

SOME STATISTICAL METHODS FOR THE
ANALYSIS OF CHEMICAL REACTION RATE DATA

by

Anna Christine Miller

Thesis submitted to the Graduate Faculty of the
Virginia Polytechnic Institute
in candidacy for the degree of
MASTER OF SCIENCE
in
Statistics

October 1960

Blacksburg, Virginia

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION AND BACKGROUND.	4
1.1 Chemical Background.	4
1.2 Prony's Method	6
1.3 Hartley's Method of 'Internal Least Squares'	7
II. FIRST ORDER REACTIONS.	10
2.1 Estimation of Parameters	10
2.2 Tests of Significance.	12
2.3 Numerical Example.	13
2.3a Maximum Likelihood Estimate	19
III. SECOND ORDER REACTIONS	33
3.1 Estimation of Parameters	33
3.2 Test for Irreversibility	34
3.3 Numerical Example.	35
3.4 Analysis for Any Reaction of Integral Order.	38
IV. CONSECUTIVE REACTIONS.	45
4.1 Disregarding Reverse Reactions	45
4.1a Numerical Example	47
4.2 Including Reverse Reactions.	49
4.2a Two Numerical Examples.	57
V. SUMMARY AND CONCLUSIONS.	65
VI. BIBLIOGRAPHY	67



CHAPTER	PAGE
VII. ACKNOWLEDGEMENTS	68
VIII. VITA	69

I. INTRODUCTION AND BACKGROUND

1.1 Chemical Background

The rate of a chemical reaction is defined to be the number of units of mass of some substance which is formed or transformed per unit of time per unit of volume. Especially in the field of organic chemistry where many reactions take place at a conveniently measurable rate, and where several possible reactions are often competing for available material, the study and prediction of reaction rates is a matter of practical importance as well as one of theoretical interest.

The classical theory of chemical kinetics has led essentially to the development of a wholly experimental science in which reaction rate measurements are made and interpreted in terms of the various mathematical formulations resulting from the theory. By the law of mass action the rate of a chemical reaction is proportional to the concentration of each of the reacting substances, and since these reactants are being continually consumed, their concentration, and therefore the rate of reaction must both vary with elapsed time. In actual magnitude the rate will also depend upon the temperature, pressure, presence of catalysts, and properties of the reacting substances. It is no simple matter to state

the general rate of a given reaction; and so in practice the rate is considered at a specific instant, and by mathematical means it is possible to obtain certain constants that are characteristic of the combined influences of the reaction at each temperature. It is this so-called specific reaction rate constant or velocity constant which will subsequently be estimated and subjected to statistical testing.

Many chemical reactions are reversible to a measurable extent; i.e., the products of a reaction are capable of again reacting to form the original reactants. According to some theories all reactions are reversible, and the so-called irreversible reaction is only a special case in which the reverse rate is so small as to be undetectable, and its rate constant may be considered zero. The reversible reaction ultimately finds a point of dynamic equilibrium where the forward and reverse rates are equal and opposite. A change in external conditions will favor the reaction in one direction or the other and thus cause a shift in the point of equilibrium. A knowledge of the magnitude of rate constants is desirable so that one may control the conditions necessary to yield a maximum in the desired end product, and suppress undesired reactions.

1.2 Prony's Method

The estimation of parameters in reaction equations can, theoretically, be done by the method of maximum likelihood in each case; but even for the simplest cases, the numerical work is very complicated. An approximate method, attributed to Prony (7, 10), implies that the differential quotient found in equations describing many physical phenomena may be replaced by a difference quotient.

Prony proposed that a function $k(x)$, specified by a table of numerical values, may be represented approximately by a sum of exponentials, e.g.

$$(1.2.1) \quad k(x) = Pe^{px} + Qe^{qx} + \dots + Ve^{vx}$$

where $P, Q, \dots, V, p, q, \dots, v$ are constants to be determined. The values of $k(x)$ are k_0, k_1, k_2, \dots , corresponding to values $0, W, 2W, \dots$, of an independent variable x . If $k(x)$ were represented exactly by a sum of μ exponentials, then $k(x)$ would satisfy a linear difference equation of the form

$$(1.2.2) \quad Ak_{n+\mu} + Bk_{n+\mu-1} + \dots + Mk_n = 0$$

where the roots of the algebraic equation

$$(1.2.3) \quad A\alpha^\mu + B\alpha^{\mu-1} + \dots + M = 0$$

would be e^{pw}, e^{qw}, \dots . Prony's method was to write down a set of linear equations

$$Ak_{\mu} + Bk_{\mu-1} + \dots + Mk_0 = 0$$

$$Ak_{\mu+1} + Bk_{\mu} + \dots + Mk_1 = 0$$

$$Ak_{\mu+2} + Bk_{\mu+1} + Ck_{\mu} + \dots + Mk_2 = 0$$

etc.,

for which values of A, B, ..., M could be found by the method of least squares; and with these values an equation of the type (1.2.3) was formed and solved for e^{pw} , e^{qw} , ..., e^{vw} . Since w is known to be the increment of x, p, q, ..., v are determined, and a set of linear equations could be formed to find the coefficients P, Q, ..., V.

This is the deterministic approach; however, we shall denote the error in gain of $k(x)$ in consecutive intervals by ϵ_i , and use the least squares approach to the problem in accordance with the assumption that the ϵ_i are independent and have the same variance.

1.3 Hartley's Method of "Internal Least Squares"

Another method for the estimation of parameters has been proposed by Hartley in view of the fact that many scientific laws are generated by simple linear relationships between the function and its first or second derivative, e.g., the exponential law of diminishing returns

$$(1.3.1) \quad y = \hat{y}(1 - e^{kx})$$

is generated by the first order differential equation

$$(1.3.2) \quad dy/dx = k\hat{y} + ky \quad .$$

Rather than fit a condition involving the differential coefficient to an empirical series of observed y_i , Hartley used the equivalent finite difference equation of the form

$$(1.3.3) \quad Y_i - Y_{i-1} = bY_i + a \quad .$$

Integrating (or summing) (1.3.3) gives the linear regression equation

$$(1.3.4) \quad y_j = bY_j + ax_j + c$$

where the progressive sum $Y_j = \sum_{i=0}^j y_i$ is an independent variable, and x_j is a second independent variable. Any such regression equation in which the dependent variable y is related to its own repeated sums, e.g., $Y_j = \sum_{i=0}^j y_i$, is termed internal regression. Estimates for a , b , c are obtained by applying least squares to the "integrated" equation (1.3.4) (see (4)).

This method though somewhat simpler than the maximum-likelihood method in the first order reaction, becomes likewise intractable when more complicated reactions are involved.

After a demonstration of the usefulness of Prony's approximation, by comparing it with the maximum likelihood

results in a first order reversible reaction, we will apply it to the more complicated reactions mentioned above.

II. FIRST ORDER REACTION

2.1 Estimation of Parameters

The order of a reaction indicates the power of concentration of reactants* to which the reaction rate is proportional. Consider the first order irreversible reaction $A \xrightarrow{k} B$, and the reversible reaction $A \xrightleftharpoons[k_2]{k_1} B$, first order in both directions. In the irreversible case the reaction rate is expressed by the differential equation

$$(2.1.1) \quad \text{Rate} = dx/dt = k(A_0 - x)$$

where x is the decrease in initial concentration of A at time t , A_0 is the initial concentration, and k is the rate constant. Similarly for the reversible case, the net rate is

$$(2.1.2) \quad dx/dt = k_1(A_0 - x) - k_2(B_0 + x) \quad ,$$

where k_1 is the rate constant for the forward reaction, k_2 that of the opposing reaction, and B_0 is the concentration of B at the beginning of the reaction. Assuming that the concentration of product B is initially zero, (2.1.2) then reduces to

$$(2.1.3) \quad dx/dt = k_1 A_0 - (k_1 + k_2)x \quad .$$

* In actual practice some modifications are required, and "activities" replace "concentrations", (9).

The simplest method of estimation of rate constants was proposed by Prony (10). It amounts to approximating the differential quotient dx/dt by a difference quotient $\Delta x/\Delta t$.

Equation (2.1.1) then becomes

$$(2.1.4) \quad (\Delta x/\Delta t)_i \approx k(A_0 - x_i) + \epsilon_i$$

or

$$(\Delta x/\Delta t)_i = y_i \approx \alpha + \beta x_i + \epsilon_i$$

where

$$\alpha = kA_0 \quad \text{and} \quad \beta = -k \quad .$$

Likewise, for the reversible process

$$(2.1.5) \quad (\Delta x/\Delta t)_i \approx k_1 A_0 - (k_1 + k_2)x_i + \epsilon_i$$

or

$$(\Delta x/\Delta t)_i = y_i \approx \alpha + \beta x_i + \epsilon_i$$

with $\alpha = k_1 A_0$ and $\beta = -(k_1 + k_2)$.

It is seen that both reversible and irreversible cases now have the form of a simple regression model.

Assuming that each y_i is an observation on a random variable Y_i which is normally distributed with constant variance σ^2 and mean $(\alpha + \beta x_i)$, the method of least squares may be used to estimate the parameters α and β . We minimize the expression

$$\sum \epsilon_i^2 = \sum (y_i - \alpha - \beta x_i)^2 \quad ;$$

and obtain the results

$$(2.1.6) \quad \hat{\beta} = (\Sigma x_i y_i - \frac{1}{n} \Sigma x_i \Sigma y_i) / (\Sigma x_i^2 - [\Sigma x_i]^2 / n)$$

and

$$(2.1.7) \quad \hat{\alpha} = \frac{1}{n} \Sigma y_i - \hat{\beta} \Sigma x_i / n ,$$

where n is the number of differences, (i.e., the number of observations minus one).

2.2 Tests of Significance

It is now possible to formulate and test hypotheses to determine whether a particular reaction follows a reversible or irreversible mechanism. For the irreversible case, $\beta = -k$ and $\alpha = A_0 k$, so a null hypothesis can be stated as

$$H_0: \beta = -\alpha/A_0 \text{ or } \alpha + A_0 \beta = 0 .$$

If the reaction is reversible, $\alpha = k_1 A_0$ and $\beta = -(k_1 + k_2)$, which implies that

$$A_0 \beta + \alpha = -A_0 k_2 .$$

Since the specific reaction rates are non-negative, the alternative to the above hypothesis is thus

$$\alpha + A_0 \beta < 0 ,$$

and $-\hat{\alpha}/A_0 - \hat{\beta}$ is an estimate of k_2 .

The appropriate test is a one-sided t-test with n-2 degrees of freedom and rejection region marked by the lower tail,

$$(2.2.1) \quad t_{n-2} = (\hat{\alpha} + A_0 \hat{\beta}) / s_{\hat{\alpha} + A_0 \hat{\beta}}$$

where s^2 is the estimate of the variance of $(\hat{\alpha} + A_0 \hat{\beta})$. It can be obtained more easily if we introduce the familiar notation

$$y_i = \mu + \beta(x_i - \bar{x}) + \epsilon_i$$

where \bar{x} is the arithmetic mean of the x_i , and rewrite the hypothesis as

$$H_0: \mu + \beta(A_0 - \bar{x}) = 0$$

versus the alternative

$$\mu + \beta(A_0 - \bar{x}) < 0 \quad .$$

Since $\hat{\mu}$ and $\hat{\beta}$ are uncorrelated,

$$\text{var}[\hat{\mu} + \hat{\beta}(A_0 - \bar{x})] = \sigma^2 \left[\frac{1}{n} + (A_0 - \bar{x})^2 / s_{xx} \right]$$

where $s_{xx} = \sum x_i^2 - (\sum x_i)^2 / n$, and hence $s_{\hat{\alpha} + A_0 \hat{\beta}}^2$, the square of the denominator in (2.2.1) is

$$s_{\hat{\alpha} + A_0 \hat{\beta}}^2 = s_e^2 \left[\frac{1}{n} + (A_0 - \bar{x})^2 / s_{xx} \right] \quad ,$$

where s_e^2 is the mean-square error.

2.3 Numerical Example

As an illustration we first calculate a demonstration study with $.1 \text{ min}^{-1}$ as a hypothetical value for k in the irreversible case, and set the initial concentration A_0 , equal

to one mole/liter. Solving the differential equation $dx/dt = k(A_0 - x_1)$; with the knowledge that when t is zero, x will also be zero, we obtain the exact result under this assumption,

$$x = A_0 [1 - e^{-kt}] \quad .$$

Values of x calculated for integral values of t from 1 through 20 are recorded in Table 2.1, and the curve of time versus x is plotted in Figure I as the lower solid line.

From Table 2.1 we see that for $t = 20$, $x = 0.864665$.

The solution of the reversible differential equation for x is

$$x = \frac{k_1 A_0}{k_1 + k_2} [1 - e^{-(k_1 + k_2)t}] \quad .$$

For a second demonstration we will now assume that a reversible reaction produced the same amount x , after twenty minutes as the aforementioned irreversible reaction. Thus, we must choose k_1 and k_2 in such a way that

$$\frac{k_1 A_0}{k_1 + k_2} [1 - e^{-(k_1 + k_2)20}] = 0.864665 \quad ;$$

so that after twenty minutes reaction time, in both cases the same amount of A_0 has disappeared. In each of our examples A_0 equals one mole/liter. With an arbitrary choice of k_1 as

0.15 min^{-1} , (k_1+k_2) will have to be 0.16738 and k_2 is thus 0.01738. Values of x for this irreversible case are recorded in Table 2.3 and the data are plotted as the upper solid line in Figure I.

A scattering of points which serve to demonstrate the presence of random errors about each of the two ideal curves is also shown in Figure I. These points typify experimental data which might be observed for each of the two reactions under study, and are recorded for the irreversible and reversible cases in Tables 2.2 and 2.4 respectively. Four sets of data are now available--two without and two with error about ideal curves. To demonstrate Prony's method in this case, $(\Delta x/\Delta t)_i$ was obtained by taking increments of concentration and time between successive observations. The corresponding x_i were taken as the means of two consecutive observations.

The following totals and parameter values were computed:

Irreversible theoretical data

$$\Sigma x_i = 11.346154 \qquad \Sigma x_i y_i = 0.373823$$

$$\Sigma y_i = 0.864665 \qquad \Sigma y_i^2 = 0.049043$$

$$\Sigma x_i^2 = 7.604814$$

$$\hat{\beta} = -0.099917 \qquad \hat{\alpha} = 0.099917$$

Irreversible "experimental" data

$$\begin{aligned}\Sigma x_i &= 11.407 & \Sigma x_i y_i &= 0.368255 \\ \Sigma y_i &= 0.858 & \Sigma y_i^2 &= 0.071712 \\ \Sigma x_i^2 &= 7.659193 \\ \hat{\beta} &= -0.105015 & \hat{\alpha} &= 0.102795\end{aligned}$$

Reversible theoretical data

$$\begin{aligned}\Sigma x_i &= 12.744695 & \Sigma x_i y_i &= 0.373809 \\ \Sigma y_i &= 0.86465 & \Sigma y_i^2 &= 0.066965 \\ \Sigma x_i^2 &= 9.182486 \\ \hat{\beta} &= -0.166970 & \hat{\alpha} &= 0.149629\end{aligned}$$

Reversible "experimental" data

$$\begin{aligned}\Sigma x_i &= 12.751 & \Sigma x_i y_i &= 0.312512 \\ \Sigma y_i &= 0.768 & \Sigma y_i^2 &= 0.071174 \\ \Sigma x_i^2 &= 9.159528 \\ \hat{\beta} &= -0.171946 & \hat{\alpha} &= 0.148024\end{aligned}$$

The values of $\hat{\alpha}$ and $\hat{\beta}$ obtained by the approximate method of difference quotients for the exact curves are very close to the exact values (.0999 instead of .1000 in the irreversible case; .1496 instead of .1500 and .1670 instead of

.1674 in the irreversible case). It is thus seen that the substitution of difference for differential quotients does not seriously affect results even if they extend, as in the present example, over a wide range of the reaction (considerably beyond half-life).

Next it will be illustrated how the somewhat artificial assumption of errors with common variance in the differential equation (or in the difference equation) affects the results of the scattered data which were obtained under the assumption that approximately normally distributed errors occur in the measurement of yields. Before this study is made we will formally test the hypothesis of irreversibility in each case.

For the irreversible exact data

$$t_{18} = 0$$

and as expected we accept the hypothesis of an irreversible reaction. From the corresponding "experimental" situation

$$t_{18} = -0.00222/0.01608 = -0.138$$

and again the hypothesis of a first order irreversible reaction is accepted.

In the reversible example involving "experimental" data about the curve

$$t_{18} = -0.023922/0.010522 = -2.274 \quad .$$

From a table of values for the t distribution with 18 d.f. we reject at the .05 level if $t < -1.734$, and thus reject with 95% confidence, (and correctly in this case) that this data was obtained from an irreversible reaction.

In Table 2.5 and Figure II, the theoretical curves (solid lines) are compared with the curves estimated from the scattered points (broken lines) by Prony's method. The equation of the broken lines is given by the differential equation

$$\begin{aligned} dx/dt &= \alpha + \beta x \quad , \\ 1/\beta \log(\alpha + \beta x) &= t + \log c \quad , \\ x &= 1/\beta (ce^{\beta t} - \alpha) \quad ; \end{aligned}$$

and since at time $t = 0$, $x = 0$, we have $c = \alpha$, and $x = \alpha/\beta(e^{\beta t} - 1)$; and the equation becomes

$$(2.3.1) \quad x = -\hat{\alpha}/\hat{\beta}(1 - e^{\hat{\beta}t}) \quad .$$

If it should happen that $\hat{\beta}$ is equal to zero, this would be an indication that the reaction cannot be of the first order but must be of zero order, i.e., the rate does not depend upon the concentration of the reactant.

By equation (2.3.1) x values for each set of $\hat{\alpha}$ and $\hat{\beta}$ are computed and recorded in Table 2.5, and the four resulting curves are shown in Figure II. In comparing the solid and

dotted curves of Figure II it is seen that only after the period of half life has been passed that the estimated curve based on "experimental" data in a reversible reaction begins to vary noticeably from that of the exact data.

2.3a Maximum Likelihood Estimate

In order to study the question of whether this discrepancy is due to sampling effect or the crudeness of the Prony approximation, we shall apply the more respectable, but very complicated, method of maximum likelihood to the "experimental" data for the reversible reaction. As before

$$x = -\alpha/\beta(1 - e^{\beta t})$$

and

$$x_i = -\alpha/\beta(1 - e^{\beta t_i})(1 + \epsilon_i)$$

where $\text{var}(\epsilon_i) = \sigma^2$, and the ϵ_i are assumed to have a normal distribution. Taking logarithms we have

$$(2.3.2) \quad \log x_i = \log \alpha - \log(-\beta) + \log(1 - e^{\beta t_i}) + \log(1 + \epsilon_i).$$

However,

$$\log(1 + \epsilon_i) = \epsilon_i - \epsilon_i^2/2 + \epsilon_i^3/3 - \epsilon_i^4/4 + \dots,$$

and upon making the assumption that the error ϵ_i will be small, $\log(1 + \epsilon_i)$ is replaced by ϵ_i . Then,

$$\log x_i = \log \alpha - \log(-\beta) + \log(1 - e^{\beta t_i}) + \epsilon_i.$$

This is the conventional and well-known analysis of data in

which the error is assumed to be a percentage of the yield. Since $\text{var}(\epsilon_i)$ is assumed to be the same for each ϵ_i , we may apply least squares to this formula in order to obtain maximum likelihood estimates of α and β under the above-mentioned assumptions. However, the resulting transcendental equations are very difficult to handle, and we will, instead, attempt a series expansion. Note that

$$\begin{aligned} 1 - e^{\beta t_i} &= e^{\frac{1}{2}\beta t_i} [e^{-\frac{1}{2}\beta t_i} - e^{\frac{1}{2}\beta t_i}] \\ &= 2e^{\frac{1}{2}\beta t_i} \sinh(-\beta t_i/2) \end{aligned}$$

Hence

$$\log(1 - e^{\beta t_i}) = \log 2 + \frac{1}{2}\beta t_i + \log \sinh(-\beta t_i/2)$$

However,

$$\begin{aligned} \log \sinh(-\beta t_i/2) &= \log(-\beta t_i) - \log 2 + \frac{1}{6}(\beta t_i/2)^2 - \frac{1}{180}(\beta t_i/2)^4 \\ &\quad + \frac{1}{2835}(\beta t_i/2)^6 + \dots \end{aligned}$$

These first few terms of this series are deemed sufficient since $(\beta t_i/2)$ is generally a small quantity, and the coefficients of the log sinh expansion diminish very rapidly.

Making use of this, we have

$$\begin{aligned} \log x_i &= \log \alpha - \log(-\beta) + \log 2 + \beta t_i/2 + \log(-\beta) + \log t_i \\ &\quad - \log 2 + \frac{1}{6}(\beta t_i/2)^2 - \frac{1}{180}(\beta t_i/2)^4 + \frac{1}{2835}(\beta t_i/2)^6 + \epsilon_i \\ &= \log \alpha + \log t_i + \beta t_i/2 + \frac{1}{6}(\beta t_i/2)^2 - \frac{1}{180}(\beta t_i/2)^4 \\ &\quad + \frac{1}{2835}(\beta t_i/2)^6 + \epsilon_i \end{aligned}$$

Now let $\log x_i - \log t_i = y_i$ and $\log \alpha = \mu$; then

$$(2.3.3) \quad y_i = \mu + (t_i/2)\beta + \frac{1}{6}(t_i/2)^2 \beta^2 - \frac{1}{180}(t_i/2)^4 \beta^4 + \frac{1}{2835}(t_i/2)^6 \beta^6 + \epsilon_i.$$

Least squares estimates of β are now obtained as follows:

$$\Sigma \epsilon_i^2 = \Sigma [y_i - \mu - (t_i/2)\beta - \frac{1}{6}(t_i/2)^2 \beta^2 + \frac{1}{180}(t_i/2)^4 \beta^4 - \frac{1}{2835}(t_i/2)^6 \beta^6]^2$$

$$\partial \Sigma \epsilon_i^2 / \partial \mu = -2 \Sigma [y_i - \mu - (t_i/2)\beta - \frac{1}{6}(t_i/2)^2 \beta^2 + \frac{1}{180}(t_i/2)^4 \beta^4 - \frac{1}{2835}(t_i/2)^6 \beta^6].$$

$$(2.3.4) \quad \hat{\mu} = \bar{y} - \frac{1}{n} \Sigma (t_i/2) \hat{\beta} - \frac{1}{6n} \Sigma (t_i/2)^2 \hat{\beta}^2 + \frac{1}{180n} \Sigma (t_i/2)^4 \hat{\beta}^4 - \frac{1}{2835n} \Sigma (t_i/2)^6 \hat{\beta}^6.$$

$$\partial \Sigma \epsilon_i^2 / \partial \beta = 2 \Sigma [y_i - \mu - (t_i/2)\beta - \frac{1}{6}(t_i/2)^2 \beta^2 + \frac{1}{180}(t_i/2)^4 \beta^4 - \frac{1}{2835}(t_i/2)^6 \beta^6]$$

$$\times [-(t_i/2) - \frac{1}{3}(t_i/2)^2 \beta + \frac{1}{45}(t_i/2)^4 \beta^3 - \frac{2}{945}(t_i/2)^6 \beta^5].$$

$$(2.3.5) \quad \Sigma (y_i - \hat{\mu})(t_i/2) = [-\frac{1}{3} \Sigma (y_i - \hat{\mu})(t_i/2)^2 + \Sigma (t_i/2)^2] \hat{\beta} + \frac{1}{2} \Sigma (t_i/2)^3 \hat{\beta}^2 + [\frac{1}{45} \Sigma (y_i - \hat{\mu})(t_i/2)^4 + \frac{1}{18} \Sigma (t_i/2)^4] \hat{\beta}^3 - \frac{1}{36} \Sigma (t_i/2)^5 \hat{\beta}^4 - [\frac{2}{945} \Sigma (y_i - \hat{\mu})(t_i/2)^6 + \frac{1}{180} \Sigma (t_i/2)^6] \hat{\beta}^5 + \frac{1}{405} \Sigma (t_i/2)^7 \hat{\beta}^6 + \frac{101}{170,100} \Sigma (t_i/2)^8 \hat{\beta}^7 - \frac{1}{51,030} \Sigma (t_i/2)^{10} \hat{\beta}^9 + \frac{2}{2,679,075} \Sigma (t_i/2)^{12} \hat{\beta}^{11}.$$

Substituting $\hat{\mu}$ from (2.3.4) into (2.3.5) we have

$$\begin{aligned}
 (2.3.6) \quad & -\Sigma Y_i(t_i/2) + \frac{\Sigma Y_i \Sigma(t_i/2)}{n} \\
 & + [\Sigma(t_i/2)^2 - \frac{[\Sigma(t_i/2)]^2}{n} - \frac{1}{3} \Sigma Y_i(t_i/2)^2 + \frac{1}{3} \frac{\Sigma Y_i \Sigma(t_i/2)^2}{n}] \hat{\beta} \\
 & + \frac{1}{2} [\Sigma(t_i/2)^3 - \frac{\Sigma(t_i/2) \Sigma(t_i/2)^2}{n}] \hat{\beta}^2 + [\frac{1}{45} \{\Sigma Y_i(t_i/2)^4 \\
 & \quad - \frac{\Sigma Y_i \Sigma(t_i/2)^4}{n}\} + \frac{1}{18} \{\Sigma(t_i/2)^4 - \frac{[\Sigma(t_i/2)^2]^2}{n}\}] \hat{\beta}^3 \\
 & + \frac{1}{36} [\frac{\Sigma(t_i/2) \Sigma(t_i/2)^4}{n} - \Sigma(t_i/2)^5] \hat{\beta}^4 \\
 & \quad + [\frac{-2}{945} \{\Sigma Y_i(t_i/2)^6 - \frac{\Sigma Y_i \Sigma(t_i/2)^6}{n}\} \\
 & - \frac{1}{180} \{\Sigma(t_i/2)^6 - \frac{\Sigma(t_i/2)^2 \Sigma(t_i/2)^4}{n}\}] \hat{\beta}^5 \\
 & \quad + \frac{1}{405} [\Sigma(t_i/2)^7 - \frac{\Sigma(t_i/2) \Sigma(t_i/2)^6}{n}] \hat{\beta}^6 \\
 & + [\frac{101}{170,100} \Sigma(t_i/2)^8 - \frac{4}{8505} \frac{\Sigma(t_i/2)^2 \Sigma(t_i/2)^6}{n} \\
 & \quad - \frac{1}{8100} \frac{[\Sigma(t_i/2)^4]^2}{n}] \hat{\beta}^7 \\
 & - \frac{1}{51,030} [\Sigma(t_i/2)^{10} - \frac{\Sigma(t_i/2)^6 \Sigma(t_i/2)^4}{n}] \hat{\beta}^9 \\
 & \quad + \frac{2}{2,679,075} [\Sigma(t_i/2)^{12} - \frac{[\Sigma(t_i/2)^6]^2}{n}] \hat{\beta}^{11} = 0 .
 \end{aligned}$$

By Newton's method only one root of (2.3.6) was found to be close to the true value of $-(k_1+k_2)$; it is $\hat{\beta} = -.2000$, and from (2.3.4) $\mu = -1.7844$ with antilog $\hat{\alpha} = 0.1679$.

The estimates of k_1 and k_2 which are respectively 0.1679 and 0.0321 as found by maximum likelihood are in this instance not as close to the original values as those computed by the previous method. This is a probable indication that the effect due to sampling or measurement error in the rate data is greater than that introduced by a particular method of estimation.

Values for x as defined by this latter set of parameter estimates are recorded in Table 2.6 and the corresponding curve plotted in Figure III as the broken line. The exact curve from Figure I is shown as a solid line in Figure III for comparison.

TABLE 2.1

Theoretical data for irreversible reaction

t(min.)	x	$(\Delta x/\Delta t)_i$	x_i
0	0.000000	.095163	.047582
1	.095163	.086106	.138216
2	.181269	.077913	.220226
3	.259182	.070498	.294431
4	.329680	.063789	.361575
5	.393469	.057719	.422329
6	.451188	.052227	.477302
7	.503415	.047256	.527043
8	.550671	.042759	.572051
9	.593430	.038691	.612776
10	.632121	.035008	.649625
11	.667129	.031677	.682968
12	.698806	.028662	.713137
13	.727468	.025935	.740436
14	.753403	.023467	.765137
15	.776870	.021233	.787487
16	.798103	.019213	.807710
17	.817316	.017385	.826009
18	.834701	.015730	.842566
19	.850431	.014234	.857548
20	.864665		

TABLE 2.2

"Experimental" data for irreversible reaction

t(min.)	x	$(\Delta x/\Delta t)_i$	x_i
0	0.000	.110	.055
1	.110	.042	.131
2	.152	.153	.229
3	.305	.037	.324
4	.342	.034	.359
5	.376	.092	.422
6	.468	.045	.491
7	.513	.009	.518
8	.522	.058	.551
9	.580	.098	.629
10	.678	.008	.682
11	.686	.002	.687
12	.688	.025	.701
13	.713	.019	.723
14	.732	.073	.769
15	.805	-.015	.798
16	.790	.038	.809
17	.828	.012	.834
18	.840	.006	.843
19	.846	.012	.852
20	.858		

TABLE 2.3

Theoretical data for reversible reaction

t(min.)	x	$(\Delta x/\Delta t)_i$	x_i
0	0.00000	.13812	.06906
1	.13812	.11682	.19653
2	.25494	.09884	.30436
3	.35378	.08359	.39558
4	.43737	.07066	.47270
5	.50803	.05986	.53796
6	.56789	.05059	.59318
7	.61848	.04226	.63961
8	.66074	.03674	.67911
9	.69748	.03062	.71279
10	.72810	.02590	.74105
11	.75400	.02191	.76496
12	.77591	.01854	.78518
13	.79445	.01567	.80228
14	.81012	.01326	.81675
15	.82338	.01121	.82898
16	.83459	.00949	.83924
17	.84408	.00803	.84810
18	.85211	.00679	.85550
19	.85890	.00575	.86178
20	.86465		

TABLE 2.4

"Experimental" data for reversible reaction

t(min.)	x	$(\Delta x/\Delta t)_i$	x_i
0	0.000	.151	.076
1	.151	.125	.214
2	.276	.067	.310
3	.343	.072	.379
4	.415	.120	.475
5	.535	.043	.556
6	.578	.008	.582
7	.586	.012	.642
8	.698	.009	.702
9	.707	.027	.720
10	.734	.003	.736
11	.737	.026	.750
12	.763	.054	.790
13	.817	-.020	.807
14	.797	-.002	.796
15	.795	.028	.809
16	.823	.029	.838
17	.852	.003	.854
18	.855	.000	.855
19	.855	.013	.862
20	.868		

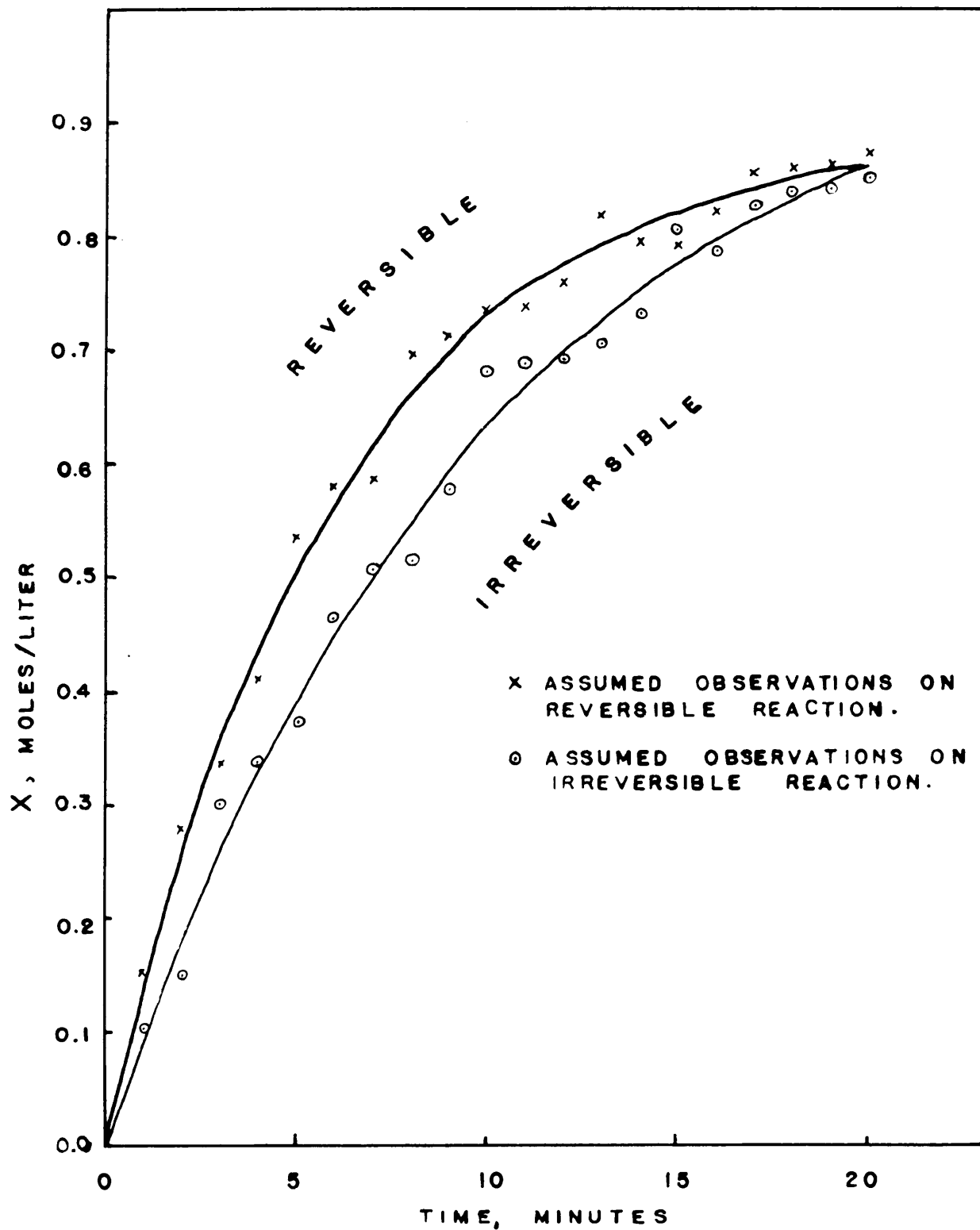


FIGURE I. IDEAL FIRST-ORDER REACTIONS.

TABLE 2.5

Values of $x = -a/b[1 - e^{-bt}]$ as calculated

from constants obtained by approximate method

t	x, Theoretical irreversible	x, "Experimental" irreversible	x, Theoretical reversible	x, "Experimental" reversible
0	0.00000	0.00000	0.00000	0.00000
1	.09509	.09759	.13780	.13599
2	.18114	.18544	.25441	.25052
3	.25900	.26453	.35309	.34692
4	.32946	.33575	.43661	.42812
5	.39322	.39985	.50727	.49649
6	.45091	.45758	.56706	.55405
7	.50311	.50954	.61757	.60252
8	.55037	.55633	.66049	.64333
9	.59313	.59845	.69672	.67770
10	.63181	.63637	.72748	.70663
11	.66686	.67051	.75334	.73101
12	.69850	.70126	.77530	.75152
13	.72717	.72893	.79388	.76888
14	.75311	.75381	.80961	.78334
15	.77659	.77628	.82291	.79559
16	.79783	.79647	.83419	.80590
17	.81705	.81466	.84370	.81459
18	.83445	.83102	.85176	.82189
19	.85019	.84576	.85860	.82806
20	.86444	.85903	.86437	.83324

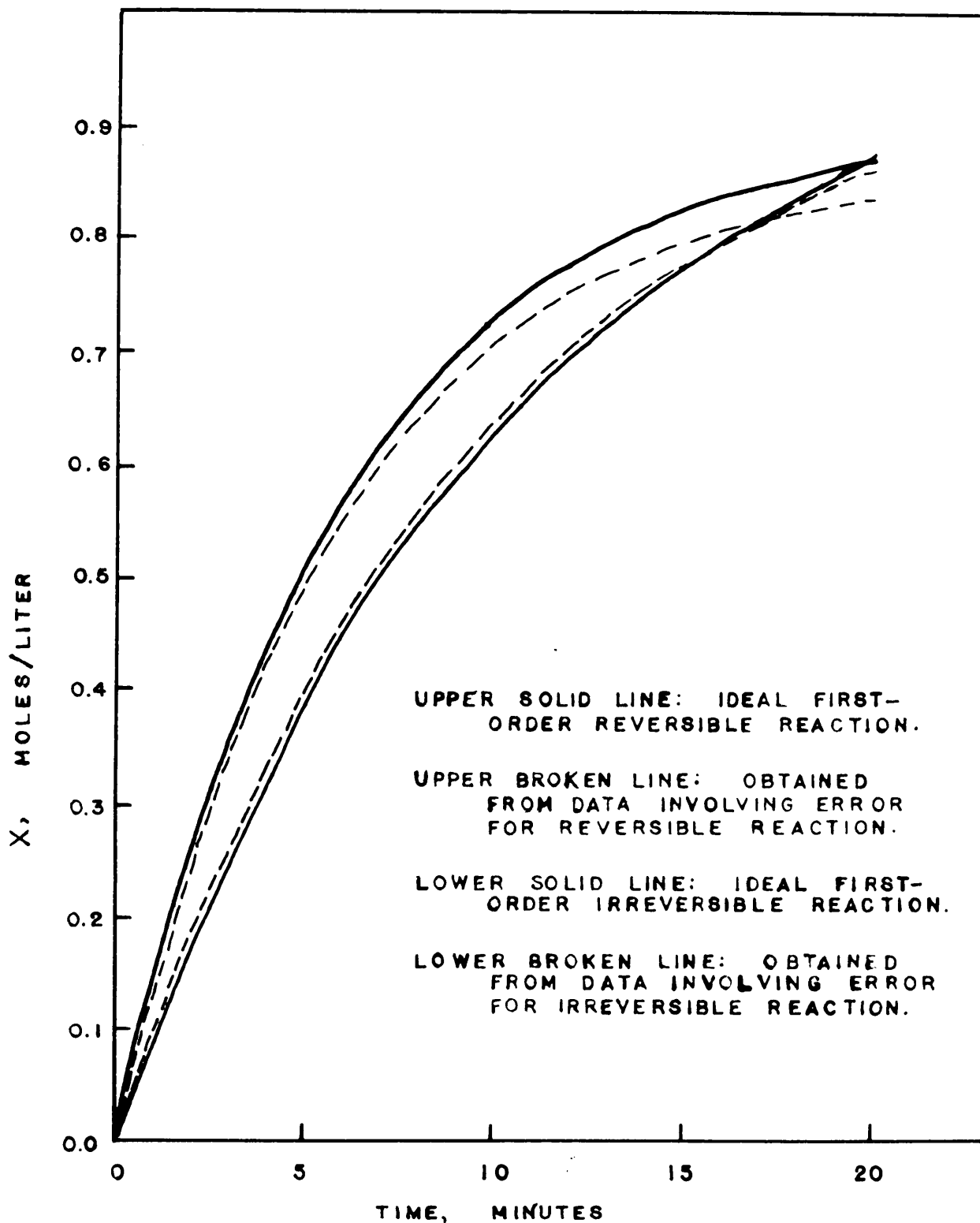


FIGURE II. CURVES OBTAINED BY PRONY'S
METHOD OF APPROXIMATION.

TABLE 2.6

Values of $x = -a/b[1 - e^{-bt}]$ as calculated
from constants obtained by maximum likelihood

<u>t</u>	<u>x</u>
0	0
1	.15214
2	.27678
3	.37875
4	.46229
5	.53066
6	.58664
7	.63248
8	.67001
9	.70073
10	.72847
11	.74648
12	.76334
13	.77700
14	.78845
15	.79770
16	.80509
17	.81149
18	.81656
19	.82072
20	.82412

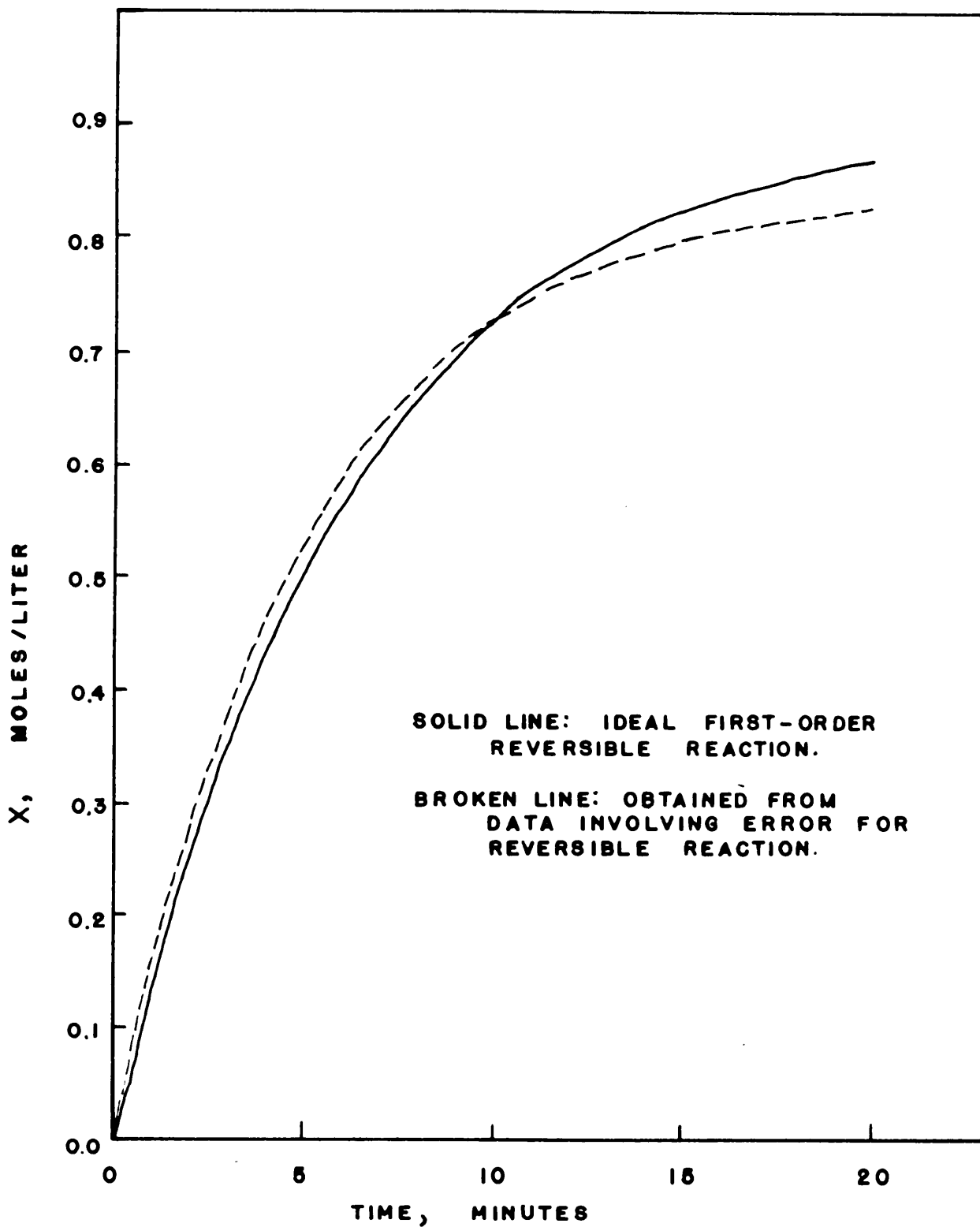


FIGURE III. CURVE OBTAINED BY MAXIMUM LIKELIHOOD METHOD.

III. SECOND ORDER REACTIONS

3.1 Estimation of Parameters

The study presented in Chapter II of this thesis seems to indicate that the very simple Prony method of estimating constants in the rate equation may produce some useful results even though the assumption of a homoscedastic error in the differential equation would appear rather arbitrary. Even for a simple first-order reversible reaction the method of maximum likelihood is very difficult to apply so that, for more complicated processes, we will restrict ourselves to the estimation of parameters by the simple Prony method only, and formulate hypotheses accordingly.

Consider a situation in which a second order reaction is opposed by one of the first order; e.g.



The rate equation is expressed by

$$(3.1.1) \quad dx/dt = k_1 (A_0 - x) (B_0 - x) - k_2 (C_0 + x) .$$

When $A_0 = B_0$, and $C_0 = 0$; equation (3.1.1) becomes

$$(3.1.2) \quad dx/dt = k_1 (A_0 - x)^2 - k_2 x$$

which is approximated by

$$(3.1.3) \quad (\Delta x / \Delta t)_i = y_i * k_1 (A_0 - x_i)^2 - k_2 x_i + \epsilon_i .$$

Again the assumption is made that the y_i are independent, normally distributed and have a common variance. Using the method of least squares and conventional matrix notation we have

$$\begin{bmatrix} \Sigma(A_0 - x_i)^4 & \Sigma x_i (A_0 - x_i)^2 \\ \Sigma x_i (A_0 - x_i)^2 & \Sigma x_i^2 \end{bmatrix} \begin{bmatrix} \hat{k}_1 \\ \hat{k}_2^* \end{bmatrix} = \begin{bmatrix} \Sigma y_i (A_0 - x_i)^2 \\ \Sigma x_i y_i \end{bmatrix}$$

where k_2^* is defined as $-k_2$.

3.2 Test for Irreversibility

The irreversible reaction corresponding to this situation has the rate equation

$$(3.2.1) \quad dx/dt = k_1 (A_0 - x)^2 \quad .$$

Therefore to test for irreversibility it is sufficient to test the null hypothesis

$$H_0: k_2^* = 0$$

against the alternative

$$H_1: k_2^* < 0 \quad ;$$

remembering that k_2^* is $-k_2$.

$$\text{var } \hat{k}_2^* = \frac{\Sigma(A_0 - x_i)^4 \sigma^2}{\Sigma x_i^2 \Sigma(A_0 - x_i)^4 - [\Sigma x_i (A_0 - x_i)^2]^2}$$

In the t-test σ^2 will be estimated by s^2 , the sum of squares of the errors of estimation, which is obtained by

$$s^2 = \frac{1}{n-2} [\sum Y_i^2 - \hat{k}_1 \sum Y_i (A_0 - x_i)^2 - \hat{k}_2^* \sum x_i y_i]$$

As before the test indicated is a one-tailed t-test with n-2 degrees of freedom

$$t_{n-2} = \hat{k}_2^* / \sqrt{\frac{s^2 \sum (A_0 - x_i)^4}{\sum x_i^2 \sum (A_0 - x_i)^4 - [\sum x_i (A_0 - x_i)^2]^2}}$$

and we reject if the observed $t < t_{n-2, \alpha}^{**}$

3.3 Numerical Example

The data for this illustration as shown in Table 3.1 was derived from a problem ((7), p. 58) in which the reaction is believed to be second order in the forward direction and first order in reverse. Initial concentrations of the two reactants were each 1.2 lb. mole/cu. ft., and there was no product present at the beginning of the reaction. The rate constant for the forward reaction was 1.5 mole⁻¹ cu. ft. min⁻¹. The net rate of reaction is zero when x, the amount of each reactant gone, is 0.80; i.e., at x = 0.80 the system has reached dynamic equilibrium. In order to compare this with an irreversible reaction, which was not assumed in the

** $t_{n-2, \alpha}^{**}$ denotes the α 'th percentage point on the t-distribution, i.e., the lower tail which is, in general, a negative number.

recorded data, we took the point $x = 0.75$ which is attained after 1.5 minutes, and introduced an irreversible reaction for which

$$x = A_{\infty} - A_{\infty} / (A_{\infty} kt + 1) \quad ; \quad (\text{from (3.2.1)})$$

and k has the value 0.932704. Data for this irreversible reaction are in Table 3.2.

Both reversible (upper line) and irreversible (lower line) curves are shown in Figure IV; and as in the previous example scattered points indicating experimental error are also shown about each curve and recorded in Tables 3.3 and 3.4.

The totals, parameter estimates, and tests for mechanism reversibility are as follows:

Reversible theoretical data

$$\Sigma x_i^2 = 3.791250 \quad \Sigma (A_{\infty} - x_i)^4 = 9.993646 \quad \Sigma (A_{\infty} - x_i)^2 y_i = 13.979280$$

$$\Sigma y_i^2 = 19.803622 \quad \Sigma x_i y_i = 3.777288 \quad \Sigma (A_{\infty} - x_i)^2 x_i = 3.285452$$

$$\hat{k}_1 = 1.498063 \quad \hat{k}_2^* = -0.301886 = -\hat{k}_2$$

$$\text{Estimated var}(\hat{k}_2^*) = 0.00004206 \quad **$$

$$t = -0.301886 / 0.006485 = -46.551$$

** With this small error, a t-test would not usually be made.

Since the hypothesis of irreversibility is rejected at the .05 level for $t < -1.753$ we must very definitely reject in this case, and assume that the reaction is reversible, in accordance with the underlying model. Actually, the mean-square due to error is here due to error of approximation only, and as is seen above ($\text{var } \hat{k}_2^*$), it is remarkably small.

Reversible "experimental" data

$$\Sigma x_i^2 = 5.265880 \quad \Sigma (A_{\circ} - x_i)^4 = 6.506723 \quad \Sigma (A_{\circ} - x_i)^2 y_i = 9.057205$$

$$\Sigma y_i^2 = 14.5584 \quad \Sigma x_i y_i = 3.343360 \quad \Sigma (A_{\circ} - x_i)^2 x_i = 3.298120$$

$$\hat{k}_1 = 1.567917 \quad \hat{k}_2^* = -0.347106$$

$$\text{var}(\hat{k}_2^*) = 0.028156$$

$$t = -.347100 / .167798 = -2.068594$$

Again we reject that this data is from an irreversible reaction and correctly decide in favor of a reversible reaction even though, as Figure IV indicates, the errors are quite considerable, and the two reaction equations are quite close.

Irreversible theoretical data

$$\Sigma x_i^2 = 4.332294 \quad \Sigma (A_{\circ} - x_i)^4 = 8.666490 \quad \Sigma (A_{\circ} - x_i)^2 y_i = 8.076136$$

$$\Sigma y_i^2 = 7.526000 \quad \Sigma x_i y_i = 3.068968 \quad \Sigma (A_{\circ} - x_i)^2 x_i = 3.293768$$

$$\hat{k}_1 = 0.931934 \quad \hat{k}_2^* = -0.000140$$

$$\text{var}(\hat{k}_2^*) = 6.4925 \times 10^{-8}$$

$$t = -.000140 / .000254 = -0.5512$$

Therefore the hypothesis of irreversibility is accepted.

As before we see that the error of approximation is very small for theoretical data.

Irreversible "experimental" data

$$\Sigma x_i^2 = 4.37817375 \quad \Sigma (A_{\odot} - x_i)^4 = 8.509683 \quad \Sigma (A_{\odot} - x_i)^2 x_i = 3.309282$$

$$\Sigma y_i^2 = 9.6357 \quad \Sigma x_i y_i = 3.128605 \quad \Sigma (A_{\odot} - x_i)^2 y_i = 8.237865$$

$$\hat{k}_1 = .977489 \quad \hat{k}_2^* = -.024253$$

$$\text{var}(\hat{k}_2^*) = .03578167$$

$$t = -.024253 / .189161 = -.1282$$

Again we accept an irreversible reaction.

3.4 Analysis for Any Reaction of Integral Order

The theory of this chapter can be extended to cover reactions of any integral order in either direction by adjusting the powers of $(A_{\odot} - x)$ and x .

$$(3.4.1) \quad dx/dt = k_1 (A_{\odot} - x)^p - k_2 x^q$$

where p is the order of the forward reaction and q that of the reverse reaction. Equation (3.4.1) is approximated by

$$(\Delta x / \Delta t)_i = k_1 (A_{\odot} - x_i)^p - k_2 x_i^q + \epsilon_i$$

and by the method of least squares we have

$$\begin{bmatrix} \Sigma(A_0 - x_i)^{2p} & \Sigma x_i^q (A_0 - x_i)^p \\ \Sigma x_i^q (A_0 - x_i)^p & \Sigma x_i^{2q} \end{bmatrix} \begin{bmatrix} \hat{k}_1 \\ \hat{k}_2^* \end{bmatrix} = \begin{bmatrix} \Sigma(A_0 - x_i)^p y_i \\ \Sigma x_i^q y_i \end{bmatrix}$$

where \hat{k}_2^* is equal to $-\hat{k}_2$.

As before the hypothesis to be tested for irreversibility is $k_2^* = 0$ vs. $k_2^* < 0$.

$$\text{var}(\hat{k}_2^*) = \frac{\Sigma(A_0 - x_i)^{2p} \sigma^2}{\Sigma x_i^{2q} \Sigma(A_0 - x_i)^{2p} - [\Sigma x_i^q (A_0 - x_i)^p]^2} = c\sigma^2, \text{ say;}$$

where σ^2 is estimated by s^2 obtained as

$$s^2 = \frac{1}{n-2} [\Sigma y_i^2 - \hat{k}_1 \Sigma y_i (A_0 - x_i)^p - \hat{k}_2^* \Sigma x_i^q y_i]$$

The test statistic for a one-tailed t-test in this general case is

$$t_{n-2} = \hat{k}_2^* / s\sqrt{c}$$

and we reject for $t < t_{n-2, \alpha}$.

TABLE 3.1

Theoretical data for reversible reaction

t(min.)	x	$(\Delta x/\Delta t)_i$	x_i	$(A_0 - x_i)$
0.0000	0.00	2.0661	.025	1.175
.0242	.05	1.8727	.075	1.125
.0509	.10	1.6949	.125	1.075
.0804	.15	1.5244	.175	1.025
.1132	.20	1.3550	.225	.975
.1501	.25	1.1792	.275	.925
.1925	.30	1.0638	.325	.875
.2394	.35	.9058	.375	.825
.2946	.40	.7418	.425	.775
.3620	.45	.6676	.475	.725
.4369	.50	.5241	.525	.675
.5323	.55	.4112	.575	.625
.6539	.60	.3060	.625	.575
.8173	.65	.2339	.6625	.5375
.9242	.675	.1868	.6875	.5125
1.0580	.700	.1417	.7125	.4875
1.2344	.725	.0982	.7375	.4625
1.4891	.750			

TABLE 3.2

Theoretical data for irreversible reaction

t(min.)	x	$(\Delta x/\Delta t)_i$	x_i	$(A_{\infty} - x_i)$
0.00	0.000	1.272	.0318	1.1682
.05	.0636	1.144	.0922	1.1078
.10	.1208	1.034	.1466	1.0534
.15	.1725	.940	.1960	1.0040
.20	.2195	.821	.2606	.9394
.30	.3016	.695	.3364	.8636
.40	.3711	.595	.4008	.7992
.50	.4306	.515	.4564	.7436
.60	.4821	.451	.5046	.6954
.70	.5272	.397	.5470	.6530
.80	.5669	.353	.5846	.6154
.90	.6022	.316	.6180	.5820
1.00	.6338	.284	.6480	.5520
1.10	.6622	.256	.6750	.5250
1.20	.6878	.234	.6995	.5005
1.30	.7112	.213	.7218	.4782
1.40	.7325	.196	.7412	.4588
1.489	.7500			

TABLE 3.3

"Experimental" data for reversible reaction

t(min.)	x	$(\Delta x/\Delta t)_i$	x_i	$(A_0 - x_i)$
0.00	0.000	2.36	.059	1.141
.05	.118	.98	.142	1.058
.10	.167	1.66	.208	.992
.15	.250	1.72	.293	.907
.20	.336	.74	.373	.827
.30	.410	.39	.430	.770
.40	.449	.74	.481	.719
.50	.523	.82	.564	.636
.60	.605	.15	.612	.588
.70	.620	.18	.629	.571
.80	.638	.09	.642	.558
.90	.647	.20	.657	.543
1.00	.667	.52	.693	.507
1.10	.719	.09	.724	.476
1.20	.728	.02	.729	.471
1.30	.730	.08	.734	.466
1.40	.738	.08	.742	.458
1.50	.746			

TABLE 3.4

"Experimental" data for irreversible reaction

$t(\text{min.})$	x	$(\Delta x/\Delta t)_i$	x_i	$(A_0 - x_i)$
0.00	0.000	1.62	.040	1.160
.05	.081	1.16	.110	1.090
.10	.139	.38	.148	1.052
.15	.158	1.56	.197	1.003
.20	.236	.37	.254	.946
.30	.273	1.03	.324	.876
.40	.376	.76	.414	.786
.50	.452	.36	.470	.730
.60	.488	.17	.496	.704
.70	.505	.54	.532	.668
.80	.559	.66	.592	.608
.90	.625	.22	.636	.564
1.00	.647	.05	.650	.550
1.10	.652	.52	.678	.525
1.20	.704	.13	.710	.490
1.30	.717	.10	.722	.478
1.40	.727	.28	.741	.459
1.50	.755			

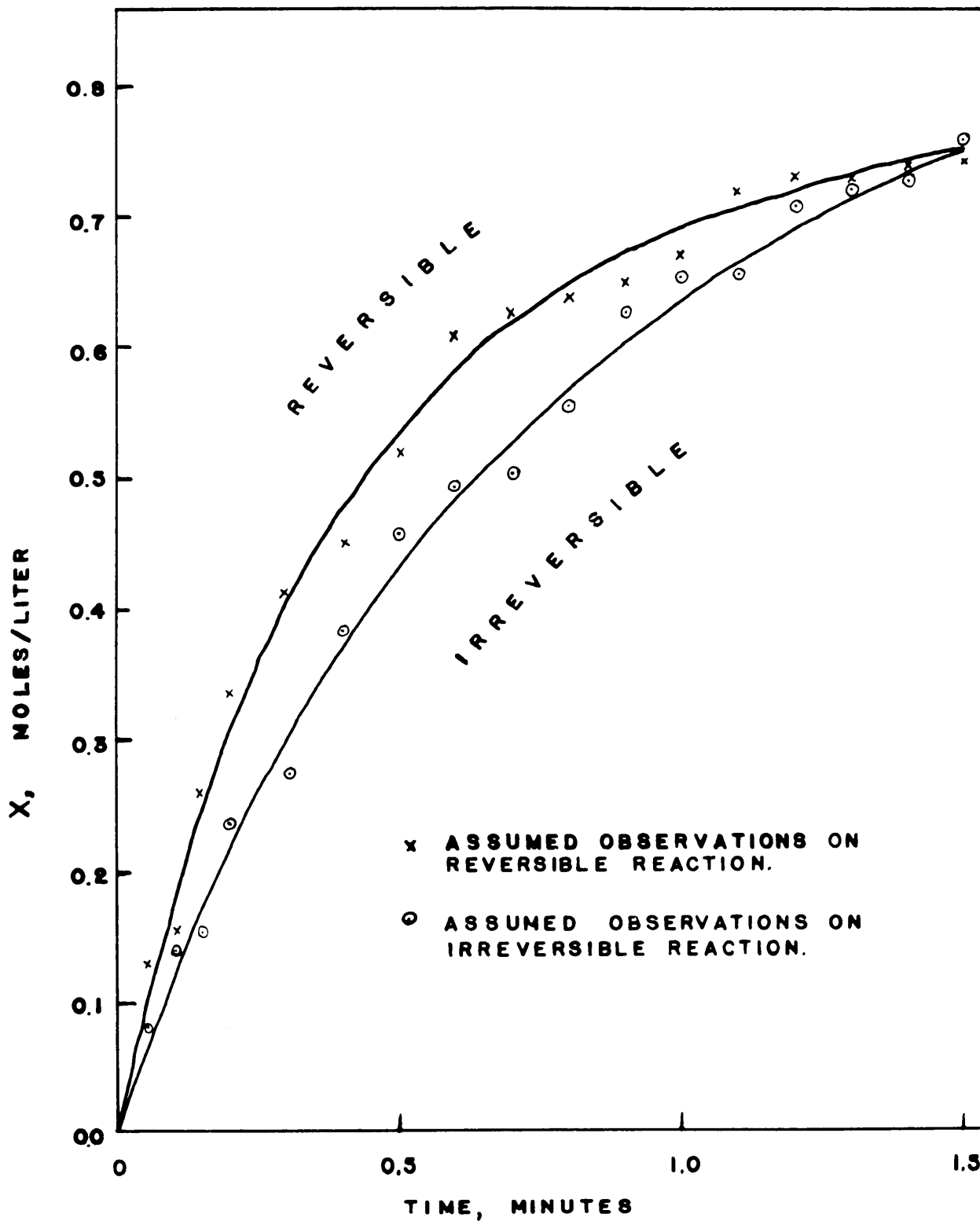
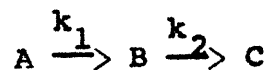


FIGURE IV. IDEAL SECOND-ORDER REACTIONS.

IV. CONSECUTIVE REACTIONS

4.1 Disregarding Reverse Reactions

Another type of reaction, the consecutive reaction, takes place when the products of one reaction undergo further chemical change to yield other products. The simplest example of this occurs when each of two consecutive reactions is of the first-order. This situation is represented by



where k_1 is the rate constant of the first reaction $A \rightarrow B$, and k_2 that of $B \rightarrow C$. For the present we shall neglect any reverse reactions.

The rate of disappearance of A at any time is given by

$$(4.1.1) \quad -da/dt = k_1 a$$

where a is the concentration of remaining reactant A. The rate of formation of C is

$$(4.1.2) \quad dc/dt = k_2 b ;$$

so that the net rate affecting the concentration of intermediate product B is

$$(4.1.3) \quad db/dt = k_1 a - k_2 b ;$$

i.e. the rate of formation of B from A, minus the rate of conversion of B to produce C. The material balance

$$(4.1.4) \quad A_{\circ} + B_{\circ} + C_{\circ} = a + b + c$$

always exists; and therefore by measuring any two components, say A and C, the amount of B present at any time can be found by subtraction.

In order to find estimates for k_1 and k_2 we take the two following expressions where differences have been used in place of differentials:

$$\begin{aligned} -(\Delta a / \Delta t)_i &= y_{1i} \approx k_1 a_i + \epsilon_{1i} \\ (\Delta b / \Delta t)_i &= y_{2i} \approx k_1 a_i - k_2 b_i + \epsilon_{2i} \end{aligned} \quad .$$

Here a_i and b_i are taken as the means of the two consecutive observations defining y_{1i} and y_{2i} respectively.

In matrix notation we have the model

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} \underline{a} & \underline{0} \\ \underline{a} & \underline{-b} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \quad ;$$

and the normal equations are

$$\begin{bmatrix} 2\Sigma a_i^2 & -\Sigma a_i b_i \\ -\Sigma a_i b_i & \Sigma b_i^2 \end{bmatrix} \begin{bmatrix} \hat{k}_1 \\ \hat{k}_2 \end{bmatrix} = \begin{bmatrix} \Sigma a_i Y_{1i} + \Sigma a_i Y_{2i} \\ -\Sigma b_i Y_{2i} \end{bmatrix} \quad .$$

For this reaction a test for irreversibility is useless since this was assumed in setting up the model; but it may be of some interest to test for equality of the two rate

constants. We test the null hypothesis $H_0: k_1 = k_2$ against the alternative $k_1 \neq k_2$ using a two-tailed t-test.

$$\text{Var}(\hat{k}_1 - \hat{k}_2) = \left(\frac{2\sum a_i^2 + \sum b_i^2 - 2\sum a_i b_i}{2\sum a_i^2 \sum b_i^2 - [\sum a_i b_i]^2} \right) \sigma^2 = c_1 \sigma^2, \text{ say;}$$

with σ^2 estimated by s^2 where

$$s^2 = \frac{1}{2n-4} [\sum Y_{1i}^2 + \sum Y_{2i}^2 - \hat{k}_1 (\sum a_i Y_{1i} + \sum a_i Y_{2i}) + \hat{k}_2 \sum b_i Y_{2i}]$$

The test statistic to be used here is

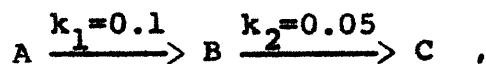
$$t = (\hat{k}_1 - \hat{k}_2) / s \sqrt{c_1} \quad ,$$

and the rejection region indicated is

$$t < t_{2n-4, \frac{1}{2}\alpha} \quad \text{and} \quad t > t_{2n-4, 1-\frac{1}{2}\alpha} \quad .$$

4.1a Numerical Example

As an example of the consecutive first-order reaction let us take



where there is initially one mole/liter of A and none of B or C (2). The amounts of A, B, and C as calculated by solving differential equations (4.1.1) and (4.1.2), and using the relationship of (4.1.4) are recorded at various times in Table 4.1 and shown graphically in Figure V. "Experimental" observations indicating errors about the curves are also

given in Table 4.2.

The totals, estimates of k_1 and k_2 , and test for equality of rate constants are as follows:

Theoretical data

$$\Sigma a_i^2 = 2.41821370 \quad \Sigma a_i y_{1i} + \Sigma a_i y_{2i} = 0.41054641$$

$$\Sigma b_i^2 = 2.10993875 \quad \Sigma b_i y_{2i} = 0.03279490$$

$$\Sigma a_i b_i = 1.40507160 \quad \Sigma y_{1i}^2 + \Sigma y_{2i}^2 = 0.03925623$$

$$\hat{k}_1 = 0.099649 \quad \hat{k}_2 = 0.050816$$

$$t = .048833/.000452 = 108.038$$

The critical region for $\alpha = .05$ is $t > 2.042$ and $t < -2.042$; and thus we would reject the hypothesis that $k_1 = k_2$. The error of approximation by difference equation is, again, very small.

"Experimental data"

$$\Sigma a_i^2 = 2.49882175 \quad \Sigma a_i y_{1i} + \Sigma a_i y_{2i} = 0.40793445$$

$$\Sigma b_i^2 = 2.13980825 \quad \Sigma b_i y_{2i} = 0.03494465$$

$$\Sigma a_i b_i = 1.38912450 \quad \Sigma y_{1i}^2 + \Sigma y_{2i}^2 = 0.04136409$$

$$\hat{k}_1 = 0.094058 \quad \hat{k}_2 = 0.044730$$

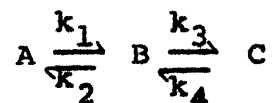
$$t = .049328/.008693 = 5.674$$

As before we reject $k_1 = k_2$, and the data indicate that k_2

determines the overall production rate of C, in accordance with the model.

4.2 Including Reverse Reactions

In the preceding section it was assumed that there were no reverse reactions in the consecutive reactions. Let us consider now the case in which a first order reaction is followed by another first order reaction and each is possibly accompanied by a reverse reaction also of the first order.



The differential rate equations for the three participants are:

$$(4.2.1) \quad \frac{da}{dt} = -k_1 a + k_2 b$$

$$(4.2.2) \quad \frac{db}{dt} = k_1 a - k_2 b - k_3 b + k_4 c$$

$$(4.2.3) \quad \frac{dc}{dt} = k_3 b - k_4 c$$

If B and C are absent at the beginning of the reaction it is possible to eliminate b from (4.2.1) and (4.2.3) by the relationship $A_0 = a+b+c$, and to rewrite them as

$$\frac{da}{dt} = -k_1 a + k_2 (a_0 - a - c)$$

and

$$\frac{dc}{dt} = k_3 (a_0 - a - c) - k_4 c$$

TABLE 4.1

Theoretical data for consecutive reaction

t	A	B	C	$-Y_{1i} = Y_i = \left(\frac{\Delta a}{\Delta t}\right)_i$	$Y_{2i} = \left(\frac{\Delta b}{\Delta t}\right)_i$	$Z_i = \left(\frac{\Delta c}{\Delta t}\right)_i$	a_i	b_i	c_i
0	1.000	0.000	0.000	-.0905	.0860	.0045	.9095	.0860	.0045
2	.819	.172	.009	-.0745	.0630	.0115	.7445	.2350	.0205
4	.670	.298	.032	-.0605	.0420	.0185	.6095	.3400	.0505
6	.549	.382	.069	-.0500	.0300	.0200	.4990	.4120	.0890
8	.449	.442	.109	-.0405	.0170	.0235	.4085	.4590	.1325
10	.368	.476	.156	-.0317	.0073	.0243	.3205	.4870	.1925
13	.273	.498	.229	-.0250	.0000	.0250	.2480	.4980	.2540
15	.223	.498	.279	-.0200	-.0050	.0250	.2030	.4930	.3040
17	.183	.488	.329	-.0160	-.0073	.0233	.1590	.4770	.3640
20	.135	.466	.399	-.0106	-.0112	.0218	.1085	.4280	.4535
25	.0821	.410	.508	-.0065	-.0148	.0192	.0659	.3780	.5560
30	.0498	.346	.604	-.0039	-.0124	.0164	.0400	.3150	.6450
35	.0302	.284	.686	-.0024	-.0100	.0124	.0248	.2590	.7170
40	.0183	.234	.748	-.0014	-.0092	.0106	.0147	.2110	.7745
45	.0111	.188	.801	-.00087	-.0074	.0082	.0089	.1695	.8215
50	.00674	.151	.842	-.00053	-.0062	.0068	.0054	.1355	.8590
55	.00409	.120	.876	-.00032	-.0050	.0054	.0033	.1075	.8895
60	.00248	.095	.903						

TABLE 4.2

"Experimental" data for consecutive reaction

t	A	B	C	$-y_{1i} = y_i = \left(\frac{\Delta a}{\Delta t}\right)_i$	$y_{2i} = \left(\frac{\Delta b}{\Delta t}\right)_i$	$z_i = \left(\frac{\Delta c}{\Delta t}\right)_i$	a_i	b_i	c_i
0	1.000	0	0	-.0765	.0750	.0015	.9235	.0750	.0015
2	.847	.150	.003	-.0695	.0520	.0175	.7775	.2020	.0205
4	.708	.254	.038	-.0980	.0740	.0240	.6100	.3280	.0620
6	.512	.402	.086	-.0235	.0225	.0010	.4885	.4245	.0870
8	.465	.447	.088	-.0355	.0150	.0205	.4295	.4620	.1085
10	.394	.477	.129	-.0463	.0087	.0377	.3245	.4900	.1855
13	.255	.503	.242	-.0270	-.0010	.0280	.2280	.5020	.2700
15	.201	.501	.298	-.0045	.0020	.0025	.1965	.5030	.3005
17	.192	.505	.303	-.0097	-.0130	.0227	.1775	.4855	.3370
20	.163	.466	.371	-.0190	-.0038	.0228	.1155	.4565	.4280
25	.068	.447	.485	-.0060	-.0220	.0280	.0530	.3920	.5550
30	.038	.337	.625	-.0006	-.0178	.0184	.0365	.2925	.6710
35	.035	.248	.717	-.0026	-.0008	.0034	.0285	.2460	.7255
40	.022	.244	.734	-.0014	-.0100	.0114	.0185	.2190	.7625
45	.015	.194	.791	-.0008	-.0126	.0134	.0130	.1625	.8245
50	.011	.131	.858	-.0016	-.0028	.0044	.0070	.124	.8690
55	.003	.117	.880	-.0004	-.0004	.0008	.0020	.1160	.8820
60	.001	.115	.884						

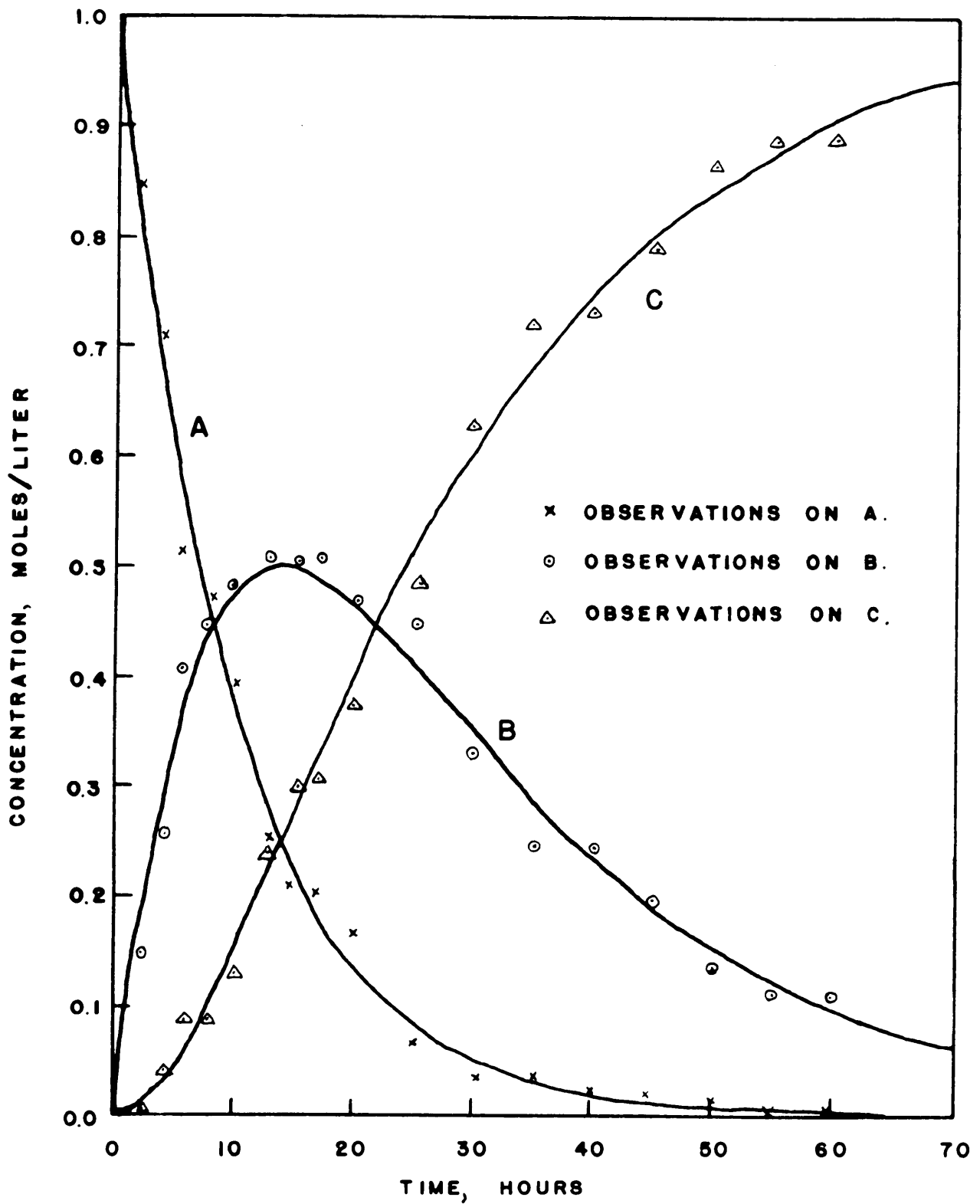


FIGURE V. CONSECUTIVE FIRST-ORDER REACTIONS WITHOUT REVERSE REACTION.

or preferably

$$da/dt = -(k_1+k_2)a + k_2(a_0-c)$$

and

$$dc/dt = k_3(a_0-a) - (k_3+k_4)c \quad .$$

Thus the rate constants k_1 and k_2 for the first reaction $A \rightleftharpoons B$ are completely separated from the constants k_3 and k_4 of the reaction $B \rightleftharpoons C$; and we have essentially two disjoint estimation problems. This step is justified from the chemical point of view since it is the concentration of B at any particular instant which affects the rate of each segment of the reaction, and the origin of B is immaterial.

Using difference quotients $(\Delta a/\Delta t)_i = y_i$ and $(\Delta c/\Delta t)_i = z_i$, we have the regression models

$$(4.2.4) \quad y_i = -(k_1+k_2)a_i + k_2(a_0-c_i) + \epsilon_{1i} \quad ,$$

and

$$(4.2.5) \quad z_i = k_3(a_0-a_i) - (k_3+k_4)c_i + \epsilon_{2i} \quad .$$

The method of least squares applied to (4.2.4) gives the normal equations

$$\begin{bmatrix} \Sigma a_i^2 & -\Sigma a_i(a_0-c_i) \\ -\Sigma a_i(a_0-c_i) & \Sigma(a_0-c_i)^2 \end{bmatrix} \begin{bmatrix} \hat{k}_1 + \hat{k}_2 \\ \hat{k}_2 \end{bmatrix} = \begin{bmatrix} -\Sigma a_i y_i \\ \Sigma(a_0-c_i) y_i \end{bmatrix}$$

from which \hat{k}_2 can be computed directly, and the estimate of

\hat{k}_1 found by subtraction of \hat{k}_2 from $(\hat{k}_1 + \hat{k}_2)$.

To determine the irreversibility of this first step of the reaction, $A \xrightleftharpoons[k_2]{k_1} B$, the hypothesis submitted for testing is $H_0: k_2 = 0$ (irreversible) vs. $k_2 > 0$ (reversible).

$$\text{Var}(\hat{k}_2) = \sigma_1^2 \Sigma a_i^2 / [\Sigma a_i^2 \Sigma (a_{0i} - c_i)^2 - [\Sigma a_i (a_{0i} - c_i)]^2] = \sigma_1^2 c_{22}, \text{ say;}$$

and here since the reaction $B \rightleftharpoons C$ is disregarded we may estimate σ_1^2 by

$$s_1^2 = \frac{1}{n-2} [\Sigma Y_i^2 + (\hat{k}_1 + \hat{k}_2) \Sigma a_i Y_i - \hat{k}_2 \Sigma (a_{0i} - c_i) Y_i]$$

The test for the proposed hypothesis is again a one-sided t-test where

$$t = \hat{k}_2 / s_1 \sqrt{c_{22}}$$

with $n-2$ degrees of freedom, which should be rejected for values of $t > t_{n-2, 1-\alpha}$.

Next we consider the second step of the reaction, $B \xrightleftharpoons[k_4]{k_3} C$.

By least squares we have from (4.2.5) the normal equations

$$\begin{bmatrix} \Sigma (a_{0i} - a_i)^2 & -\Sigma (a_{0i} - a_i) c_i \\ -\Sigma (a_{0i} - a_i) c_i & \Sigma c_i^2 \end{bmatrix} \begin{bmatrix} \hat{k}_3 \\ \hat{k}_3 + \hat{k}_4 \end{bmatrix} = \begin{bmatrix} \Sigma (a_{0i} - a_i) z_i \\ -\Sigma c_i z_i \end{bmatrix}$$

which are solved for \hat{k}_3 and $(\hat{k}_3 + \hat{k}_4)$, with \hat{k}_4 being found by subtraction.

The irreversible mechanism is tested as before by the

hypothesis that the reverse rate constant is zero, $k_4 = 0$, against the alternative $k_4 > 0$.

$$\begin{aligned} \text{Var}(\hat{k}_4) &= \text{var}[(\hat{k}_3 + \hat{k}_4) - \hat{k}_3] \\ &= \sigma_2^2 \frac{\Sigma(a_{\bullet} - a_i)^2 + \Sigma c_i^2 - 2\Sigma(a_{\bullet} - a_i)c_i}{\Sigma(a_{\bullet} - a_i)^2 \Sigma c_i^2 - [\Sigma(a_{\bullet} - a_i)c_i]^2} \\ &= \sigma_2^2 c' \text{ , say;} \end{aligned}$$

and in this case

$$s_2^2 = \left(\frac{1}{n-2}\right) [\Sigma z_i^2 - \hat{k}_3 \Sigma(a_{\bullet} - a_i)z_i + (\hat{k}_3 + \hat{k}_4) \Sigma c_i z_i] .$$

Once more the prescribed test is a one-tailed t-test,

$$t = \hat{k}_4 / s_2 \sqrt{c'}$$

with rejection region designated by the upper tail.

Since we are assuming independence of errors, i.e., treating the error in the $A \rightleftharpoons B$ reaction rate independent of that in the $B \rightleftharpoons C$ rate, a simultaneous test for $\hat{k}_2 = 0$ and $\hat{k}_4 = 0$ may be performed in two different ways.

(1) One way to test the joint hypothesis that both reactions are irreversible would be to perform two distinct t-tests such that

$$\text{Pr}[t_1 < c \text{ and } t_2 < c \mid k_2 = 0 \text{ and } k_4 = 0] = 1 - \alpha \text{ ,}$$

or, since we are considering the two reactions as independent

$$\text{Pr}[t_1 < c \mid k_2 = 0] \text{Pr}[t_2 < c \mid k_4 = 0] = 1 - \alpha \text{ .}$$

We would accept the hypothesis only if both reactions are

accepted as irreversible and reject if at least one of the reactions is reversible.

Since the degrees of freedom and constant c are identical for each reaction we should have

$$\Pr[t_1 < c \mid k_2=0] = \Pr[t_2 < c \mid k_4=0] = \sqrt{1-\alpha}$$

and therefore $c = t_{\sqrt{1-\alpha}, n-2}$, where $t_{\sqrt{1-\alpha}}$ denotes the upper value of the t-distribution corresponding to probability $\sqrt{1-\alpha}$.

(2) A second method for making the simultaneous test would be to assume that the magnitude of the errors, indicated by σ^2 , is the same for both branches of the reaction and to take

$$(4.2.6) \quad F = (2n-4) (\hat{k}_2^2/c_{22} + \hat{k}_4^2/c') / 2s^2$$

with $(2, 2n-4)$ degrees of freedom, where $s^2 = \frac{1}{2}(s_1^2 + s_2^2)$. As before the alternatives are one-sided and we want to reject if at least one of the estimates should be rejected as zero, and accept if both are to be accepted. We may obtain an approximation to the appropriate level for the F-test by the following reasoning: If the two t-tests were independent (which they are not quite here)

$$\Pr[F > c_f] = \beta$$

would imply

$$\Pr[-c_1 < t_1 < c_1] \Pr[-c_2 < t_2 < c_2] = 1-\beta \quad .$$

Further, taking $c_1=c_2$, we would have

$$\Pr[-c < t < c] = \sqrt{1-\beta} .$$

This, however, implies that

$$\Pr[t < c] = \frac{1}{2}(1 + \sqrt{1-\beta})$$

which we want to be equal to $\sqrt{1-\alpha}$.

Hence,

$$1 + \sqrt{1-\beta} = 2\sqrt{1-\alpha}$$

$$\sqrt{1-\beta} = 2\sqrt{1-\alpha} - 1$$

$$1-\beta = 4(1-\alpha) - 4\sqrt{1-\alpha} + 1$$

$$\beta = 4[\sqrt{1-\alpha} - (1-\alpha)]$$

If $\alpha = .05$, then $\beta = 4[\sqrt{.95} - .95] = .10$.

The effective doubling of the probability level, a conventional method, can thus be explained by the fact that for small β ,

$$\sqrt{1-\beta} \approx 1-\beta/2 .$$

In some cases the level β may not be found in tables of the F-distribution; we may then take the observed F and compute

$$(4.2.7) \quad B = 1/[1 + \frac{1}{n-2} F]$$

which is to be compared with tables of the lower tail of the Incomplete Beta Distribution, with parameters $p=n-2$, $q=1$, (6).

4.2a Two Numerical Examples

As a first example consider again the data of Figure V and Tables 4.1 and 4.2 where k_2 and k_4 are both actually zero.

The estimates and tests as described in this section are carried out with the following results:

Theoretical data

$\Sigma a_i y_i$	= -0.24093406	$\Sigma c_i z_i$	= 0.09881375
$\Sigma (a_{\circ} - c_i)^2$	= 7.34866825	$\Sigma (a_{\circ} - a_i)^2$	= 10.67221370
$\Sigma (a_{\circ} - c_i) y_i$	= -0.38086144	$\Sigma (a_{\circ} - a_i) z_i$	= 0.20524471
$\Sigma (a_{\circ} - c_i) a_i$	= 3.82435455	$\Sigma (a_{\circ} - a_i) c_i$	= 6.57885455
Σa_i^2	= 2.41821370	Σc_i^2	= 4.60366825
$\hat{k}_1 + \hat{k}_2$	= 0.099841	$\hat{k}_3 + \hat{k}_4$	= 0.050549
\hat{k}_2	= 0.000131	\hat{k}_3	= 0.050393
\hat{k}_1	= 0.099710	\hat{k}_4	= 0.000156

We shall omit tests of significance for estimates obtained from theoretical data which contain no errors since, as was stated previously, both the estimates of \hat{k}_2 and \hat{k}_4 , and the estimate of the variance are different from zero because of the approximation only.

"Experimental data"

$\Sigma a_i y_i$	= - 0.23763415	$\Sigma c_i z_i$	= 0.09149625
$\Sigma (a_{\circ} - c_i)^2$	= 7.41687900	$\Sigma (a_{\circ} - a_i)^2$	= 10.63982175
$\Sigma (a_{\circ} - c_i) y_i$	= - 0.37166725	$\Sigma (a_{\circ} - a_i) z_i$	= 0.19063370
$\Sigma (a_{\circ} - c_i) a_i$	= 3.88794625	$\Sigma (a_{\circ} - a_i) c_i$	= 6.54844625

$$\Sigma a_i^2 = 2.49882175 \quad \Sigma c_i^2 = 4.59687900$$

$$\Sigma y_i^2 = 0.02549547 \quad \Sigma z_i^2 = 0.00601606$$

$$\hat{k}_1 + \hat{k}_2 = 0.092903 \quad \hat{k}_3 + \hat{k}_4 = 0.045597$$

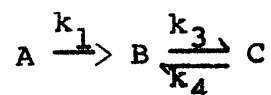
$$\hat{k}_2 = -0.001411 \quad \hat{k}_3 = 0.045980$$

$$\hat{k}_1 = 0.094314 \quad \hat{k}_4 = -0.000383$$

Since both \hat{k}_2 and \hat{k}_4 are negative in this instance the joint hypothesis of irreversibility would of course be accepted by the first method of testing.

In the second method if we choose α equal to .05, then β , the level of the F-test will be .10, and we reject the hypothesis for values of F greater than $F(2, 2n-4) = .402$. For our estimates F is equal to 0.327 which is less than the critical value, and so the hypothesis is again accepted as true at the .05 level.

As a second illustration consider a first-order reaction followed by another first-order reaction which, however, is accompanied by a reverse reaction.



where $k_1 = 0.10$, $k_3 = 0.10$, $k_4 = 0.05$, and k_2 is taken as zero (2).

Data for this reaction are found in Figure VI and Table 4.3 with "experimental" observations in Table 4.4. Results for

this data are as follows:

Theoretical data

$\Sigma a_i y_i$	= -0.24093406	$\Sigma c_i z_i$	= 0.06837210
$\Sigma (a_{\circ} - c_i)^2$	= 7.02124175	$\Sigma (a_{\circ} - a_i)^2$	= 10.67221370
$\Sigma (a_{\circ} - c_i) y_i$	= -0.36344191	$\Sigma (a_{\circ} - a_i) z_i$	= 0.15695771
$\Sigma (a_{\circ} - c_i) a_i$	= 3.65512900	$\Sigma (a_{\circ} - a_i) c_i$	= 6.04962900
Σa_i^2	= 2.41821370	Σc_i^2	= 3.55619950
Σy_i^2	= 0.02400557	Σz_i^2	= 0.00574325
$\hat{k}_1 + \hat{k}_2$	= 0.100370	$\hat{k}_3 + \hat{k}_4$	= 0.162313
\hat{k}_2	= 0.000488	\hat{k}_3	= 0.106715
\hat{k}_1	= 0.099882	\hat{k}_4	= 0.055598

Again we will omit tests of significance for data obtained from the exact curves, because errors are errors of approximation only, and are very small.

"Experimental" data

$\Sigma a_i y_i$	= -0.23763415	$\Sigma c_i z_i$	= 0.07165660
$\Sigma (a_{\circ} - c_i)^2$	= 6.88243725	$\Sigma (a_{\circ} - a_i)^2$	= 10.63982175
$\Sigma (a_{\circ} - c_i) y_i$	= -0.35526365	$\Sigma (a_{\circ} - a_i) z_i$	= 0.16385760
$\Sigma (a_{\circ} - c_i) a_i$	= 3.68120100	$\Sigma (a_{\circ} - a_i) c_i$	= 6.067886
Σa_i^2	= 2.49882175	Σc_i^2	= 3.57843725
Σy_i^2	= 0.02549547	Σz_i^2	= 0.00834455

$$\begin{array}{ll} \hat{k}_1 + \hat{k}_2 = 0.08986244 & \hat{k}_3 + \hat{k}_4 = 0.18479070 \\ \hat{k}_2 = -0.00355408 & \hat{k}_3 = 0.12078647 \\ \hat{k}_1 = 0.09341652 & \hat{k}_4 = 0.06400423 \end{array}$$

If we choose α equal to 0.05 for the joint t-test we reject the hypothesis if either value observed for t is greater than $t_{\sqrt{.95}, 15}$ which is approximately 2.131.

$$\text{Var}(\hat{k}_2) = s_1^2 c_{22} = (.00019190)(0.68522013) = .00013149$$

$$t_1 = -.003554 / .011467 = -.3099$$

We accept that the reaction $A \rightarrow B$ is irreversible.

$$\text{Var}(\hat{k}_4) = s_2^2 c' = (.00011962)(1.6597568) = .0001984$$

$$t_2 = .064004 / .014090 = 4.543$$

Here we must reject the hypothesis that the reaction $B \rightarrow C$ is irreversible, and hence reject the joint hypothesis at the .05 level.

For the second method of testing in which the magnitude of error is assumed to be the same for each branch of the reaction we compute $F = 239.467$. This quite definitely indicates that the hypothesis, $k_2=0$ and $k_4=0$, should be rejected.

TABLE 4.3

Theoretical data for consecutive reaction

t	A	B	C	$y_i = \frac{\Delta A}{\Delta t}$	$z_i = \frac{\Delta C}{\Delta t}$	a_i	c_i
0	1.000	0.000	0.000	-.0905	.0065	.9095	.0065
2	.819	.168	.013	-.0745	.0255	.9615	.0385
4	.670	.266	.064	-.0605	.0260	.9100	.0900
6	.549	.335	.116	-.0500	.0340	.8500	.1500
8	.449	.367	.184	-.0405	.0270	.7890	.2110
10	.368	.394	.238	-.0317	.0233	.7270	.2730
13	.273	.419	.308	-.0250	.0240	.6680	.3320
15	.223	.421	.356	-.0200	.0245	.6195	.3805
17	.183	.412	.405	-.0160	.0213	.5635	.4365
20	.135	.397	.468	-.0106	.0136	.4980	.5020
25	.0821	.382	.536	-.0065	.0088	.4420	.5580
30	.0498	.370	.580	-.0039	.0052	.4070	.5930
35	.0302	.364	.606	-.0024	.0038	.3845	.6155
40	.0183	.357	.625	-.0014	.0030	.3675	.6325
45	.0111	.349	.640	-.00087	.0014	.3565	.6435
50	.00674	.346	.647	-.00053	.0012	.3500	.6500
55	.00409	.343	.653	-.00032	.0008	.3450	.655
60	.00248	.340	.657				

TABLE 4.4

"Experimental" data for consecutive reaction

t	A	B	C	$y_i = \frac{\Delta A}{\Delta t}$	$z_i = \frac{\Delta C}{\Delta t}$	a_i	c_i
0	1.000	0.000	0.000	-.0765	.0110	.9235	.0110
2	.847	.131	.022	-.0695	.0075	.7775	.0295
4	.708	.255	.037	-.0980	.0335	.6100	.0705
6	.512	.384	.104	-.0235	.0530	.4885	.1570
8	.465	.325	.210	-.0355	.0290	.4295	.2390
10	.394	.338	.268	-.0463	.0073	.3245	.2790
13	.255	.455	.290	-.0270	.0500	.2280	.3400
15	.201	.409	.390	-.0045	.0150	.1965	.4050
17	.192	.388	.420	-.0097	.0100	.1775	.4350
20	.163	.387	.450	-.0190	.0122	.1155	.4805
25	.068	.421	.511	-.0060	.0180	.0530	.5560
30	.038	.361	.601	-.0006	.0022	.0365	.6065
35	.035	.353	.612	-.0026	.0006	.0285	.6135
40	.022	.363	.615	-.0014	.0028	.0185	.6220
45	.015	.356	.629	-.0008	.0048	.0130	.6410
50	.011	.336	.653	-.0016	.0000	.0070	.6530
55	.003	.344	.653	-.0004	.0028	.0020	.6600
60	.001	.332	.667				

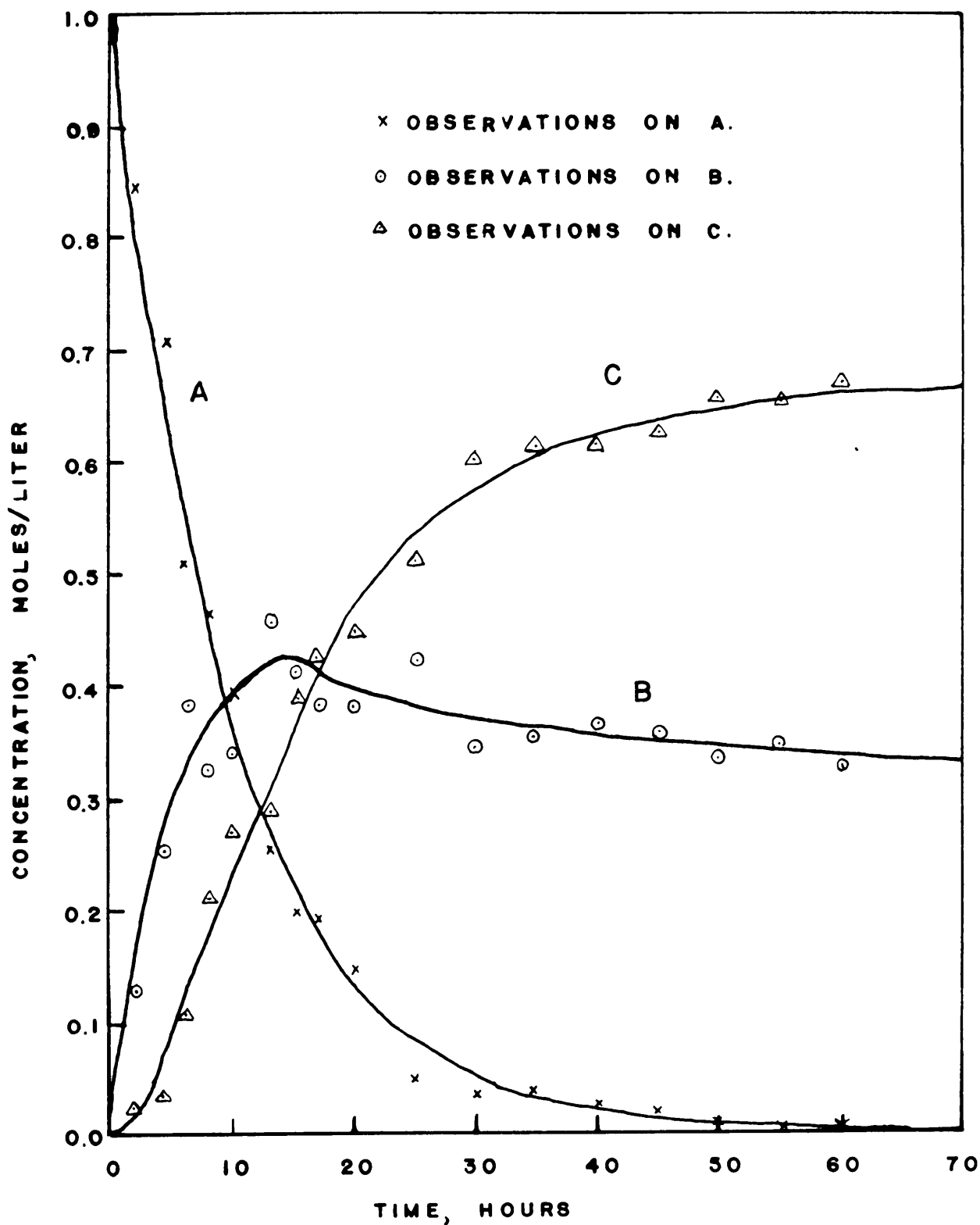


FIGURE VI. CONSECUTIVE FIRST-ORDER REACTIONS WITH REVERSE REACTION.

V. SUMMARY AND CONCLUSIONS

In this thesis an attempt has been made to explore the use of an approximate method, that of Prony, for the estimation of rate constants in chemical kinetics. The discussion has been restricted to some simple reversible and consecutive reactions. The proposed methods could be extended to more complex situations since it has the advantage that the rate equation need not be integrated.

The method of least squares for obtaining the slope of a straight line in the calculation of rate constants has sometimes been regarded as questionable (Roseveare, 1931 (8)). The use of $\Delta x/\Delta t$ for dx/dt results in a different kind of approach from that previously used. In the above-mentioned paper, least squares techniques were applied after the differential rate equation had been integrated. For example, for a first-order reaction the rate equation $dx/dt=k(A_0-x)$ was solved to give $x = A_0(1-e^{-kt})$, and the least squares estimate obtained as

$$k = \frac{\sum p t \sum \ln(a_0-x) - \sum p \sum [p t \ln(a_0-x)]}{[\sum p t]^2 - \sum p \sum t^2}$$

Here p is a weighting factor which at any point is inversely proportional to the square of the probable error. The

probable error of $(a_0 - x)$ was considered as a constant quantity rather than a constant percentage error, and each point had a different weight since $(a_0 - x)$ and not $\ln(a_0 - x)$ was measured directly.

More recently least squares techniques were applied (Hougen and Watson, 1947 (5)) in an attempt to indicate which one of eighteen possible complex reactions should be used to represent a particular set of rate data. However, as pointed out by Chou (1), the choice of a mechanism based entirely on the nature of the estimated parameters is not sound. In the examples presented in this paper no attempt was made to estimate the order of the reactions; rather we specified several particular models and attempted to select an appropriate reaction mechanism among possible rivaling ones by statistical significance tests.

A number of demonstration studies showed that, even by the crude estimation techniques employed, very good estimates were always obtained, and the tests between rivaling models always led to the correct decision. In one instance, a maximum-likelihood estimate was obtained which required considerable numerical work, and which was in no way superior to the estimate obtained by the Prony method.

VI. BIBLIOGRAPHY

- (1) Chou, C. "Least Squares". Industrial and Engineering Chemistry, (May, 1958) pp. 799-802.
- (2) Daniels, F., and Alberty, R. A. Physical Chemistry, Wiley & Sons, Inc., (1955).
- (3) Glasstone, S. Textbook of Physical Chemistry, D. Van Nostrand Co., (1940).
- (4) Hartley, H. O. "The Estimation of Non-Linear Parameters by 'Internal Least Squares'". Biometrika, Volume 35, (1948), pp. 32-45.
- (5) Hougen, O. A., and Watson, K. M. Chemical Process Principles, Part III, "Kinetics and Catalysis", Wiley & Sons, Inc., (1947).
- (6) Pearson, E. S., and Hartley, H. O. Biometrika Tables for Statisticians, Volume I. Cambridge University Press, (1956).
- (7) Prony "Sur La Dilatation". Journal De L'Ecole Polytechnique, Cahier 2, An. IV (1795), pp. 28-34.
- (8) Roseveare, W. E. "Methods of Calculating and Averaging Rate Constants". Journal of the American Chemical Society, Volume 53, (1931), pp. 1651-1661.
- (9) Walas, S. M. Reaction Kinetics for Chemical Engineers, McGraw-Hill Book Company, Inc., (1959).
- (10) Whittaker, and Robinson, The Calculus of Observations, Blackie and Son, Limited, (1944).

VII. ACKNOWLEDGMENTS

The author wishes to express her sincere appreciation to Dr. Rolf E. Bargmann for his invaluable guidance and advice during the course of this work.

Gratitude is extended to Dr. Clyde Y. Kramer for his many suggestions and constructive criticisms. Appreciation is also expressed to Dr. Boyd Harshbarger for his advice, assistance, and encouragement to the graduate student.

The author is also indebted to _____ for the preparation of the final typewritten copies.

This study was supported in part by a research grant from the National Institutes of Health, under contract number 2G.

**The vita has been removed from
the scanned document**

ABSTRACT

This paper proposes and compares some statistical methods for the analysis of chemical reaction rate data. A method, similar to that of Prony, is applied for the approximation of rate constants in simple reversible and consecutive reactions. Difference quotients are substituted for differential quotients of the rate equation, and least squares techniques are employed on the resulting regression models to estimate the specific reaction rate constant. In one instance the approximation thus obtained is compared with the maximum likelihood estimate.

Furthermore, hypotheses designed to discriminate between reversible and irreversible reaction mechanisms are formulated and tested.

Several numerical demonstration examples both for exact functions and for observations involving measurement error about such function are given in order to illustrate the proposed method.