Chapter 21

GRAPH THEORETICAL METHODS FOR PHYSIOLOGICALLY BASED MODELING

HONG ZHANG

Department of Mathematical Sciences
Indian University - Purdue University
Fort Wayne, Indiana 46805

ZHEN ZHANG

Department of Biometry and Epidemiology
Medical University of South Carolina
Charleston, South Carolina 29425

I. INTRODUCTION

We propose a new mathematical approach to the physiologically based biological modeling (such as pharmacokinetic or nutritional system modeling) using a mathematical concept called graphs. A graph is defined as an object which consists of a set of vertices, a set of edges, and an incidence function which describes the incidence relation between the vertices and edges. Graph theory is a well-developed branch of mathematics and has been successfully applied to various practical problems such as electrical circuit analysis and computer network design. It has also been used in modeling biological systems (Jacquez, 1985). In many applications, the graph theoretical method is proven to be a natural and effective technique for system analyses and designs.

The graph for a physiologically based model is based on the flow diagram of the model. The vertices of the graph correspond to the junctions in the flow diagram. Each body region (or compartment) is represented by an edge in the graph. Two edges are adjacent if the corresponding body regions are linked in the flow diagram.

The graph approach is especially useful for implementing automatic modeling systems. A computer program is written to illustrate the application. The program provides a convenient graphic interface for users to set up the graph and enter the parameters. Based on the graph model the program automatically generates the equations and solves the system.

The graph method is also useful for theoretical analysis of the models. We will present some theoretical results obtained by using the graph model and applying results from graph theory. In particular, we give a method to determine a minimal set of flow rates in a model and we derive a necessary and sufficient condition for certain systems to be minimal.

II. THE GRAPH MODEL

The graph for a model is based on the flow diagram of the model. The vertices of the graph correspond to the junctions in the flow diagram. Each tissue region (or compartment) is represented by an edge in the graph. Two edges are adjacent if the corresponding body regions are linked in the flow diagram. The edges of the graph will be oriented. It is natural to orient the graph according to the directions of flows.

An example of a physiologically based model is shown in Fig. 1. The corresponding directed graph for the model is given in Fig. 2.

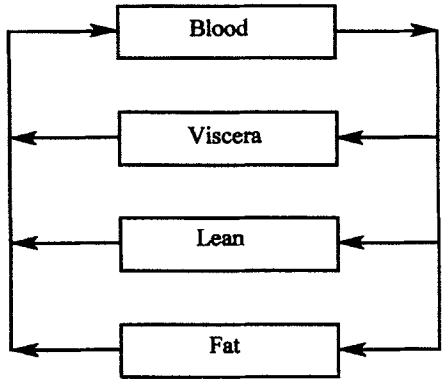

FIG. 1. A model.

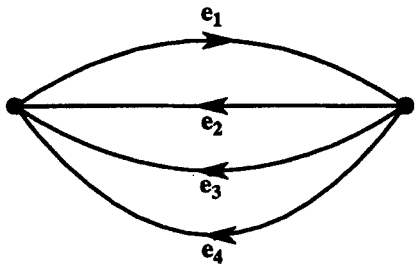

FIG. 2. Graph for model in Fig. 1.

This method of constructing the graph may seem unusual because it associates tissue regions to edges rather than vertices of the graph. However, this method will provide a more natural and mathematically convenient description of the topological structure of a physiologically based model. For conventional (not physiologically based) compartment models, representing the compartments with vertices is certainly the natural approach, since the links between compartments are usually specified separately. Physiologically based models, on the other hand, often involve more complicated mutually dependent connections among the tissue regions. For example, three or more different compartments may be connected at a common point and clearly their flow rates will be related. In this case using the vertex representation will require an additional procedure of separating the connections to several single connections between two compartments because an edge of a graph has only two end points and it cannot represent a connection among three or more compartments. The edge representation will be much more flexible because there is no restriction on the number of edges at a vertex and virtually any type of reasonable interconnections among the tissue regions can be naturally described by this method. This method will give a topological representation that is closer to the original model and that is mathematically easier to manipulate, especially when the structure of the model is complex. This type of graph model is also commonly used in electrical engineering for modeling complex electrical circuits (Deo, 1974).

The graph provides a simple description of the topological structure of the model. For simplicity, we assume that the model is flow-limited and that the tissue total concentration is combined with the equilibrium blood total concentration. Because we are mainly concerned with the modeling methodology and the generalization of the method to more complicated models is straightforward, the discussion of this simplified case is sufficient.

The mathematical equations for the model can be derived from the oriented graph of the model, the parameters of the tissue regions, and a few simple general rules. The rules are fixed for all models of a certain general category. They are not specific to a particular model and they apply to all tissue regions of the model identically. Consequently this graph model provides a simple and convenient way for computer implementations. A computer modeling program will implement the fixed set of rules. Users only need to enter the graph and the parameters for their model. The program will be able to automatically perform all the necessary calculations for any given graph and the associated parameters. There is no need to specify any special rule for any individual region. All characteristics of a tissue region are completely determined by its parameters and its relation with other tissue regions which is given by the graph.

The rules for the models we considered here are given as follows.

1. At each vertex, the sum of flow rates Q_i directed away from the vertex is equal to the sum of flow rates directed into the vertex.
2. At each vertex, the sum of mass transfer rates away from the vertex is equal to the sum of mass transfer rates into the vertex.
3. For an edge started at a vertex, the mass transfer rate directed into the edge is proportional to the flow rate on the edge.
4. For an edge terminated at a vertex, the mass transfer rate directed away from the edge is Q_iC_i/R_i, where R_i is the tissue/blood partition coefficient.

We will use the incidence matrix to represent a graph. Let G be a directed graph with m vertices and n edges. The incidence matrix $M = [M_{ij}]$ of G is a $m \times n$ matrix defined as

$$M_{ij} = \begin{cases} 1, & \text{if } v_i \text{ is the initial vertex of } e_j \\ -1, & \text{if } v_i \text{ is the terminal vertex of } e_j \\ 0, & \text{otherwise.} \end{cases}$$

We will also need to separate the positive and negative parts of the incidence matrix. Let M be the incidence matrix of a directed graph. The positive and negative parts of M, denoted by M_+ and M_-, are defined as

$$M_{ij}^{+} = \begin{cases} 1, & \text{if } M_{ij} = 1 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad M_{ij}^{-} = \begin{cases} 1, & \text{if } M_{ij} = -1 \\ 0, & \text{otherwise.} \end{cases}$$

Clearly we have $M = M_+ - M_-$.

The graph for the model shown in Fig. 1 is given in Fig. 2. Its incidence matrix, positive and negative parts are

$$M = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \end{bmatrix}$$

$$M_+ = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

$$M_- = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

We will derive the equations for a model in matrix form. The parameters for tissue regions in a model are denoted by the following matrices.

$V = diag(V_1, \ldots, V_e)$ the diagonal matrix of volumes of tissue regions.
$Q = diag(Q_2, \ldots, Q_e)$ the diagonal matrix of blood flow rates.
$R = diag(R_1, \ldots, R_e)$ the diagonal matrix of partition coefficients.
$C = (C_1, \ldots, C_e)^T$ the vector of concentrations in the tissue regions.
$In = (In_1, \ldots, In_e)^T$ the vector of rates of injections.
$Ex = (Ex_1, \ldots, Ex_e)^T$ the vector of rates of elimination.

We need an matrix operation that cannot be conveniently expressed as conventional matrix operations. This operation, denoted by A^D, acts on a square matrix A and produces a diagonal matrix with row sums of the square matrix at the diagonal entries. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix},$$

then

$$A^D = \begin{bmatrix} a_{11} + a_{12} + \ldots + a_{1n} & 0 & \ldots & 0 \\ 0 & a_{21} + a_{22} + \ldots + a_{2n} & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & a_{n1} + a_{n2} + \ldots + a_{nn} \end{bmatrix}.$$

For example, if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix},$$

then,

$$A^D = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 24 \end{bmatrix}.$$

Given the rules, incidence matrix of the graph, and parameters defined above, we may derive the differential equations for the model directly using only matrix operations.

Theorem 1 The system of differential equation for the model with the graph and parameters defined above is given by

$$V\frac{dC}{dt} = (FM_+{}^TM_- - I)QR^{-1}C + In - Ex,$$

where

$$F = Q[(M_+{}^TM_+Q)^D]^{-1}.$$

Proof. At an edge e_i of the graph, which corresponds to an tissue region in the model, the rate of change of mass satisfies the mass balance equation

$$V_i\frac{dC_i}{dt} = r_i + In_i - s_i - Ex_i,$$

where r_i and s_i denote the mass transfer rates directed into and away from the edge respectively.

Clearly s_i can be determined by Rule 4.

$$s_i = C_iQ_i/R_i.$$

To determine r_i we apply Rule 2 and Rule 3. Let the initial vertex of e_i

be v_j. Then the jth row of M_- indicates the edges directed to v_j. The sum of all mass rates directed into the vertex is given by

$$\sum_{k=1}^{n} (M_-)_{jk} C_k Q_k / R_k,$$

where $(M_-)_{jk}$ denotes the entry (j, k) of M_-. The algebraic sum of the mass rates at a vertex is 0 by Rule 2. Hence the above total sum is redistributed among the edges started at this vertex. The amount of mass transfer rate directed into an edge is proportional to the flow rate by Rule 3. Therefore, the mass transfer rate into the given edge is

$$r_i = \frac{Q_i}{\sum_{k=1}^{n} (M_+)_{jk} Q_k} \sum_{k=1}^{n} (M_-)_{jk} C_k Q_k / R_k.$$

Combining the equations into matrix form, we obtain the system equation in the theorem.

Example. Consider the model shown in Fig. 1 (cf. Bischoff, 1987).

$$M_+{}^T M_+ = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & Q_2/Q_1 & 0 & 0 \\ 0 & 0 & Q_3/Q_1 & 0 \\ 0 & 0 & 0 & Q_4/Q_1 \end{bmatrix}.$$

By theorem 1, we obtain the system equation

$$\begin{bmatrix} V_1 \cdot dC_1/dt \\ V_2 \cdot dC_2/dt \\ V_3 \cdot dC_3/dt \\ V_4 \cdot dC_4/dt \end{bmatrix} = \begin{bmatrix} -Q_1/R_1 & Q_2/R_2 & Q_3/R_3 & Q_4/R_4 \\ Q_2/R_1 & -Q_2/R_2 & 0 & 0 \\ Q_3/R_1 & 0 & -Q_3/R_3 & 0 \\ Q_4/R_1 & 0 & 0 & -Q_4/R_4 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix}$$

If the above matrix equations are expanded in component forms, we will obtain the usual differential equations for the model.

In our graph modeling technique for physiologically based systems, parallel edges can be handled naturally since we use the incidence matrices for representation. However, loops (edges with same initial and terminal

vertices) are not allowed. There is a special case that two junctions are directly connected. This may cause a loop in the graph if the junctions are identified as one vertex. We can resolve the problem by introducing a special region with volume 0 represented by an edge that corresponds to the connection between the two vertices. The above procedure can still be applied to produce the correct equations for the system. The only difference is that the equation for the special edge does not have the derivative term since the volume is 0. Hence the resulting system matrix equation is an algebraic-differential equation. This type of equations have been studied extensively and several numerical methods are available (Gear, 1971; Zwillinger 1989). This special case is illustrated by the following example.

Example. Consider the system in Fig. 3 (cf. Gibaldi and Perrier, 1982). The corresponding graph is shown in Fig. 4. Note that e_3 is the special edge used to avoid a loop. It corresponds to a region with volume 0. The incidence matrix is

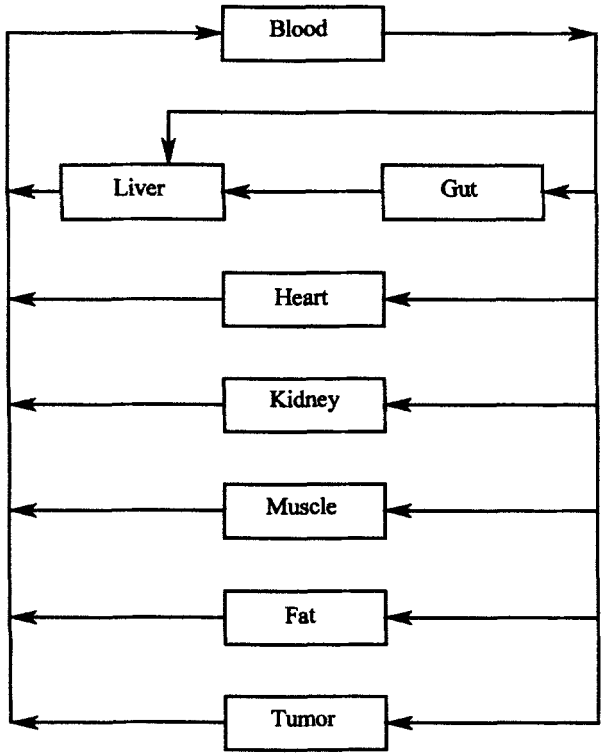

FIG. 3. Another model.

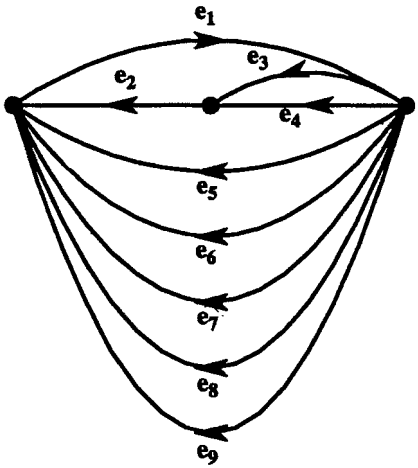

FIG. 4. Graph for model in Fig. 3.

$$M = \begin{bmatrix} 1 & -1 & 0 & 0 & -1 & -1 & -1 & -1 & -1 \\ 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

By theorem 1 the system equation can be written as

$$V_1\frac{dC_1}{dt} = \frac{Q_2}{R_2}C_2 + \frac{Q_5}{R_5}C_5 + \frac{Q_6}{R_6}C_6 + \frac{Q_7}{R_7}C_7 + \frac{Q_8}{R_8}C_8 + \frac{Q_9}{R_9}C_9 - \frac{Q_1}{R_1}C_1$$

$$V_2\frac{dC_2}{dt} = \frac{Q_3}{R_3}C_3 + \frac{Q_4}{R_4}C_4 - \frac{Q_2}{R_2}C_2$$

$$V_3\frac{dC_3}{dt} = \frac{Q_3}{R_1}C_1 - \frac{Q_3}{R_3}C_3$$

$$V_4\frac{dC_4}{dt} = \frac{Q_4}{R_1}C_1 - \frac{Q_4}{R_4}C_4$$

$$V_5\frac{dC_5}{dt} = \frac{Q_5}{R_1}C_1 - \frac{Q_5}{R_5}C_5$$

$$V_6\frac{dC_6}{dt} = \frac{Q_6}{R_1}C_1 - \frac{Q_6}{R_6}C_6$$

$$V_7\frac{dC_7}{dt} = \frac{Q_7}{R_1}C_1 - \frac{Q_7}{R_7}C_7$$

$$V_8\frac{dC_8}{dt} = \frac{Q_8}{R_1}C_1 - \frac{Q_8}{R_8}C_8$$

$$V_9\frac{dC_9}{dt} = \frac{Q_9}{R_1}C_1 - \frac{Q_9}{R_9}C_9.$$

Since $V_3 = 0$ and $R_3 = R_1 = 1$, from the third equation we have $C_3 = C_1$. Substituting the variable C_3 using this algebraic equation, we obtain the usual equation for tissue region 1.

$$V_2 \frac{dC_2}{dt} = \frac{Q_3}{R_1} C_1 + \frac{Q_4}{R_4} C_4 - \frac{Q_2}{R_2} C_2.$$

III. COMPUTER IMPLEMENTATION

This graph theoretical approach to the models has several advantages. It provides a systematical way to obtain the mathematical equations for the models. This is especially useful for computer implementation and simulation of the model.

To illustrate this application, we implemented a computer program for constructing and simulating the models. The program runs under Microsoft Windows with an easy-to-use interface. Users establish a graph model by clicking the appropriate tool buttons and drawing the vertices and edges with a mouse. Parameters for a compartment are entered through a dialogue box. The model can be conveniently edited by using operations such as move, cut, copy, and paste. Once the graph model with the associated parameters is established, the program will calculate the equations automatically. It will also solve the system numerically and plot the results of simulation.

The program achieves the generality and simplicity by applying theorem 1. It is conceptually straightforward to implement. The program maintains an incidence matrix of the model based on user inputs. Each edge object contains data structures for various parameters associated with the region. With the incidence matrix and the parameters on edges, the system equation is calculated through direct matrix operations according to theorem 1. Numerical methods are used to solve the resulting system equation.

The program can be used as a tool for teaching and research. It is available from the first author at the e-mail address zhang@cvax.ipfw.indiana.edu.

IV. ANALYSIS OF MODELS

The graph method provides a convenient mathematical tool for the analysis of the models. Here we present some results on the models obtained through the properties of their graphs.

Clearly the flow rates Q_i are not always independent because of Rule 1. A natural problem is to find a minimal set of flow rates that completely determines all flow rates in the model.

Theorem 2 A minimal set of flow rates that completely determines all flow rates contains $e - n + 1$ values.

Proof. Let $q = (Q_1, Q_2, \ldots, Q_e)^T$. Because of Rule 1, the vector q satisfies

$$M_q = 0.$$

Hence the vector q is in the cycle space of the graph. By a result in graph theory, the row space of the incidence matrix is the bond space of the graph and the nullspace of the incidence matrix (the orthogonal complement of the row space) is the cycle space of the graph. Since the graph we consider here is always connected, the cycle space of the graph has dimension $e - n + 1$ (cf. Bondy and Murty, 1976). Consequently, among the e flow rates, only $e - n + 1$ of them are linear independent. $n - 1$ values can be expressed as linear combinations of the basis vectors.

There are also techniques in graph theory to determine a basis for the cycle space, which yields a set of independent values of flow rates. One simple method involves a spanning tree of the graph. A spanning tree of a graph is a connected subgraph with the same vertex set as the original graph and with no cycles. Let T be a spanning tree of the graph. Then all values on edges not in T form an independent set which generates all values in the graph. For example, in Fig. 1, the flow rates Q_1, Q_3, and Q_4 form an independent set. The flow rate Q_2 can be expressed as a linear combination of Q_1, Q_3, and Q_4. For the model given in Fig. 3 and Fig. 4, e_2 and e_3 form a spanning tree of the graph. Hence the flow rates Q_2 and Q_3 can be expressed as linear combinations of other rates.

Another application of the graph model is to derive equivalence for certain parts of the model. In analysis and simulation of the model, it is often desirable to combine several components to an equivalent single component. The overall reduction of a model to a minimal system will be of great value not only for computations but also for theoretical study of essential characteristics of the system and comparison to the traditional compartment modeling. The graph model will facilitate the development of such equivalence transformations.

We will derive a necessary and sufficient condition for minimality of a special type of models. The models are similar to the one illustrated in Fig. 1. It contains one central region in one direction and several other regions connected to the central region in the other direction.

Theorem 3 A linear system defined above is equivalent to a system with fewer states if and only if there exist two edges satisfying

$$\frac{Q_i}{V_i R_i} = \frac{Q_i}{V_j R_j}.$$

In order to prove the theorem, we need some concepts and results from linear system theory. For convenience, we state the results here (see Kailath, 1980, for details).

A linear system is defined by

$$\frac{dx}{dt} = Ax + By$$
$$z = Cx.$$

A linear system is said to be controllable if the system can be taken to any desired state x by controlling the input function. A linear system is said to be observable if the states $x(t)$ can be determined from the observation of the output function. The following is a useful result for testing controllability and observability.

Lemma 1 (Popov–Belevitch–Hautus tests) 1. A linear system is noncontrollable if and only if there exists a nonzero row vector q such that $qA = \lambda q$ and $qB = 0$. 2. A linear system is nonobservable if and only if there exists a nonzero column vector p such that $Ap = \lambda p$ and $Cp = 0$.

The following result establishes the connection between the minimality of a system and its controllability and observability.

Lemma 2 A linear system is minimal if and only if it is both observable and controllable.

Proof of theorem 3. If the condition is satisfied, we may combined the two edges into one with the following new parameters.

$$V_* = V_i + V_j$$
$$Q_* = Q_i + Q_j$$
$$C_* = (V_i C_i + V_j C_j)/(V_i + V_j)$$
$$R_* = (V_i R_i + V_j R_j)/(V_i + V_j).$$

It is straightforward to verify that the new system is equivalent to the original system.

Conversely if the system is not minimal, then it is noncontrollable or nonobservable. By Popov–Belevitch–Hautus tests, one of the two cases is true.

Case 1: $qV^{-1}(FM_+{}^TM_- - I)QR^{-1} = \lambda q$ and $qb = 0$

$$q_1 = 0$$
$$\frac{Q_2}{V_2R_2}q_2 + \ldots + \frac{Q_m}{V_mR_m}q_m = 0$$
$$-\frac{Q_i}{V_iR_i}q_i = \lambda q_i, i = 2, \ldots, m.$$

Case 2: $V^{-1}(FM_+{}^TM_- - I)QR^{-1}p = \lambda p$ and $cp = 0$

$$p_1 = 0$$
$$\frac{Q_2}{V_1R_2}p_2 + \ldots + \frac{Q_m}{V_1R_m}p_m = 0$$
$$-\frac{Q_i}{V_iR_i}p_i = \lambda p_i, i = 2, \ldots, m.$$

Either case will lead to the condition in the theorem.

Example. For the system in Fig. 1, let $Q_3 = 2.0$, $V_3 = 1.0$, $R_3 = 0.2$, $Q_4 = 3.2$, $V_4 = 0.8$, and $R_4 = 0.4$. Then the system is not minimal since we have

$$\frac{Q_3}{V_3R_3} = 10.0 = \frac{Q_4}{Q_4R_4}.$$

According to theorem 3, compartments 3 and 4 can be combined to form a new compartment with parameters

$$V_* = V_3 + V_4 = 1.8$$
$$Q_* = V_3 + V_4 = 5.2$$
$$C_* = (V_3C_3 + V_4C_4)/(V_3 + V_4) = 0.56C_3 + 0.44\,C_4$$
$$R_* = (V_3R_3 + V_4R_4)/(V_3 + V_4) = 0.29.$$

The new system has only three tissue regions (or compartments), but it is equivalent to the original system. From the observations of compartments C_1 and C_2, the new system is indistinguishable from the original system. The new combined concentration C_* behaves as a weighted average of original values C_3 and C_4.

REFERENCES

Bischoff, K. B. (1987). Physiologically based pharmacokinetic modeling, *In* "Pharmacokinetics in Risk Assessment—Drinking Water and Health," Vol. 8, pp. 36–61. National Academy Press, Washington DC.

Bondy, J. A., and Murty, U. S. R. (1976). "Graph Theory with Applications." North-Holland Publ., New York.

Deo, N. (1974). "Graph Theory with Applications to Engineering and Computer Science." Prentice-Hall, Englewood Cliffs, NJ.

Gear, C. W. (1971). Simultaneous numerical solution of differential-algebraic equations. *IEEE Trans. Circuit Theory* **CT-18**(1), 89–95.

Gibaldi, M., and Perrier, D. (1982). "Pharmacokinetics," 2nd ed., Dekker, New York.

Jacquez, J. A. (1985). "Compartmental Analysis in Biology and Medicine," 2nd ed. Univ. of Michigan Press, Ann Arbor.

Kailath, T. (1980). "Linear Systems." Prentice-Hall, Englewood Cliffs, NJ.

Zwillinger, D. (1989). "Handbook of Differential Equations." Academic Press, San Diego, CA.