Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 2879–2883, May 1988
Applied Mathematical Sciences

Interrelations between random walks on diagrams (graphs) with and without cycles

(first passage time/absorption/cycle completion/one-way cycle fluxes)

TERRELL L. HILL

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Contributed by Terrell L. Hill, January 4, 1988

ABSTRACT **Three topics are discussed. A discrete-state, continuous-time random walk with one or more absorption states can be studied by a presumably new method: some mean properties, including the mean time to absorption, can be found from a modified diagram (graph) in which each absorption state is replaced by a one-way cycle back to the starting state. The second problem is a random walk on a diagram (graph) with cycles. The walk terminates on completion of the first cycle. This walk can be replaced by an equivalent walk on a modified diagram with absorption. This absorption diagram can in turn be replaced by another modified diagram with one-way cycles back to the starting state, just as in the first problem. The third problem, important in biophysics, relates to a long-time continuous walk on a diagram with cycles. This diagram can be transformed (in two steps) to a modified, more-detailed, diagram with one-way cycles only. Thus, the one-way cycle fluxes of the original diagram can be found from the state probabilities of the modified diagram. These probabilities can themselves be obtained by simple matrix inversion (the probabilities are determined by linear algebraic steady-state equations). Thus, a simple method is now available to find one-way cycle fluxes exactly (previously Monte Carlo simulation was required to find these fluxes, with attendant fluctuations, for diagrams of any complexity). An incidental benefit of the above procedure is that it provides a simple proof of the one-way cycle flux relation $J_{n\pm} = \Pi_{n\pm}\Sigma_n/\Sigma$, where n is any cycle of the original diagram.**

This paper is concerned with some interconnections between discrete-state, continuous-time random walks with absorption, on the one hand, and random walks on certain related graphs with cycles, on the other. There are three subtopics to be considered, all of which involve the use of cycle fluxes as in the so-called diagram method (1). The term "diagram," as used in the biophysical literature, is synonymous with "graph," as used by mathematicians. In each section, below, a general procedure is proposed and then illustrated with one or more examples. The intention here is simply to introduce the three subjects.

Mean First Passage Time in Random Walks with Absorption

The problem considered in this section is a conventional discrete-state, continuous-time random walk that starts ($t = 0$) at a particular state s and ends at one or more absorption states denoted with index κ. An arbitrary state is denoted i or j. The (stationary) transition probabilities are denoted α_{ij} (for the transition: state $i \rightarrow$ state j). The states need not be arranged, with respect to transitions, in a linear array. Also, they may be infinite in number provided that there is no "escape" to infinity in the random walk and that absorption eventually occurs at one of the states κ.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

We begin with a review of the usual approach to this kind of problem and then turn to a possibly useful alternative method that allows one to calculate the mean (first passage) time to absorption (anywhere) and also the probability of eventual absorption at each state κ.

In a very large ensemble of random walks on the same graph or diagram of states, all starting from state s at $t = 0$, let $p_i(t)$ be the fraction of walks that are in state i at t. In principle, at least, all the $p_i(t)$ can be found by solving the set of linear first-order differential equations, with constant coefficients, of the form

$$\frac{dp_i}{dt} = -p_i\sum_j \alpha_{ij} + \sum_j \alpha_{ji}p_j, \qquad [1]$$

where the sums here are over those states j that can convert to state i, or vice versa, by a transition. Then the fraction of walks that end at state κ between t and $t + dt$ is

$$(dp_\kappa/dt)dt \equiv p'_\kappa dt = \alpha_{\kappa'\kappa}p_{\kappa'}(t)dt, \qquad [2]$$

where κ' is the immediate precursor of absorption state κ. The fraction of walks that end at state κ, irrespective of the arrival time, is

$$\int_0^\infty p'_\kappa dt = p_\kappa(\infty), \qquad \sum_\kappa p_\kappa(\infty) = 1. \qquad [3]$$

The mean time to absorption at any κ is then

$$\bar{t} = \int_0^\infty t\left(\sum_\kappa p'_\kappa\right)dt. \qquad [4]$$

In some problems $\bar{t}$ can be found without solving the differential equations. The mean lifetime of those walks that end at κ is, from Eq. 3,

$$\bar{t}_\kappa = \frac{1}{p_\kappa(\infty)}\int_0^\infty tp'_\kappa dt. \qquad [5]$$

Then, from Eq. 4,

$$\bar{t} = \sum_\kappa p_\kappa(\infty)\,\bar{t}_\kappa. \qquad [6]$$

The (presumably) new method, which I now describe, is capable of finding $\bar{t}$ and the $p_\kappa(\infty)$ but not the $\bar{t}_\kappa$. (I have used the discrete-time version of this method in unpublished work.) One begins with the original kinetic diagram for the random walk with absorption, as introduced above. A simple example is shown in Fig. 1*A*, where the walk starts at $s = 0$ and ends at $\kappa = -2$ or 2. Next, one constructs the "closed diagram" from the original diagram by eliminating every absorption state κ and by directing every $\kappa' \rightarrow \kappa$ arrow to state s (instead of κ). The transition probability $\alpha_{\kappa'\kappa}$ is unchanged. As an example, Fig. 1*B* shows the closed diagram obtained

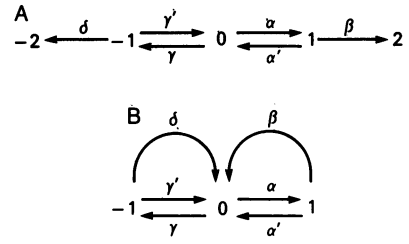

FIG. 1. (*A*) A random walk starts at state $s = 0$ at time $t = 0$. Absorption occurs at state $\kappa = -2$ or 2. (*B*) Closed diagram constructed from *A*.

from Fig. 1*A*. The closed diagram has a new one-way cycle for each absorption state in the original diagram. The steady-state properties of the closed diagram are now found by the diagram method (1). The mean steady-state cycle flux (in the one-way direction) for the cycle formed from absorption state κ is denoted J_κ (this is the mean rate of κ cycle completions). The steady-state state probabilities in the closed diagram are denoted P_i (for an arbitrary state i). Because the κ cycles are one-way cycles, we can obtain each J_κ from $J_\kappa = \alpha_{\kappa'\kappa}P_{\kappa'}$, once all the P_i are known. Thus the diagram method (1) for cycle fluxes, as applied to the closed diagram, may but need not be used. The P_i are steady-state solutions ($dP_i/dt = 0$) of linear equations such as Eq. **1** but for the closed diagram. The P_i may be obtained in any convenient way, for example, by Cramer's rule (for a finite closed diagram) or by the diagram method (1).

The physical significance of the above procedure is the following. Instead of considering, as above, a large ensemble of walks all starting at $t = 0$, to obtain ensemble averages, here we follow a single system, over a very long time. This system immediately starts a new walk from state s once a walk is completed at any κ. Thus, steady-state time averages for a single system in a long repeated walk are obtained rather than ensemble averages. P_i is the fraction of time spent in state i in the long repeated random walk. The ensemble method is more general because it includes transient properties.

For a discrete-time process, Karlin and Taylor (page 112 of ref. 2) also use a return to state s but in a rather different way (they retain κ states and introduce extra transitions $\kappa \rightarrow s$).

Having found the P_i for the closed diagram and then the J_κ from $\alpha_{\kappa'\kappa}P_{\kappa'}$, one proceeds as follows. The total mean rate of κ cycle completions is $J = \Sigma_\kappa J_\kappa$. This is also the mean rate of absorptions (at any κ) in the long repeated random walk. The reciprocal of J is the mean time per absorption: $\bar{t} = 1/J$. The fraction of walks that end with absorption at κ is the same as the fraction of κ cycle completions in the closed diagram:

$$p_\kappa(\infty) = J_\kappa/J = J_\kappa \bar{t}, \qquad J_\kappa = p_\kappa(\infty)J. \qquad [7]$$

The mean time between absorptions at κ in the repeated random walk is

$$1/J_\kappa = \bar{t}/p_\kappa(\infty). \qquad [8]$$

This is not the same as $\bar{t}_\kappa$ (Eq. **5**). If we put Eq. **7** and $\bar{t} = 1/J$ into Eq. **6**, we find

$$1 = \sum_\kappa J_\kappa \bar{t}_\kappa. \qquad [9]$$

This is a useful self-consistency relation. Also, it is easy to see that

$$P_i = \frac{1}{\bar{t}} \int_0^\infty p_i(t)dt, \qquad \bar{t} = \sum_i \int_0^\infty p_i(t)dt, \qquad [10]$$

where i refers to a non-κ state.

Let us use Fig. 1 as an example. There is "detailed balance" at steady state in Fig. 1*B*. Hence we find easily

$$P_{-1} = \gamma A/\Sigma, \qquad P_1 = \alpha B/\Sigma, \qquad P_0 = AB/\Sigma$$
$$\Sigma = \gamma A + \alpha B + AB, \quad A \equiv \beta + \alpha', \quad B \equiv \delta + \gamma'. \qquad [11]$$

Then

$$J_{-2} = \gamma\delta A/\Sigma, \qquad J_2 = \alpha\beta B/\Sigma \qquad [12]$$

$$\bar{t} = (J_{-2} + J_2)^{-1} = \Sigma/(\gamma\delta A + \alpha\beta B). \qquad [13]$$

A separate stochastic argument (omitted) leads to

$$\bar{t}_{-2} = (A^2B + \gamma A^2 + \alpha\beta A + \alpha\alpha' B)/A(\gamma\delta A + \alpha\beta B)$$
$$\bar{t}_2 = (AB^2 + \alpha B^2 + \gamma\delta B + \gamma\gamma' A)/B(\gamma\delta A + \alpha\beta B). \qquad [14]$$

Eqs. **12** and **14** are found to satisfy Eq. **9**.

Another example appears below.

Mean First Passage Time to a Cycle Completion

The section above dealt with a random walk that started at a state s and ended with absorption at one of the states κ. The mean time to absorption (at any κ) was of particular interest. In this section (as in the one below) we consider a finite diagram (graph) with at least one cycle. A random walk starts at a particular state s and ends when some one cycle is completed (in either direction) for the first time. A large ensemble of such walks is contemplated, all starting at state s and at time $t = 0$. Let an arbitrary cycle, including a particular direction (+ or −), be denoted by η. The questions that are usual for a random walk with absorption can be asked here. What fraction of the walks will end with completion of cycle η? What fraction of walks will end with cycle η, between t and $t + dt$? What is the mean lifetime of walks that end with cycle η? What is the mean time required to complete the first cycle (any η)? This problem is probably of more mathematical than physical or biological interest (but see the section below). The novel part of the solution of this problem is its translation into a random walk with absorption. After that is accomplished, the discussion in the section above applies and need not be repeated in detail here.

I use an example to illustrate and establish the general procedure. Fig. 2*A* shows a four-state diagram (graph) with three cycles (Fig. 2*B*) that are labeled a, b, and c (thus η stands for a+, a−, etc.). A line between two states in Fig. 2 represents possible transitions in either direction. The transition probabilities are α_{ij} as before. Suppose the random walk in question starts at state $s = 2$. We now construct a new diagram (the "opened diagram") from Fig. 2*A*, as in Fig. 3*A*. We start, in Fig. 3*A*, with state 2 (circled), include a line and its transition probabilities out of state 2 for each possible transition option (to state 1 or 3), and continue this process state by state, with proliferation of paths until each possible

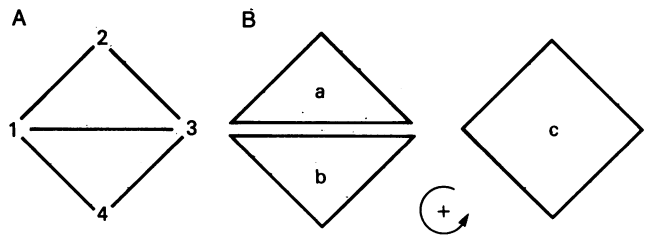

FIG. 2. (*A*) Four-state diagram. (*B*) Cycles of the diagram. The + sign indicates the positive direction for cycle completions.

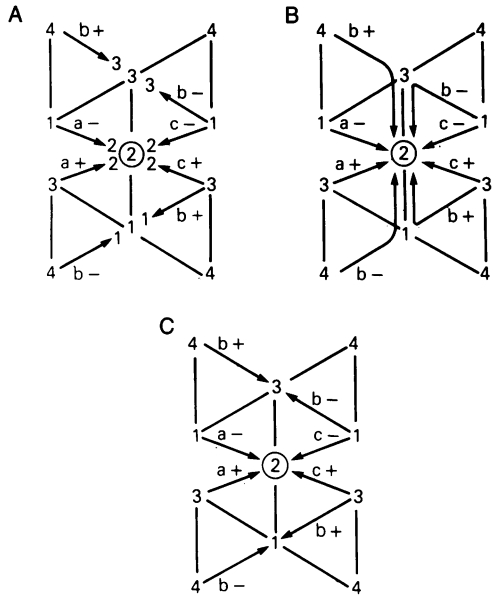

FIG. 3. (*A*) Opened diagram constructed from Fig. 2*A* when the starting state is $s = 2$. Absorption occurs at the arrowhead states, on completion of first cycles (a+, a−, etc.). (*B*) Closed diagram constructed from *A*. (*C*) Closed-2 diagram constructed from *A*.

path is ended with a one-way arrow, which indicates the first cycle completion encountered in the path. Absorption is considered to occur at the repeated state number that completes the cycle (i.e., the state number at the pointed end of the arrow). The opened diagram in Fig. 3*A* has no cycles (hence the term "opened"). That is, in graph theory nomenclature, it is a tree. The shape of Fig. 3*A* is immaterial; the interconnections between states are the only significant features. The transition probabilities are the same as in the original diagram. To distinguish among repeated states, a second index can be added.

With the opened diagram available, including its absorption states, the conventional treatment (in the section above) of a random walk with absorption can be applied. (Absorption is counted as cycle completion, as indicated by the labels on the arrows in Fig. 3*A*.) That is, linear first-order differential equations have to be solved, etc. Note that cycles b+ and b− each receive two additive contributions.

The time-averaging method in the section above can be used here as an alternative way of calculating some of the mean properties, especially the mean time to the first cycle completion (any η). For this purpose, the closed diagram derived from Fig. 3*A* (see the section above) is shown in Fig. 3*B*. There are 11 states in the closed diagram, and it includes cycles. The transition probabilities for the arrows in Fig. 3*B* are clear from Fig. 3*A*. The 11 steady-state P_i have to be found from the steady-state linear (algebraic) equations, plus the normalization of the P_i. The P_i then lead to the required cycle fluxes (see the section above). Note that the above procedure is based on two successive diagram transformations: original diagram → opened diagram → closed diagram. The second transformation is an application of the first section.

Incidentally, if the random walk in the original diagram (Fig. 2*A*) had started at state $s = 1$, the closed diagram would appear as in Fig. 4 (every cycle completion in the original diagram ends back at state 1). A second index has been added in Fig. 4, for use below. Obviously state 4 (as a starting state) is similar to state 2 and state 3 is similar to state 1.

Calculation of Cycle Fluxes in a Diagram with Cycles

The topic in the section above is primarily of mathematical

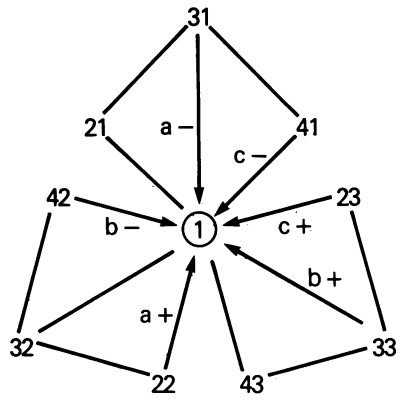

FIG. 4. Closed and closed-2 diagram, based on the original diagram of Fig. 2*A*, when the starting state is $s = 1$.

interest: in a diagram with cycles, the objects of interest are random walks that end with the *first* cycle completion. Of mathematical and also much greater practical interest, for example in biophysics (1), are random walks (on a diagram with cycles) that *continue indefinitely*. A long random walk on Fig. 2*A* would be an example. Of primary concern are the time-averaged rates of one-way cycle completions, J_η (e.g., η = a+, a−, . . . in Fig. 2*B*), and the time-averaged state probabilities p_i. In simple cases, such as Fig. 2, the diagram method (1) can provide the J_η and p_i as explicit algebraic functions of the full set of transition probabilities α_{ij}. However, if the diagram is rather complicated, only numerical results are practical: the p_i can be found from the steady-state linear algebraic equations in the p_i by matrix inversion (1) and the J_η can be found (1, 3) by Monte Carlo simulation of the random walk. What I show here is that the J_η can also be obtained (exactly, without Monte Carlo fluctuations) by simple matrix inversion if the original diagram is first opened as illustrated in Fig. 3*A* and then closed (as in Fig. 3*C*, not Fig. 3*B*).

For a given original diagram with cycles (e.g., Fig. 2*A*), we first select a starting state s (the final properties p_i and J_η do not depend on this choice, but intermediate properties do). We then construct the opened diagram as specified in the second section. However, for present purposes, the ensemble-averaged transient properties of the opened diagram are irrelevant. Here we are concerned only with time averages in a very long continuous walk. Our object, therefore, is to transform the opened diagram so as to permit a continuous walk on a new diagram that duplicates the interstate transition choices in the continuous walk on the original diagram. This is accomplished by returning each cycle-completion arrow (see Fig. 3*A*) not to state s (as in Fig. 3*B*) but to the state that originated the cycle just completed (as in Fig. 3*C*). This produces a second kind of closed diagram, denoted closed-2. Thus the transformation sequence here is: original diagram → opened diagram → closed-2 diagram. Note that Fig. 4 is both the closed diagram and the closed-2 diagram for a random walk on Fig. 2*A* that starts in state $s = 1$.

Fig. 5 presents a third example (1, 3) of a closed-2 diagram (with 24 states).

The usefulness of the closed-2 diagram arises from the fact that all of its cycles are one-way cycles. Because of this feature the state probabilities P_i of the closed-2 diagram suffice to determine all of its (one-way) cycle fluxes, just as for the J_κ in the first section. (The J_η cannot be found merely from the p_i of an arbitrary original diagram with cycles.) But even in complicated cases, it is easy to obtain the P_i of the closed-2 diagram, numerically, by computer matrix inversion. Hence, the cycle fluxes of the closed-2 diagram can be found numerically and exactly without use of Monte Carlo simulation (1, 3). The J_η of the original diagram then follow imme-

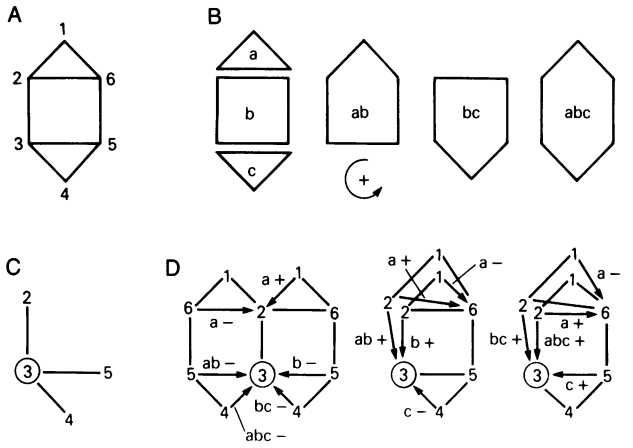

FIG. 5. (*A*) Original diagram. (*B*) Cycles of the diagram. (*C*) Possible initial steps for a walk that starts at $s = 3$. (*D*) Closed-2 diagram for the same walk. To simplify appearances, the diagram is broken into three pieces, one for each initial (or later) step out of $s = 3$. There is only *one* state 3 in the closed-2 diagram.

diately, sometimes by summation. For example, J_{b+} and J_{b-} in Fig. 2*B* each have two contributions in Fig. 3*C*. Similarly, J_{a+} and J_{a-} in Fig. 5*B* each have three contributions in Fig. 5*D*.

The closed-2 diagram is an expanded, more-detailed, version of the original diagram. The cycles of the original diagram are subdivided into one-way cycles, and some of these may be subdivided further. The states, except for the starting state, are also subdivided; a second index is needed to distinguish these but, for simplicity, a second index has not been used in Figs. 3 and 5. The four states in Fig. 2*A* are separated into 11 substates in Fig. 3*C* and into 10 substates in Fig. 4; the six states in Fig. 5*A* become 24 substates in Fig. 5*D*. Substate probabilities with the same label (first index) in the closed-2 diagram add up to the corresponding state probability in the original diagram.

The closed-2 diagram is more complicated than the original diagram because it has some history of the early steps of the random walk (prior to a cycle completion) built in, including the starting state. This history is the equivalent of the "remainders" (1, 3) that must be kept track of in a Monte Carlo simulation of a walk on the original diagram. The remainders are implicit in the closed-2 diagram and need not be considered in a walk on this diagram. In such a walk, every arrowed transition (Figs. 3*C*, 4, and 5*D*) produces a one-way cycle completion.

The transition choices and transition probabilities are the same at every step for a walk starting at state s on either the original diagram or on the closed-2 diagram. That is, the two walks are essentially identical (if the same choices are made, e.g., using the same set of random numbers in a Monte Carlo simulation). Hence, each cycle type is completed at the same rate on the two diagrams. The diagram method, applied to the original diagram, includes a proof (1) that an arbitrary cycle n has a *net* cycle flux of the form

$$J_n = (\Pi_{n+} - \Pi_{n-})\Sigma_n/\Sigma, \qquad [15]$$

where Π_{n+} is the product of transition probabilities around cycle n in the positive direction, etc. But is $J_{n+} = \Pi_{n+}\Sigma_n/\Sigma$? Monte Carlo simulation in special cases (3) verifies this one-way flux relation and proper but sophisticated proofs have been supplied (4–6). It is worth noting, however, that the closed-2 diagram itself provides another and very simple proof of the one-way flux relation because Eq. **15** in the original diagram method can be applied directly to the closed-2 diagram. In this case, for every cycle, either $\Pi_{n+} = 0$ or $\Pi_{n-} = 0$ because all cycles in the closed-2 diagram are already one-way cycles. Hence J_{n+} must be proportional to Π_{n+} and J_{n-} to Π_{n-}.

I conclude with an example that is simple enough to handle analytically. The original diagram is Fig. 2*A* and the closed-2 diagram (with $s = 1$) is Fig. 4. The state probabilities P_{ik} in Fig. 4 are found most easily in this case from the J_η and p_i of Fig. 2*A*. One can then check that the P_{ik} obtained in this way do indeed satisfy the linear steady-state algebraic equations associated with Fig. 4. For example, for state 31,

$$(\alpha_{32} + \alpha_{31} + \alpha_{34})P_{31} = \alpha_{23}P_{21} + \alpha_{43}P_{41}. \qquad [16]$$

The eight partial diagrams (1) that belong to Fig. 2*A* are shown in Fig. 6. These can be converted into directional diagrams (1) for each state i ($i = 1, 2, 3, 4$) by introducing arrows (one on each line) that lead to state i. Each arrow corresponds to a transition probability (e.g., $i \rightarrow j$ corresponds to α_{ij}). Then p_i (Fig. 2*A*) is proportional to the sum of 8 terms, each of which is a product of three transition probabilities (each partial diagram has three lines). Σ, the normalization constant, is defined as the sum of all 32 terms obtained in this way from the four states.

The diagram method (1) also gives (details omitted)

$$J_{a+} = \alpha_{21}\alpha_{13}\alpha_{32}(\alpha_{41} + \alpha_{43})/\Sigma$$
$$J_{a-} = \alpha_{12}\alpha_{31}\alpha_{23}(\alpha_{41} + \alpha_{43})/\Sigma, \qquad [17]$$

etc., for the six one-way cycle fluxes in Fig. 2 (compare Eq. **15**). In Fig. 4, we have

$$J_{a+} = \alpha_{21}P_{22}, \qquad J_{a-} = \alpha_{31}P_{31}, \qquad [18]$$

etc. Thus, from Eqs. **17** and **18**, we can deduce

$$P_{22} = \alpha_{13}\alpha_{32}(\alpha_{41} + \alpha_{43})/\Sigma$$
$$P_{31} = \alpha_{12}\alpha_{23}(\alpha_{41} + \alpha_{43})/\Sigma$$
$$P_{33} = \alpha_{14}\alpha_{43}(\alpha_{21} + \alpha_{23})/\Sigma$$
$$P_{42} = \alpha_{34}\alpha_{13}(\alpha_{21} + \alpha_{23})/\Sigma$$
$$P_{23} = \alpha_{14}\alpha_{43}\alpha_{32}/\Sigma, \qquad P_{41} = \alpha_{12}\alpha_{34}\alpha_{23}/\Sigma. \qquad [19]$$

The still unknown P_{ik} are P_{21}, P_{32}, and P_{43}. These can be obtained from

$$p_i = P_{i1} + P_{i2} + P_{i3} \qquad (i = 2, 3, 4). \qquad [20]$$

Hence, because each p_i (Fig. 2*A*) has 8 terms in the numerator, as already mentioned, P_{21} has 5 terms in the numerator, P_{32} has 4 terms, and P_{43} has 5 terms. Again, explicit details are omitted. Of course $P_1 = p_1$ (8 terms), because this state is not subdivided in Fig. 4.

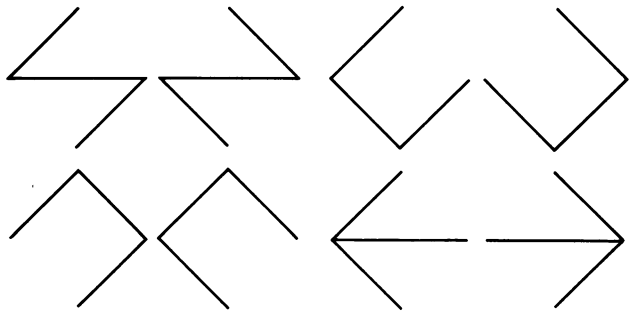

FIG. 6. The eight partial diagrams that belong to the diagram of Fig. 2*A*.

The equivalence of Figs. 2*A* and 4 can also be verified without examining the p_i and P_{ik} explicitly. For example, for state 1 in Fig. 4 (equating steady-state outflow to inflow):

$$(\alpha_{12} + \alpha_{13} + \alpha_{14})p_1 = \alpha_{21}(P_{21} + P_{22} + P_{23}) + \alpha_{31}(P_{31} + P_{32} + P_{33}) + \alpha_{41}(P_{41} + P_{42} + P_{43}) \quad [21]$$

$$= \alpha_{21}p_2 + \alpha_{31}p_3 + \alpha_{41}p_4. \quad [22]$$

Here, Eq. **21** refers to Fig. 4 (remembering that $p_1 = P_1$) and Eq. **22**, to Fig. 2*A*.

In a more complicated case, the (numerical) calculation would be done in the reverse order. The P_{ik} would be found first and then J_η and p_i would be calculated from the P_{ik}. As an example, this procedure has been applied to Fig. 5 to check the numerical results for this diagram already given in refs. 1 and 3. I am indebted to Yi-der Chen for the computer calculation.

The methods of this paper can be used also for a more complicated, hybrid kind of diagram: the original diagram includes one or more cycles and also one or more appendages that lead to absorption. This type of diagram would arise, for example, in biophysics if a poison or inhibitor molecule attaches irreversibly (leading to "absorption") to the working protein complex of the diagram in one or more of its kinetic states.

I am indebted to Dr. Doron Zeilberger for his comments on the manuscript.

1. Hill, T. L. (1977) *Free Energy Transduction in Biology* (Academic, New York).
2. Karlin, S. & Taylor, H. M. (1975) *A First Course in Stochastic Processes* (Academic, New York), 2nd Ed.
3. Hill, T. L. & Chen, Y. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 1291–1295.
4. Kohler, H.-H. & Vollmerhaus, E. (1980) *J. Math. Biol.* **9,** 275–290.
5. Qian, C., Qian, M. & Qian, M. (1981) *Sci. Sin.* **24,** 1431–1448.
6. Qian, M., Qian, M. & Qian, C. (1984) *Sci. Sin.* **27,** 470–481.