INVESTIGATION OF AN OPERATOR METHOD IN THE ANALYSIS OF BIOLOGICAL TRACER DATA

J. MYHILL, G. P. WADSWORTH, *and* G. L. BROWNELL

From the Physics Research Laboratory, Massachusetts General Hospital, Boston, and the Department of Mathematics, Massachusetts Institute of Technology, Cambridge. J. Myhill's present address is the Institute of Medical Research, Royal North Shore Hospital of Sydney, Australia.

ABSTRACT Some factors affecting the usefulness of a linear operator in the analysis of tracer data were evaluated. Application of the operator to a sum of two exponential components resulted in the separation of the rate constants with an accuracy of 10 to 15 per cent if they differed by a factor of at least 2 and the error in the data was about 2 per cent. A factor of 4 was necessary if the error in the data was 6 per cent, and of 6 if the error was 10 per cent. The ratio of amplitudes varied from near unity to equality with the ratio of rate constants. However, if the ratio of amplitudes was greater than the ratio of rate constants the method would not resolve the rate constants. Application of the operator to a sum of three exponential components was also considered.

1. INTRODUCTION

With the improvement of experimental methods in radioactive tracer studies the analysis of the data obtained is of increasing importance. In biological systems the tracer material may be simultaneously undergoing several processes, such as diffusion, excretion, or interaction. In order to reduce these complexities it is often both valid and of value to represent the biological system by a simplified compartmental model. The model consists of a series of compartments within which the isotope may mix by flow or diffusion between compartments, or be lost from the system by secretion from one or more of them. Such a model is valid to the extent that the results calculated on the basis of the model agree with those actually obtained in a real biological system.

A further simplification is afforded if it can be assumed that the system is in a steady state, that is, that the relative amount of stable material in each compartment remains unchanged during the course of the experiment. This implies that the inflow of stable material into the system equals the outflow. It has been previously

shown that the behavior of labelled substances in steady state compartmental systems can be represented by linear differential equations (Sheppard and Householder, 1951; Berman and Schoenfeld, 1956) and that the solution for the activity *versus* time curve in each compartment consists of a sum of exponentials:

$$f(t) = \sum_{i=0}^{n} N_i e^{-\lambda_i t}$$

where $f(t)$ is the specific activity, t is the time, the N_i are the amplitudes, the λ_i are the rate constants of the exponential terms, e is the base of Naperian logarithms, and the number of exponential terms, n, is equal to the number of compartments. The problem of resolving the λ_i, and estimating their values, remains very difficult. Several methods have been proposed, ranging through least-squares curve fitting with sums of exponentials to transform analysis. These will be considered in the discussion.

Compartmental models for biological systems are, in most cases, established on the basis of known physiologic properties. In the study of iodine metabolism with iodine-131, for example, the existence of compartments representing iodide in the plasma, organically bound iodine in the thyroid, and bound iodine in the plasma may be inferred from a qualitative knowledge of the system (Brownell, 1951; Riggs, 1952), though in all human subjects this model is not strictly valid (Oddie, 1955). In other systems, such as calcium or strontium metabolism, the existence of compartments other than the plasma ionic calcium or strontium is not suggested by known physiologic properties. Some evidence for the existence of compartments comes from tracer specific activity curves, which generally are limited to observations of plasma and urine. Both curves usually decrease monotonically and have a form suggestive of a sum of decreasing exponential terms (Krane *et al.*, 1956). The demonstration of a finite number of decreasing exponential terms would be important because it would imply that calcium metabolism may be represented by a finite number of compartments. The magnitude of such compartments and the rates of transfer would also be of direct use in the understanding of calcium metabolism.

It would be desirable to have a method of analysis which would yield information on whether there are exponential terms in the data and, if so, the number of terms present. The method should not explicitly depend upon the time interval over which data are collected or the magnitude of the error in the data points, though obviously a restricted time interval, or large error in the data, will render an accurate analysis more difficult.

One approach to the separation of exponential terms is through use of a linear operator. While it is known that this method is extremely sensitive to small error in the data (Lanczos, 1956), it seemed worthwhile to investigate its range of usefulness and to examine some possible refinements of the technique.

There was thus no attempt, in the studies described below, to compare different methods of obtaining the exponential terms, but only an investigation of one of them. Since this method essentially uses observations equally spaced in time, there was no consideration of the statistical design of an experiment that employs observations so spaced in time that the separation of the exponential terms is optimally facilitated. Some work in this direction has been initiated by Box (1960).

Rather than apply maximum likelihood theory, which promised very difficult functions, and also because the effect of smoothing (which, while often a hazardous procedure, has been found to be very beneficial in transform methods) was to be investigated, a simulation approach was employed.

Experimental data were simulated at equally spaced time intervals, and in other respects so as to correspond to experimental situations usually encountered, and then analyzed. The results of the analysis were compared with the known values of the parameters used to simulate the data. The extent to which the values obtained in the analysis agreed with those used to simulate the data was taken as a measure of the usefulness of the method. The simulation and analysis of the data were performed using a digital computer.

2. METHODS

(*i*) *Linear Operator.* The basis of the linear operator method is the fundamental relation which exists for the general class of linear functions; namely, a linear combination of the past is related to the future by the equation

$$f(t) = \sum_{m=1}^{k} A_m f(t - \tau m) \tag{1}$$

where k is determined by the type and number of linear functions comprising $f(t)$. In our case $f(t)$ is a function of the form

$$f(t) = \sum_{i=1}^{n} N_i e^{-\lambda_i t} \tag{2}$$

where N_i and λ_i are constants and where τ is considered a constant and is the distance between equally spaced points on the time axis at which it is presumed the data are observed. If we assume the relation (1) is true for our data the number of terms k are the same as the number of components n in equation (2). Again when it is assumed that the $f(t)$ is in the form (2), the A_m are functions only of the λ_i. When one assumes a certain number of components n in equation (2), it is then possible to set up sufficient equations for the A_m in terms of the λ_i by merely equating equivalent functions on both sides of equation (1). These derived relations are as follows, if we assume $\tau = 1$:

$$A_1 = +\sum_{i=1}^{n} e^{-\lambda_i}$$

$$A_2 = -\sum_{i} \sum_{j \neq i} e^{-\lambda_i} e^{-\lambda_j}$$

$$A_3 = +\sum_{i=1}^{n} \sum_{j \neq i} \sum_{k \neq i, \neq j} e^{-\lambda_i} e^{-\lambda_j} e^{-\lambda_k}$$

$$\vdots$$

$$A_n = (-1)^{n+1} \prod_{i=1}^{n} e^{-\lambda_i} \quad \text{if} \quad \tau = 1. \tag{3}$$

These would follow from the polynomial equation

$$x^n - A_1 x^{n-1} - A_2 x^{n-2} - \cdots - A_n = 0 \quad \text{where} \quad x = e^{-\lambda}.$$

As indicated above, the N_i are not determined by this technique, since in general the linear operator is independent of magnitude and phase, but their determination by a least-squares procedure at a later time is trivial. Obviously if there are n exponential components in the data (equation 2) then exactly n simultaneous equations of the form (1) are necessary to give the values of the A_m utilizing perfect data. The equations (3) have been previously given, for example, by Lanczos (1956), and some implications of the method will be described below.

(*ii*) *Two Components.* As an example of these equations we will present specifically the case of two exponential components. Equation (1) becomes

$$f(t) = A_1 f(t - 1) + A_2 f(t - 2) \tag{1a}$$

where equation (2) is now

$$f(t) = N_1 e^{-\lambda_1 t} + N_2 e^{-\lambda_2 t} \tag{2a}$$

and $\tau = 1$ throughout.

Using two equations (requiring four equally spaced data points) of the form (1*a*), substituting from (2*a*) for each $f(t)$ and equating coefficients of like exponentials, the values of A_1 and A_2 are found to be

$$A_1 = e^{-\lambda_1} + e^{-\lambda_2}$$

and

$$A_2 = -e^{-\lambda_1} e^{-\lambda_2} \tag{3a}$$

(*iii*) *Imperfect Data.* If the data are imperfect,—that is, represented by equation (2) plus a superimposed random error occurring at each point t,—then, obviously, n equations, of the form (1), will be sufficient to solve for the A_m as above but will yield values for the parameters which are inaccurate. More than n equations will be inconsistent. If more than n equations of the form (1) are obtained successively between the available equally spaced imperfect data points, one can solve for the A_m by using all the equations in a least-squares method provided the usual assumptions are satisfied.

If there are only n components in the data, but greater than n are assumed when writing the equation (1), then it can be easily shown that with perfect data a least-squares solution yields a singular matrix and the solution thus fails. With imperfect data the matrix is, instead, nearly singular and ridiculous values for the rate constants may result, or the solution may still fail in that the calculated values of the A_m have the wrong sign or lie outside a valid range. For example, if $n = 2$, then (from equation 3*a*) the conditions $0 < A_1 < 2$ and $-1 < A_2 < 0$ must be satisfied. If allowed values of A_m happen to be calculated, it is still possible that the solution may fail through

the polynomial in the A_m's failing to have real positive roots. This is required since each root is of the form $e^{-\lambda}$.

(*iv*) *Data Cut-off.* In all tracer experiments the data are terminated at a finite time. This has been called "data cut-off." The reasons for the termination of the data lie in the experimental situation, quite apart from questions of how the data should be analyzed. The reason may be, for example, that the radio-isotope specific activity has fallen to an unmeasurable level or that the patient has become unavailable for further study. Data cut-off is the most troublesome feature for some transform methods of tracer data analysis which involve integrations over infinite time (Callahan and Pizer, 1964).

Since the linear operator method involves a finite number of equations based on a finite number of data points, it is obvious that it is not explicitly affected by data cut-off and no special extrapolation procedures are necessary. However, no method can extract more information from data than is present in them and the results discussed below are seen to be more satisfactory if the data are extended to later times. To anticipate a result, in the case when the amplitudes (N_i) are roughly in the same ratio as the rate constants (λ_i), the data need to be extended to a time such that the last measured data point is 5 per cent, or less, of the first measured point (Table V) in order to resolve rate constants differing by a factor of 2 when 2 per cent error is present in each data point. If experimental reasons prohibit this, then the rate constants cannot be resolved.

(*v*) *Smoothing of the Data.* Raw data were generally smoothed in order to minimize the error input to the operator procedure. The trial of data smoothing was suggested, in the absence of *a priori* support from statistical theory, by the greatly improved results obtained by the transform methods for tracer data analysis when data smoothing was employed (Brownell and Callahan, 1963; Callahan and Pizer, 1964). It is shown in the results below that data smoothing improves the results obtained with the linear operator method at all levels of data error from 0 to 10 per cent.

What smoothing function to use was more difficult to decide. In the absence of any definite criterion it was decided to smooth the data points five at a time. It was thought that smoothing too many points at a time might eliminate meaningful detail from the curve, and thus mitigate against resolving two exponential terms with rate constants of nearly equal value. Smoothing too few points at a time would of course lessen the smoothing introduced. The points were smoothed with a single exponential. Apart from the fact that this would be expected to give a better fit than smoothing with a straight line (also a two parameter function), there is little to justify the choice. Functions with more parameters (*e.g.*, a parabola) were avoided since we did not wish to risk allowing the smoothing curve to follow random fluctuations in the data points too closely, or even to give a zero sum of squares by the fitted curve oscillating through every data point (*e.g.*, Worsley and Lax, 1962). Our choices are admittedly arbitrary and were made empirically with a view to obtaining improved results. There was provision in the computer program for altering both the number of data points smoothed at a time and the type of smoothing function used.

A data point was thus smoothed by fitting a single exponential term through the sequence of five points including the two points on each side of the point being smoothed. This was done consecutively for all data points. The first two and last two points were smoothed from the curves which smoothed the third and third last point respectively. Thus for p points there were $p - 4$ smoothing curves each of the form $Ne^{-\alpha t}$ with different N and α. Since, with the exception of the terminating points, only one point was

read from each curve no extraneous exponential terms were introduced into the data by this procedure.

The smoothing was effected by a simple least-squares procedure. This is valid since the deviations of the observed values from the expected values of the data points were independent and subject to approximately the same variance. In order to simulate experimental situations, the error in each data point was assumed to be proportional to the expected value of the data point, that is, a constant percentage error was assumed throughout. Since the points used when smoothing are adjacent, the absolute errors are approximately constant. In some experimental situations a constant absolute error may be more realistic, and provision for this assumption was incorporated in the computer program. However, no runs were carried out under this condition since a constant per cent error obtains in most of the tracer data to which we ultimately wish to apply our programs.

(*vi*) *Prony's Method.* When in order to solve for the n values of A_m, equations of the form (1) are set up for all data points, p, $p - n$ of these equations are inconsistent owing to the error in the data, as discussed previously. In order to utilize all of the equations, and thus all of the data, and also in an attempt to minimize the error, a quantity (being a function of the A_m) of the form

$$\sum_{i=n+1}^{p} \left[f(t_i) - \sum_{m=1}^{n} A_m f(t_i - m\tau) \right]^2 \tag{4}$$

was minimized. This is a least-squares interpretation of the method first given by Prony (1795). The method is complicated, however, by the fact that the data generally extend over one or two orders of magnitude.

(*vii*) *Weighting of the Data.* The expression (4) does not incorporate a strictly valid least-squares procedure, although it is of the correct form. Firstly, the assumptions of classical statistical theory in setting up an expression of this form involve that the $f(t_i)$ on the left of the minus sign contain error whereas the quantities on the right hand side do not. Here all terms contain error. Secondly, each term should be weighted inversely as its relative variance. There are, however, other difficulties.

When the magnitude of data points varies over a range of one or two decades, as in the case of most biological data, it is obvious that equations of the form of (1) set up between points of low magnitude (the tail of the curve) contribute very little to the sums of cross-products formed when minimizing the expression (4). In order to ensure that each equation contributed about equally to the sums of cross-products the numerical coefficients (data points) in each equation of the form (1) were divided throughout by the lowest coefficient (data point) in each equation. Each equation was thus essentially unchanged, but the coefficients became of similar magnitude in all equations. This led to the minimization of expression (5) instead of expression (4). The practical effect of this weighting procedure was that the precise magnitude of low value points, that is, the exact shape of the tail of the curve, had some influence on the sums of cross-products and thus on the calculated values of the rate constants, λ_i. Without weighting, the shape of the tail did not affect the solution for the λ_i appreciably.

A further advantage of this procedure is that, if a constant per cent error is assumed throughout the data, the terms in expression (5) contain an approximately constant absolute error, and thus the required assumptions of a least-squares fit are more nearly observed. This is more accurately the case if the weighting procedure mentioned under Results *iii* is followed. The results were found to be no better with the latter method,

however, and the weighting described here was used because it was easier to perform on the computer.

The solution for the set of weighted equations was taken as the A_m which minimized the expression

$$\sum_{i=n+1}^{p} \left[\frac{f(t_i)}{f(t_i - n\tau)} - \sum_{m=1}^{n} A_m \frac{f(t_i - m\tau)}{f(t_i - n\tau)} \right]^2 \tag{5}$$

These may be found analytically. The λ_i are uniquely determined by the A_m through equation (3).

The minimization of the expression

$$\sum_{i=n+1}^{p} \left| \frac{f(t_i)}{f(t_i - n\tau)} - \sum_{m=1}^{n} A_m \frac{f(t_i - m\tau)}{f(t_i - n\tau)} \right| \tag{6}$$

was also evaluated. This had to be done numerically.

3. RESULTS

(*i*) *Computer Procedure.* The procedure followed on the computer in evaluating the method was to generate perfect data points with known functions, being sums of exponential components. The time range to be studied was then set to correspond with some experimental situation; *e.g.,* such that the last point taken was in value 5 per cent of the initial point. A number of components, *n,* was then assumed in the analysis. Random normal error of a certain per cent standard deviation was added to all data points. An analysis as described in Methods was carried through for the rate constants. Random normal error was again added to the perfect data points. An analysis was again made, and so on, until the mean values of the rate constants obtained, and their standard deviations, could be calculated. The complete procedure was repeated with several magnitudes of random error, several assumptions on the number (n) of components, and of course several functions.

Except when otherwise stated, the first step in an analysis was to smooth the "observed" points of the simulated data as under Methods *v*. The weighted equations were then set up as described in Methods *vii*. The solution for the A_m [this generally minimized the expression (5)] was then found, and the polynomial in the A_m (Methods *i*) solved to give the roots $e^{-\lambda_i}$ and thus the λ_i.

The approach was programmed for the IBM 7090. A solution for n components was deemed to have failed (Methods *iii*) if the matrix obtained in solving for the A_m was singular, or so nearly singular that underflow or overflow occurred on the computer, or if any of the A_m had the wrong sign, or if the roots of the polynomial were not all real, positive. The computer program solved the polynomial in the complex plane. If the imaginary part of a root were greater than 0.001 times the real part, then the root was considered to be not purely real.

If a solution for n components succeeded, but for $n + 1$ components failed, then

n components were considered to be present in the data. It was decided to investigate what variables in the method facilitated the analysis, what ratio of rate constants permitted their values to be resolved, and what factors would ensure that spurious rate constants were not found.

(*ii*) *Minimizing Technique Which Best Estimated the* λ_i. A comparsion between the λ_i obtained when expressions (5) and (6) were minimized was made on the function $f(t) = 0.1e^{-0.1t} + 0.01e^{-0.01t}$, over the range $0 < t < 80$, when normally distributed random error was added. The standard deviation of the error was varied from 1 through 5 per cent. Unsmoothed, but weighted data points were used for this comparison. The analysis was repeated ten times at each error value and the mean λ_i and the SD of an individual calculated λ_i was computed. The results in Table I show that the SD's were generally smaller when the expression (5)

TABLE I

COMPARISON OF VALUES OF RATE CONSTANTS WHICH MINIMIZED DIFFERENT EXPRESSIONS

Error in data	Sum of squared differences minimized				Sum of absolute differences minimized			
	λ_1		λ_2		λ_1		λ_2	
%	*mean*	SD	*mean*	SD	*mean*	SD	*mean*	SD
0	0.100		0.0100		0.100		0.0100	
1	0.117	± 0.006	0.0154	± 0.0017	0.119	± 0.013	0.0194	± 0.0070
2	0.173	± 0.028	0.0256	± 0.0026	0.166	± 0.047	0.0255	± 0.0099
3	0.260	± 0.047	0.0293	± 0.0028	0.265	± 0.094	0.0307	± 0.0087
4	0.422	± 0.111	0.0329	± 0.0019	0.397	± 0.122	0.0311	± 0.0091
5	0.676	± 0.239	0.0343	± 0.0009	0.932	± 0.578	0.0327	± 0.0044

was minimized, as would be expected, and that there was no systematic difference between the mean λ_i obtained by the two methods.

The minimization of expression (6) was very time consuming on the computer. Since function (5) could be minimized analytically it was a faster procedure. It yielded equally good mean values, and yielded lower SD's. This function was minimized in most of the remaining studies, though with unsmoothed data terminating at early times this technique failed to give a solution when the other technique succeeded. This occurred when unsmoothed points were analyzed from $f(t) = 0.1e^{-0.1t} + 0.01e^{-0.01t}$, $0 < t < 40$, with 4 per cent error added to the perfect data points.

No studies were made of the minimization of expressions other than forms (5) or (6).

(*iii*) *Comparison of Two Weighting Procedures.* In addition to the weighting procedure described in Methods *vii* another simple technique was investigated using the function, $f(t) = 0.1e^{-0.1t} + 0.01e^{-0.01t}$, $0 < t < 40$, data

points unsmoothed. A straight line was fitted to the curve on semi-log paper, its equation was $f(t') = 0.086e^{-0.063t'}$. At each time the data point from $f(t)$, with random error added as before, was divided by the corresponding point on $f(t')$ and the operator equations were set up among the weighted points so obtained. The results did not differ appreciably, though the method of dividing by the lowest data point described previously gave slightly better results.

(*iv*) *Data Cut-off.* The rate constants estimated from data terminated early in time were shown to be subject to greater error than if the data extended to later times. Indeed, if the data cut-off was too early and the error in the data too great, a solution was completely impossible. Table II compares the rate constants obtained under two cut-off conditions for unsmoothed data points. At lower cut-off the solutions were poorer at 1 and 2 per cent error in the data, and the

TABLE II

EFFECT OF DATA CUT-OFF ON SOLUTIONS FOR THE RATE CONSTANTS

Function = $0.1e^{-0.1t} + 0.01e^{-0.01t}$, with normally distributed random error added. No smoothing

Error added	$0 < t < 40$				$0 < t < 80$			
	λ_1		λ_2		λ_1		λ_2	
%	*mean*	SD	*mean*	SD	*mean*	SD	*mean*	SD
0	0.100		0.010		0.100		0.010	
1	0.393	± 0.049	0.057	± 0.021	0.117	± 0.006	0.015	± 0.002
2	1.58	± 0.63	0.062	± 0.001	0.173	± 0.028	0.026	± 0.003
3	1.83*	± 0.64	0.062	± 0.001	0.260	± 0.047	0.029	± 0.003
4	No solution		No solution		0.422	± 0.111	0.033	± 0.002
5	No solution		No solution		0.676	± 0.239	0.034	± 0.001

*Only 3 solutions out of 10 trials. Everywhere else 10 solutions out of 10 trials were obtained.

method failed to find solutions at errors of 4 and 5 per cent. This emphasizes the desirability of accumulating data over as long a time span as possible.

(*v*) *Error in the Data and Smoothing.* It has been obvious from Tables I and II that the greater the error in the data the greater the error in the calculated rate constants. The error sometimes became so large as to make a solution impossible as obtained in Table II. Smoothing the data greatly reduced the error in the rate constants. In Table III the results are seen to be greatly improved over those in Table II (the four right-hand columns). The calculated mean λ_i values were much closer to the true values (at 3 per cent error a difference of 1 per cent instead of 160 per cent in λ_1) and the standard deviations of the λ_i were generally much smaller. Smoothing was employed in all the remaining studies.

(*vi*) *Effect of Varying the Interval between Data Points.* If the interval between data points was very small the errors introduced into the operator equation

TABLE III

CALCULATED RATE CONSTANTS FOR SMOOTHED DATA POINTS

Function = $0.1e^{-0.1t} + 0.01e^{-0.01t}$, $0 < t < 80$, with normally distributed random error added and the points then smoothed

Error in data	λ_1		λ_2	
%	*mean*	SD	*mean*	SD
0	0.100		0.0100	
1	0.0992	± 0.0016	0.0101	± 0.0007
2	0.0992	± 0.0024	0.0106	± 0.0012
3	0.1013	± 0.0035	0.0115	± 0.0014
4	0.1060	± 0.0062	0.0129	± 0.0021
5	0.1086	± 0.0080	0.0132	± 0.0028
10	0.1348	± 0.0244	0.0185	± 0.0046

were so great as to prohibit any solution at all being found. This was to be expected since, as a constant per cent error is maintained in the data points, it is obvious that the probability that an imperfect data point at $t = t_1$ should be less in value than one at $t = t_2$, where $t_1 < t_2$, increased as $t_1 \rightarrow t_2$. If equations of the form (1) are used with $f(t - 1) < f(t)$ it is clear that valid solutions for the A_m will not be obtained.

Provided this situation was not reached, a larger number of points used with a given time span were found to give more certain results. This was illustrated in Table IV where a division into 30 intervals gave only one spurious three component solution. Division into less intervals gave more spurious three component solutions and thus greater uncertainty as to the number of exponential terms actually present. However, in the situation where the number of components was known, a division into 10 or 15 intervals gave the most accurate estimate of the values of the two rate constants.

This suggests the possibility of analyzing with a large number of intervals to ascertain the number of components present, and then increasing the interval size (using fewer intervals) to refine the values of the calculated λ_i.

(*vii*) *Separation of Two Components.* Considering the situation where the amplitude ratios were the same as the rate constant ratios (Table V), if the error in the data was 2 per cent, as in some good calcium tracer experiments, rate constants differing by a factor of 2 were resolved and the λ_i calculated with an average accuracy of 10 to 15 per cent. If the error in the data was 6 per cent a rate constant ratio of 4 was required for the same accuracy. If the error was 10 per cent a ratio of about 6 was necessary. In each case the components were resolved using 30 intervals and the values of the λ_i refined using 10 intervals.

If the ratio of amplitudes was greater than the ratio of rate constants (Table VI) accurate values of the rate constants were not obtained by the operator method. If the amplitude ratio was less than the ratio of rate constants, however, rate constant

TABLE IV

EFFECT OF VARYING THE INTERVAL BETWEEN DATA POINTS

Function = $0.1e^{-0.1t} + 0.01e^{-0.01t}$, $0 < t < 80$. Smoothed Data

	30 intervals				15 intervals				10 interva			
Error in data	No. of 2 component solutions	λ_1	λ_2	No. of spurious 3 component solutions	No. of 2 component solutions	λ_1	λ_2	No. of spurious 3 component solutions	No. of 2 component solutions	λ_1	λ_2	No. of spurious 3 component solutions
%		*mean*	*mean*			*mean*	*mean*			*mean*	*mean*	
2	10	0.101	0.0106	0	10	0.096	0.0105	2	10	0.098	0.0129	0
4	10	0.103	0.0122	1	10	0.101	0.0108	0	10	0.097	0.0128	1
6	10	0.111	0.0149	0	10	0.095	0.0101	3	10	0.102	0.0132	1
10	10	0.135	0.0185	0	10	0.106	0.0123	2	10	0.106	0.0161	0

TABLE V

SEPARATION OF TWO EXPONENTIAL COMPONENTS

(Ratio of rate constants equal to ratio of amplitudes)

Function: $ae^{-at} + 0.01e^{-0.01t}$, last point 5 per cent of the first point. Smoothed points

		10 intervals				30 intervals				60 intervals			
Value of a	Error in data	No. of 2 component solutions	λ_1	λ_2	No. of 3 component solutions*	No. of 2 component solutions	λ_1	λ_2	No. of 3 component solutions*	No. of 2 component solutions	λ_1	λ_2	No. of 3 component solutions*
	%		*mean*	*mean*			*mean*	*mean*			*mean*	*mean*	
0.10	2	10	0.100	0.0170	5	10	0.104	0.0124	0	10	0.129	0.0233	0
$0 < t < 62$	6	10	0.105	0.0176	0	10	0.127	0.0228	0	10	0.354	0.0394	0
	10	10	0.108	0.0215	2	10	0.176	0.0292	0	10	0.816	0.0439	0
0.08	2	10	0.077	0.0134	7	10	0.081	0.0109	0	10	0.098	0.0186	0
$0 < t < 82$	6	10	0.083	0.0137	1	10	0.096	0.0180	0	10	0.261	0.0299	0
	10	10	0.080	0.0151	2	10	0.124	0.0229	1	10	0.559	0.0338	0
0.06	2	10	0.058	0.0119	1	10	0.060	0.0109	0	10	0.0791	0.0168	0
$0 < t < 108$	6	10	0.059	0.0123	2	10	0.082	0.0162	0	10	0.278	0.0248	0
	10	10	0.063	0.0131	4	10	0.106	0.0178	0	10	0.478	0.0259	0
0.04	2	10	0.039	0.0119	3	10	0.042	0.0112	0	10	0.0492	0.0406	0
$0 < t < 143$	6	10	0.040	0.0116	2	10	0.069	0.0160	0	10	0.340	0.0200	0
	10	10	0.044	0.0130	2	10	0.096	0.0165	2	10	0.689	0.0203	0
0.02	2	10	0.0216	0.0116	4	10	0.047	0.0134	0				
$0 < t < 212$	6	10	0.0282	0.0118	2	10	0.164	0.0140	2				
	10	10	0.0346	0.0120	4	10	0.168	0.0136	2				
0.015	2	10	0.0240	0.0116	2	10	0.116	0.0121	1				
$0 < t < 245$	6	8	0.0479	0.0119	0	8	0.211	0.0123	0				
	10	9	0.0854	0.0121	0	6	0.251	0.0122	0				
0.011	2	10	0.0340	0.0105	0	10	0.167	0.0105	0				
$0 < t < 284$	6	8	0.0415	0.0101	0	8	0.183	0.0105	0				
	10	9	0.0394	0.0107	1	8	0.140	0.0105	1				

*These solutions are spurious.

TABLE VI

SEPARATION OF TWO COMPONENTS

(Ratio of rate constants not necessarily equal to ratio of amplitudes).

Function = $ae^{-bt} + 0.01e^{-0.01t}$, last point 3 per cent of first point. Smoothed points. 40 intervals

Ratio of rate constants $100b$	Error in data	Ratio of amplitudes, $100a$.											
		10		8		6		4		2		1.5	
		λ_1	λ_2	λ_1	λ_2	λ_1	λ_2	λ_1	λ_2	λ_1	λ_2	λ_1	λ_2
	%	*mean*	*mean*	*mean*	*mean*	*mean*	*mean*	*mean*	*mean*	*mean*	*mean*	*mean*	*mean*
10	2	0.099	0.0106	—	—	—	—	0.091	0.0099	—	—	—	—
	6	0.112	0.0137	—	—	—	—	0.100	0.0104	—	—	—	—
8	2	—	—	—	—	—	—	—	—	0.0698	0.0098	—	—
	6	—	—	—	—	—	—	—	—	0.0746	0.0103	—	—
6	2	—	—	—	—	0.0609	0.0105	—	—	—	—	0.0507	0.0099
	6	—	—	—	—	0.0754	0.0134	—	—	—	—	0.0672	0.0109
4	2	0.036	0.031	—	—	—	—	0.0422	0.0107	—	—	—	—
	6	0.161	0.025	—	—	—	—	0.0680	0.0144	—	—	—	—
2	2	—	—	0.155	0.017	—	—	—	—	0.0560	0.0132	—	—
	6	—	—	0.296	0.017	—	—	—	—	0.196	0.0134	—	—
1.5	2	—	—	—	—	0.283	0.0137	—	—	—	—	0.167	0.0120
	6	—	—	—	—	0.358	0.0137	—	—	—	—	0.232	0.0120

values accurate to 10 per cent or better were obtained even with 6 per cent error in the data.

(*viii*) *Separation of Three Components.* The separation of three components was more difficult since the order of the matrix evaluated in solving for the A_m increased as the assumption on the number of components present in the data increased. When every element of a matrix contains error it is well known that the difficulty of accurate evaluation increases geometrically with increasing order.

A second difficulty was that when a single exponential was fitted to a set of five points in the smoothing procedure the smoothed values were considerably in error if the data points were well separated. This is clear if one considers that a sum of exponential terms plotted on semilog paper is a curve, and fitting a single exponential to five of them is approximating a curve by a straight line on the semilog paper. The point being smoothed is, on the average, increased in value and this effect is worse in those regions of the curve which have higher curvature. Thus, the true ratio between the λ_i was on the average lessened by the smoothing procedure before the linear operator was applied.

A further difficulty was that a larger number of data points (or time intervals) within a given time span was needed to adequately define three exponential components, in contrast with two. But it has been shown above that better values of the λ_i were obtained with fewer time intervals, so some compromise had to be made in selecting the number of intervals used in the analysis.

Using the function $f(t) = 0.1e^{-0.1t} + 0.01e^{-0.01t} + 0.001e^{-0.001t}$ with random normal error added, no solutions could be obtained if the data were terminated such that the last point was in value 3 per cent or 2 per cent of the initial point. If the last point was 1 per cent of the initial point, however, a λ_i ratio of 6 or 10 was resolvable with about 2 per cent error in the data, and a ratio of 4 generally resolvable with 1 per cent error in the data (Table VII). The values of the λ_i obtained were not accurate.

If the ratio of the λ_i was maintained at 10, but the ratio of amplitudes reduced to 1, then the rate constants could be resolved sometimes (6 out of 10 trials) but not always.

(*ix*) *Possible Improvement of the Methods.* There are some possibilities of improving these results. Firstly, if a larger number of points are taken, and smoothing effected among these, and then a smaller number of intervals taken over the same time span by using, for example, every fourth smoothed point, then the second difficulty referred to above would be somewhat reduced.

The question of what smoothing function to use is very important and requires more investigation. A fit with a sum of exponentials piecewise along the data may further reduce the difficulty of falsely raising the values of points in regions of higher curvature.

TABLE VII

SEPARATION OF THREE COMPONENTS

Function = $ae^{-at} + be^{-bt} + 0.001e^{-0.001t}$, $a = 1000b^2$. Smoothed data. Twenty intervals

Ratio value of b	Error added	No. of 3 component solutions	λ_1	λ_2	λ_3	No. of spurious 4 component solutions
	%		*mean*	*mean*	*mean*	
0.01	1	10	0.156	0.105	0.0066	0
	3	6	0.187	0.108	0.0066	0
0.006	1	10	0.066	0.023	0.0023	0
	3	10	0.069	0.023	0.0023	0
0.004	1	9	0.045	0.011	0.00146	0
	3	7	0.030	0.011	0.00143	0

Another possibility is that of examining the data in sections by the operator method. When one component is determined accurately by an analysis, for example, of the tail of some data, by assuming two components in the operator, then this calculated rate constant could be entered as a constant in the further analyses.

Operator methods would generally be more applicable to real data if they were concerned with points distributed logarithmically in time or if the interval sizes between points were increasing at a constant rate. The difficulty here is to devise a method which is economical of machine time since such an operator will not be linear.

4. DISCUSSION

At the outset of most tracer experiments two questions must generally be answered: Is a compartmental analysis useful? Is the system in a steady state? While mathematical techniques are available for analyzing tracer data without assuming the existence of compartments (Bergner, 1964) it seems both justified and useful to employ mathematical compartments if there is biological evidence for them. In the absence of biological evidence, it seems generally impossible to obtain from tracer specific activity curves rigorous mathematical evidence for the existence of compartments.

It seems that the existence of a steady state in a compartmental biological system must be decided on biological grounds, aided by stable isotope and tracer measurements within the compartments. Rigorous mathematical criteria for determining a stable state system from specific activity curves have not been yet fully developed. Mathematical techniques for treating compartmental systems not in a steady state are available (Sheppard, 1948). If a biological system is in a

steady state, and there is instantaneous mixing of isotope in all compartments, then the existence of n exponential components in an experimental specific activity curve is evidence that there are n compartments in the biological system (Sheppard and Householder, 1951; Berman and Schoenfeld, 1956). The identification of these mathematical compartments with real physiological compartments is often difficult, especially in calcium studies (Krane *et al.*, 1956; Glass and Nordin, 1963). However even in the case of a steady state system with a finite number of physical or chemical compartments, if intracompartmental concentration gradients exist, further exponential components which are not interpretable as corresponding to extra compartments will be introduced (Hart, 1957). If it may be assumed that there are no intracompartmental concentration gradients, then the analysis of the system resolves into the determination of the amplitudes and rate constants of the exponential terms which contribute to the specific activity curves. The determination of the number of exponential components, if any, that may be present in a specific activity curve is difficult, and it is always safer to work with curves from several compartments.

Analog computers have been used extensively in testing proposed models by the comparison of simulated curves and data by eye; (*e.g.* Brownell *et al.*, 1953). This technique is most effective in analyzing a system in which several compartments can be observed and in which the specific activity curves undergo considerable variation. With the recent development of reliable commercial analog computers, and of the digitalization of some of the steps by using digital-analog computers, this technique continues to be of value.

Curve peeling methods (Perl, 1960) are useful if the data extends over a sufficiently long time interval for the slowest component to be accurately estimated. However, the accuracy of estimation of each rate constant decreases for successively estimated constants, and a decision must be made each time as to where each exponential term ceases to contribute to the data, and this further increases the error. Further, this method does not provide objective evidence of the existence of a finite number of exponential terms.

Fitting an observed curve with a sum of exponential terms is usually done by commencing with an assumption of one exponential component and increasing the number of components until the optimum fit is obtained (Worsley and Lax, 1962; Rogers, 1962). This procedure is best performed on an automatic digital computer and can be incorporated as part of a more general automated system of tracer data analysis (Berman, Weiss, and Shahn, 1962; Berman, Shahn, and Weiss, 1962). The number of components assumed to be present in the tracer curve is that number an excess over which does not decrease the sum of squares of the residuals of the observed data points compared with the calculated data points. In addition the fit given by this number of components must not result in any systematic deviation of the calculated curve from the experimental points either at an end or

elsewhere within the time span over which the data are collected. It is sometimes very difficult to determine by these means the exact number of components present in the data, particularly if the rate constants differ by a small amount. The existence of error in the data points further complicates the analysis considerably. Rogers (1962) finds that it is very difficult to determine the decay constants for components which: (*a*) have very long or very short half-lives compared with the time covered in the experiment; (*b*) are very similar in half-life to other components (considerable difficulties occur when one is less than 1.5 times another); (*c*) contribute only a very small portion to the total decay curve. Worsley and Lax (1962) find that the values of the rate constants obtained are very sensitive both to error in the data and the precise nature of the curve fitted to the data points. Derived physiological parameters at zero time, however, were much more stable.

The method of Prony (1795) as applied by Lanczos (1956) also yielded unstable values for the rate constants. Householder (1950, 1964) modified Prony's method to achieve statistical rigor, but feels that it is inconclusive that any improvement was obtained by using the modified method.

Unfortunately, precise information on the conditions under which rate constants were resolvable, including the ratios between the rate constants, is not available in the above references for comparison with the methods outlined in this paper. By these methods, for two exponential components and an experimental error of 2 per cent in each data point, a ratio of 2 allowed the rate constants to be resolved. An error of 6 per cent required a ratio of 4, and 10 per cent a ratio of 6. With three components and 2 per cent error a ratio of about 6 was necessary, but a greater investigation into the analysis of three component curves is required.

Methods of tracer data analysis involving mathematical transforms have been recently investigated. One method involving the use of a modified Fourier transform (Gardner *et al.*, 1959) is intrinsically capable of resolving components whose rate constants differ by small amounts, even when large error is present in the data, but this method requires that the observed data points extend over a considerable time interval, generally much greater than that obtained in experimental practice (Callahan and Pizer, 1964). A method using a modified Laplace transform (Brownell *et al.*, 1961; Brownell and Callahan, 1963) is useful for data with large experimental error and a reasonable time interval of data collection, but is intrinsically incapable of resolving exponential components whose rate constants differ by less than a ratio of 4.

Other methods for resolving exponential components involving various integrals or polynomial expansions have been suggested but seem to be equally sensitive to error in the data (Landahl, 1963; Lanczos, 1963).

The major disadvantage of many of the integral transform techniques and of Prony's operator is that the methods require that the data points be essentially equally spaced in time. If the data points are smoothed, which is always necessary

in any integral procedure, then this necessity is bypassed. However, most of these methods continue to weight the data curves equally over all time, whereas it would be more desirable to weight the data in each time region in proportion to the number of data points observed in that region and according to their variances.

In most biological experiments the data points are measured approximately evenly on the logarithmic time scale. Thus many points are taken early in an experiment, perhaps at intervals of a few minutes, while later in the experiment, say after 30 days, the points may be taken daily or even weekly. It would be desirable to have a method which operated on discrete data points themselves and which was applicable to data points distributed approximately evenly in the logarithmic time scale. The distribution of data points required by the method of analysis could then be easily replicated in an experiment.

None of the methods of resolving exponential components have been subjected to a complete error analysis. Consequently it is not known which method is best under each of a variety of circumstances. The variable circumstances are: magnitude and type of error in the data, time interval over which the data is collected, values of rate constants and values of amplitudes. It is probable that different methods would have merit under different circumstances. A complete error analysis of all methods is called for, either by means of maximum likelihood theory or by using a simulation approach similar to that employed here.

Received for publication, November 26, 1963.

REFERENCES

BERGNER, P.-E. E., 1964, *J. Theoret. Biol.*, **6**, 137.
BERMAN, M., and SCHOENFELD, R., 1956, *J. Appl. Physics*, **27**, 1361.
BERMAN, M., SHAHN, E., and WEISS, M. F., 1962, *Biophysic. J.*, **2**, 275.
BERMAN, M., WEISS, M. F., and SHAHN, E., 1962, *Biophysic. J.*, **2**, 289.
BOX, G. E. P., 1960, *Ann. New York Acad. Sc.*, **86**, 792.
BROWNELL, G. L., 1951, *J. Clin. Endocrinol.*, **11**, 1095.
BROWNELL, G. L., and CALLAHAN, A. B., 1963, *Ann. New York Acad. Sc.*, **108**, 172.
BROWNELL, G. L., CAVICCHI, R. V., and PERRY, K. E., 1953, *Rev. Scient. Instr.*, **24**, 704.
BROWNELL, G. L., GREENFIELD, E., and BRENTANI, P., 1961, Radioaktive Isotope *in* Klinik und Forschung, **4**, *Strahlentherapie*, suppl. 45, 1.
CALLAHAN, A. B., and PIZER, S. M., 1964, data in preparation.
GARDNER, D. G., GARDNER, J. C., LAUSH, G., and MEINKE, W. W., 1959, *J. Chem. Physics*, **31**, 978.
GLASS, H. I., and NORDIN, B. E. C., 1963, *Physics Med. and Biol.*, **8**, 387.
HART, H. E., *Bull. Math. Biophysics*, 1957, **19**, 61.
HOUSEHOLDER, A. S., 1950, ORNL-455, Oak Ridge, Tennessee, Oak Ridge National Laboratory.
HOUSEHOLDER, A. S., private communication, 1964.
KRANE, S. M., BROWNELL, G. L., STANBURY, J. B., and CORRIGAN, H., 1956, *J. Clin. Inv.*, **35**, 874.
LANCZOS, C., 1956. Applied Analysis, Englewood Cliffs, New Jersey, Prentice Hall, Inc., 272.
LANCZOS, C., private communication, 1963.
LANDAHL, H. D., 1963, *Ann. New York Acad. Sc.*, **108**, 331.

ODDIE, T. H., MESCHAN, I., and WORTHAM, J., 1955, *J. Clin. Invest.*, **34**, 95, 106.
PERL, W., 1960, *Internat. J. Appl. Radiation and Isotopes*, **8**, 211.
PRONY, A. L., 1795, *J. école polytech. Paris*, **2**, 29.
RIGGS, D. S., 1952, *Pharmacol. Rev.*, **4**, 284.
ROGERS, P. C., 1962, Technical Report No. 76, Laboratory for Nuclear Science, Massachusetts Institute of Technology, Cambridge, Massachusetts.
SHEPPARD, C. W., 1948, *J. Appl. Physics*, **19**, 70.
SHEPPARD, C. W., and HOUSEHOLDER, A. S., 1951, *J. Appl. Physics*, **22**, 510.
WORSLEY, B. H., and LAX, L. C., 1962, *Biochim. et Biophysica Acta*, **59**, 1.