \lambda_k) + \sum_{i=2}^{k} F(\lambda_1, \lambda_2, \ldots, \lambda_i + 1, \ldots, \lambda_k) \\
&\qquad + F(1, \lambda_2, \ldots, \lambda_k) + F(1, \lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_k - 1) \\
&\quad = F(1, \lambda_1, \ldots, \lambda_k) + \sum_{i=1}^{k} F(\lambda_1, \lambda_2, \ldots, \lambda_i + 1, \ldots, \lambda_k) \\
&\qquad + [F(1, \lambda_2, \ldots, \lambda_k) + F(1, \lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_k - 1) - F(1, \lambda_1, \ldots, \lambda_k)].
\end{aligned}
$$

In order to establish (7) we must show that

$$F(1, \lambda_2, \ldots, \lambda_k) + F(1, \lambda_1 - 1, \ldots, \lambda_k - 1) - F(1, \lambda_1, \ldots, \lambda_k) = 0. \qquad (*)$$

Now by (6), $F(1, \lambda_1, \ldots, \lambda_k) = F(\lambda_1, \ldots, \lambda_k) + F(\lambda_1 - 1, \ldots, \lambda_k - 1)$, thus the right-hand side of (*) is

$$
\begin{aligned}
&[F(1, \lambda_2, \ldots, \lambda_k) - F(\lambda_1, \lambda_2, \ldots, \lambda_k)] + [F(1, \lambda_1 - 1, \ldots, \lambda_k - 1) \\
&\qquad - F(\lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_k - 1)] \\
&\qquad \overset{(6)}{=} F(1, \lambda_2, \ldots, \lambda_k) - F(\lambda_1, \lambda_2, \ldots, \lambda_k) + F(\lambda_1 - 2, \lambda_2 - 2, \ldots, \lambda_k - 2) \\
&\qquad = 0.
\end{aligned}
$$

The last step follows from the fact that $F(\lambda_1, \lambda_2, \ldots, \lambda_k) - F(1, \lambda_2, \ldots, \lambda_k)$ enumerates the $(\mu_1, \ldots, \mu_k)$ with $0 \leqslant \mu_1 \leqslant \cdots \leqslant \mu_k$ and $2 \leqslant \mu_i \leqslant \lambda_i$ which is equivalent to $0 \leqslant \mu_i - 2 \leqslant \lambda_i - 2$, the number of which is $F(\lambda_1 - 2, \ldots, \lambda_k - 2)$.

Case Ib: $\lambda_1 = 1$, $\lambda_2 > 1$. Here the $|R|$ of both $(\lambda_2, \ldots, \lambda_k)$ and $(\lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_k - 1)$ is $|R| - 1$. We have

$$
\begin{aligned}
(|R| + 1)&F(\lambda_1, \lambda_2, \ldots, \lambda_k) \\
&\overset{(6)}{=} (|R| + 1)[F(\lambda_2, \ldots, \lambda_k) + F(\lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_k - 1)] \\
&= F(\lambda_2, \ldots, \lambda_k) + |R|F(\lambda_2, \ldots, \lambda_k) + F(\lambda_1 - 1, \ldots, \lambda_k - 1) \\
&\quad + |R|F(\lambda_1 - 1, \ldots, \lambda_k - 1) \\
&\underset{\text{and } \lambda_1 = 1}{\overset{\text{inductive hypothesis}}{=}} F(\lambda_2, \ldots, \lambda_k) + \sum_{i=2}^{k} F(\lambda_2, \ldots, \lambda_i + 1, \ldots, \lambda_k) + F(1, \lambda_2, \ldots, \lambda_k) \\
&\quad + F(\lambda_1, \lambda_2 - 1, \ldots, \lambda_k - 1) + \sum_{i=2}^{k} F(\lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_i, \ldots, \lambda_k - 1) \\
&\quad + F(\lambda_1 - 1, \ldots, \lambda_k - 1) \\
&\underset{\lambda_1 = 1}{\overset{(6)\text{ and}}{=}} F(\lambda_1 + 1, \lambda_2, \ldots, \lambda_k) + \sum_{i=2}^{k} F(\lambda_1, \ldots, \lambda_i + 1, \ldots, \lambda_k) + F(1, \lambda_1, \ldots, \lambda_k) \\
&= \sum_{i=1}^{k} F(\lambda_1, \ldots, \lambda_i + 1, \ldots, \lambda_k) + F(1, \lambda_1, \ldots, \lambda_k). \quad \square
\end{aligned}
$$

Case Ic: $\lambda_1 = \lambda_2 = 1$. Here the $|R|$ of $(\lambda_2, \ldots, \lambda_k)$ is $|R|$ while the $|R|$ of $(\lambda_1 - 1, \ldots, \lambda_k - 1)$ is $|R| - 1$. The proof is similar to the previous cases.

Case II: $s_1 = 0$ and $s_M = 0$. Write $s = (\alpha, 0)$ where $\alpha = (s_1, \ldots, s_{M-1})$; in this case also $k(s) = |R|$. We have to show that

$$(|R| + 1)w(s) = \sum_{s' \to s} w(s'). \tag{8}$$

But from Case I for $(s, 1) = (\alpha, 0, 1)$ we have, since $k(\alpha, 0, 1) = |R| + 2$,

$$(|R| + 2)w(\alpha, 0, 1) = \sum_{s' \to (\alpha,0,1)} w(s') = \sum_{s' \to (\alpha,0)} w(s', 1) + w(\alpha, 1, 0) - w(\alpha, 1).$$

But since $w(s, 1) = w(s)$ and $w(\alpha, 0, 1) = w(\alpha, 1, 0) - w(\alpha, 1)$, (8) is established.

Case III: $s_1 = 1$, $s_M = 1$. This case is similar to Case II where we use instead $w(0, s) = w(s)$ and $w(0, 1, \alpha) = w(1, 0, \alpha) - w(0, \alpha)$.

Case IV: $s_1 = 1$, $s_M = 0$. Here $k(s) = |R|$. This case follows from Case III in the same way that Case II followed from Case I.

Lemma 4. *Let*

$$U_N = \left\{(a_1, \ldots, a_{2N}) \in \{0, 1\}^{2N},\ \sum_{i=1}^{r} a_i \leqslant r/2,\ r = 1, \ldots, 2N - 1,\ \sum_{i=1}^{2N} a_i = N\right\},$$

$$
\begin{aligned}
A_M &= \bigcup_{s \in \{0,1\}^M} W(s) \\
&= \left\{(t, s) \in \{0, 1\}^{2M};\ \sum_{i=1}^{r} t_i \leqslant \sum_{i=1}^{r} s_i,\ r = 1, \ldots, M - 1,\ \sum_{i=1}^{M} t_i = \sum_{i=1}^{M} s_i\right\}.
\end{aligned}
$$

The mapping $\sigma\colon A_M \to U_{M+1}$ *defined by*

$$\sigma(t, s) = (0, t_1, 1 - s_1, t_2, 1 - s_2, \ldots, t_i, 1 - s_i, \ldots, t_M, 1 - s_M, 1)$$

is a bijection.

Proof.

$$0 + t_1 + (1 - s_1) + \cdots + t_r + (1 - s_r) \leqslant \frac{2r+1}{2}$$

and

$$0 + t_1 + (1 - s_1) + \cdots + t_r \leqslant r - 1 + 1 = r = 2r/2.$$

Corollary 5.

$$\sum_{s \in \{0,1\}^M} w(s) = C_{M+1} = \binom{2M+2}{M+1} \Big/ (M+2).$$

Proof.

$$\sum_{s \in \{0,1\}^M} w(s) = \left| \bigcup_{s \in \{0,1\}^M} W(s) \right| = |A_M|.$$

By Lemma 4, $|A_M| = |U_{M+1}|$. But U_N (N integer) is the set of (0, 1) sequences with N 0's and N 1's such that at every component the number of 1's to the left never exceeds the number of 0's, namely the famed "ballot sequences" whose number is well known to be

$$C_N = \binom{2N}{N} \Big/ (N+1)$$

(see e.g. Mohanty [4], p. 2, Bertrand [1] or Shapiro [6]).

Lemma 6. *Let r be the number of* 0*'s in* $s = (s_1, \ldots, s_M)$ *and let* λ_i $(i = 1, \ldots, r)$ *be the number of* 1*'s to the left of the i-th zero. Then*

$$w(s) = |W(s)| = \det\left[\binom{\lambda_i + 1}{j - i + 1}\right]_{r \times r}.$$

Proof. This is a known result due to Narayana [5] which can be found in Mohanty [4], p. 32.

Proof of Theorem 1. Combine Lemma 3, Corollary 5 and Lemma 6 with the observation that $\sum_{s \in \{0,1\}^M} P(s) = 1$.

Proof of Theorem 2. By Lemma 4, $\bigcup_{s_r = 0} W(s)$ is in one-one correspondence with the set $(a_1, \ldots, a_{2M+2}) \in U_{M+1}$ with $a_{2r+1} = 1$. U_{M+1} can be identified with the set of "legal bracketings" obtained by replacing 0 by "(" and 1 by ")" (see Mohanty [4], p. 4). Then the set $\bigcup_{s_r = 0} W(s)$ is in one-one correspondence with the legal bracketings which have a right parenthesis in the $(2r + 1)$-st place. This right parenthesis has a left parenthesis companion which may be in one of the places $\{2r - 2i + 2\}$ $(i = 1, 2, \ldots, r)$. The portion between the places $2r - 2i + 2$ and

$2r + 1$ (inclusive) is also a legal bracketing of length $2i$, and there are C_i possibilities. What is left by deleting the part between the places $2r - 2i + 2$ and $2r + 1$ is a legal bracketing of length $2(M + 1 - i)$, which has C_{M+1-i} possibilities. Thus

$$\left| \bigcup_{s_r = 0} W(s) \right| = \sum_{i=1}^{r} C_i C_{M+1-i}.$$

The rest follows from Corollary 5.

Acknowledgement. We wish to thank Terrell Hill for introducing us to this fascinating problem. We also wish to thank S. Roy Caplan for introducing us to Hill. Thanks are also due to Jane Legrange for introducing us to Caplan. Finally we (especially the second author) wish to thank Destiny for introducing us to Jane Legrange.

In addition we are very grateful to the referee for very helpful comments.

References

1. Bertrand, T.: Solution d'un probleme. C. R. Acad. Sci. Paris **105**, 369 (1887)
2. Hill, Terrell L.: Steady-state kinetics of a linear array of interlocking reactions. In: Statistical mechanics and statistical methods in theory and applications. (Landman, U., ed.). New York: Plenum 1977
3. Hill, Terrell L.: Free energy transduction in biology. New York: Academic Press 1977
4. Mohanty, Sri G.: Lattice path counting and applications. New York: Academic Press 1979
5. Narayana, T. V.: A combinatorial problem and its application to probability theory. J. Indian Soc. Agric. Statist. **7**, 169–178 (1955)
6. Shapiro, L. W.: A Catalan Triangle. Discrete Mathematics. **14**, 83–90 (1976)

Received February 1/Revised May 1, 1982