ANL-7113 ANL-7113

Argonne National Laboratory

A METHOD FOR AUTOMATIC ANALYSIS OF DATA ON THE KINETICS OF FAST CHEMICAL REACTIONS IN THE FORM OF PICTURES OF OSCILLOSCOPE TRACES

by

Myran C. Sauer, Jr.

LEGAL NOTICE

Printed in USA. Price $1.00. Available from the Clearinghouse for Federal Scientific and Technical Information, National Bureau of Standards, U. S. Department of Commerce, Springfield, Virginia

ANL-7113
Chemistry
(TID-4500, 47th Ed.)
AEC Research and
Development Report

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60440

A METHOD FOR AUTOMATIC ANALYSIS OF DATA ON THE KINETICS OF FAST CHEMICAL REACTIONS IN THE FORM OF PICTURES OF OSCILLOSCOPE TRACES

by

Myran C. Sauer, Jr.

Chemistry Division

October 1965

Operated by The University of Chicago
under
Contract W-31-109-eng-38
with the
U. S. Atomic Energy Commission

TABLE OF CONTENTS

LIST OF FIGURES

A METHOD FOR AUTOMATIC ANALYSIS OF DATA ON THE KINETICS OF FAST CHEMICAL REACTIONS IN THE FORM OF PICTURES OF OSCILLOSCOPE TRACES

by

Myran C. Sauer, Jr.

ABSTRACT

The CHLOE system for direct digitization of photographic data has been used in conjunction with the Control Data Corporation 3600 Computer system to analyze data, in the form of pictures of oscilloscope traces, on the kinetics of fast chemical reactions. All measurements and analytical processes were done by machine in a completely automatic method.

I. INTRODUCTION

An intermediate stage in many experiments is the display of curves on a chart-recorder graph, an oscilloscope trace, or some other similar record. Oftentimes, laborious and tedious analyses must be performed "by hand." This report describes a method for automatically analyzing pictures of oscilloscope traces. In principle, of course, the method can be applied to analysis of other similar data.

A prerequisite for such automatic analysis of pictures is a machine that determines the x, y coordinates of significant features on the picture. Such a machine has been developed in the Applied Mathematics Division by a group led by Donald Hodges.* A brief description is given in Section III of this report. Essentially, the machine, referred to as the CHLOE system, produces a magnetic-tape record of the x, y coordinates, at specified intervals, of the various curves on the picture.

A program has been written to use the 3600 Computer to analyze such magnetic tape records.

---

*Details concerning this machine may be obtained from Donald Hodges, Applied Mathematics Division, Argonne National Laboratory.

II. DESCRIPTION OF THE PHOTOGRAPHIC PROCESSES INVOLVED

The oscilloscope trace (Tektronix Type 555 dual-beam oscilloscope) was photographed by using Polaroid Land Picture Roll, 3000 speed/Type 47 film which produces a $4\frac{3}{16} \times 3\frac{1}{4}$-in. picture, such as that shown in Fig. 1. Since the CHLOE system requires the data to be on 35-mm film, the Polaroid pictures were photographed with an Exa 35-mm, single-lens, reflex camera using Kodak fine-grain positive film (P651-1). The camera was used with a set of extension tubes, which lengthened the camera-lens distance by 5/8 in. Under these conditions, the camera had to be focused on infinity to bring the Polaroid picture into focus. By means of the lamps and arrangement shown in Fig. 2, the shutter speed was set at 1/25 sec at $f/3\frac{1}{2}$ to obtain the correct exposure.

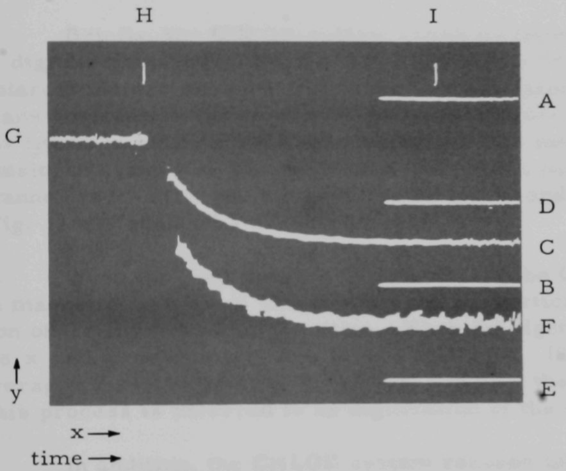

Fig. 1

Polaroid Picture of the Oscilloscope Traces Representing a Typical Transient Formation

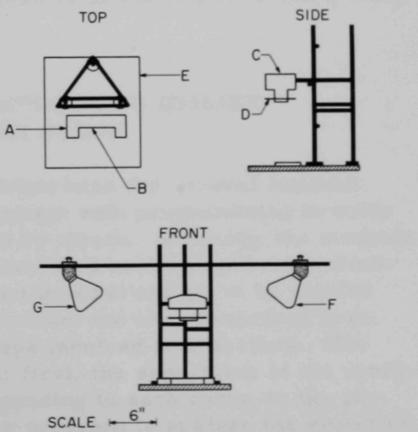

Fig. 2

Apparatus for Taking 35-mm Photographs of Polaroid Pictures.*

A--3/8-in.-thick Lucite, screwed to E
B--Flat edge against which Polaroid Picture is placed
C--Camera
D--Black cardboard to eliminate glare and reflections
E--1/2-in.-thick Masonite
F and G--General Electric Reflector Photolamps, PH 500/32R7

The supporting framework was constructed from Flexaframe rods (1/2-in. diam), supports, and clamps. The clamps are not shown. The camera was fastened to the aluminum rod by threading the rod so that it fit into the tripod fitting on the side of the camera.

*For clarity, all components of the apparatus are not shown in every view.

The Polaroid pictures were placed (see Fig. 2) in the Lucite frame, A, and held flat by a Lucite plate (not shown) which fit into the frame, while the 35-mm pictures were taken. A frame of black paper on this Lucite plate defined the edges of the picture. With the method described in Section IV, it was important to have the horizontal lines (Fig. 1) nearly parallel to the edge of the 35-mm film. Therefore, the Polaroid picture was always placed with the top edge against the flat back edge (B) of the Lucite frame. The position of the camera was then adjusted relative to the Lucite frame until the horizontal lines had nearly zero slope (less than 0.01), as measured in the process of computation.

III. DESCRIPTION OF THE CHLOE SYSTEM OF DIGITIZATION

Briefly, the CHLOE system works as follows: The frame of film to be digitized, corresponding, for example, to the 35-mm picture of the Polaroid picture shown in Fig. 1, is held in a fixed position while a light beam scans horizontally or vertically at specified intervals. The number of scans per frame in either direction is variable. For the work described here, the density of scans was 256 per frame. The main part of the picture was scanned vertically, and the area containing H and I at the top of the picture (Fig. 1) was scanned horizontally.

When the light beam falls on a curve, the CHLOE system registers on magnetic tape the x (horizontal) and y (vertical) coordinates of the position of the light beam at that time. When the light beam passes off the curve, the x and y coordinates are again registered. (Subsequent programming averages these two sets of coordinates to give the "midpoint" of the curve.) This process is referred to as digitization of the curves.

In addition, the CHLOE system records information indicating the start of a new picture, the number of the picture in the sequence being read, and the various conditions of scanning.

IV. PROGRAMMING THE COMPUTER TO UTILIZE THE RECORD FROM CHLOE

The program is not described in detail here for several reasons. First of all, the author was not familiar enough with programming to write a program that could be easily interpreted by others. Secondly, the methods used are simple in principle and can be described easily in general terms. Finally, the program as it stands now is not general enough to be applied to the analysis of pictures that differ much from the ones described here. Therefore, only the general idea of the steps involved is described. The description is divided into two main parts; first, the separation of the array of x, y coordinates into the groups corresponding to each curve on the picture, and second, a brief description of the program to analyze the resulting

The Polaroid pictures were placed (see Fig. 2) in the Lucite frame, A, and held flat by a Lucite plate (not shown) which fit into the frame, while the 35-mm pictures were taken. A frame of black paper on this Lucite plate defined the edges of the picture. With the method described in Section IV, it was important to have the horizontal lines (Fig. 1) nearly parallel to the edge of the 35-mm film. Therefore, the Polaroid picture was always placed with the top edge against the flat back edge (B) of the Lucite frame. The position of the camera was then adjusted relative to the Lucite frame until the horizontal lines had nearly zero slope (less than 0.6%, as measured in the process of computation.

III. DESCRIPTION OF THE CHLOE SYSTEM OF DIGITIZATION

Briefly, the CHLOE system works as follows. The frame of film to be digitized, corresponding, for example, to the 35-mm picture of the Polaroid picture shown in Fig. 1, is held in a fixed position while a light beam scans horizontally or vertically at specified intervals. The number of scans per frame in either direction is variable. For the work described here, the density of scans was 256 per frame. The main part of the picture was scanned vertically, and the area containing M and I at the top of the picture (Fig. 1) was scanned horizontally.

When the light beam falls on a curve, the CHLOE system registers on magnetic tape the x (horizontal) and y (vertical) coordinates of the position of the light beam at that time. When the light beam passes off the curve, the x and y coordinates are again registered. (Subsequent programming averages these two sets of coordinates to give the "midpoint" of the curve.) This process is referred to as digitization of the curves.

In addition, the CHLOE system records information indicating the start of a new picture, the number of the picture in the sequence being read, and the various conditions of scanning.

IV. PROGRAMMING THE COMPUTER TO UTILIZE THE RECORD FROM CHLOE

The program is not described in detail here for several reasons. First of all, the author was not familiar enough with programming to write a program that could be easily interpreted by others. Secondly, the methods used are simple in principle and can be described easily in general terms. Finally, the program as it stands now is not general enough to be applied to the analysis of pictures that differ much from the ones described here. Therefore, only the general idea of the steps involved is described. The description is divided into two main parts: first, the separation of the array of x, y coordinates into the groups corresponding to each curve on the picture, and second, a brief description of the program to analyze the resulting curves.

A. Arrangement of x, y Coordinates from CHLOE into Groups Corresponding to the Curves on the Picture

The first step is to get the data for a given picture from the magnetic tape into the computer. This was accomplished by using several tape-handling subroutines.*

On a particular vertical scan of a picture, CHLOE may occasionally not register a set of coordinates for a curve if the contrast is not high enough at that part of the curve. The program must, of course, allow for this. The method starts with the first point of the array (from the vertical scanning), which is from the right-hand side of Fig. 1, since successive scans moved to the left. Each succeeding point is tested to see if it belongs to a new curve or to one already started. After moving approximately one-fifth of the width of the picture, the process is stopped, and curves with less than N points are discarded (N was usually set at 15 to 20). In this way, the initial parts of curves A-F of Fig. 1 are obtained. (The method does not work if any of the curves cross or overlap.) Further points in the array are then tested by extrapolating the last N points belonging to each curve by a least-squares straight line "drawn" through these points. This process is repeated, always using the last N points as a basis for extrapolation and testing of new points, until all points in the array have been considered. If no additional points are found for a curve during a change in x of about 5% of the picture width, that curve is considered to have ended and is not tested further. Curve G on the left-hand side of Fig. 1 is determined in a similar manner, starting with the last point in the array.

Lines H and I in Fig. 1 are determined from the horizontal scan data, which are set to cover only the area above curve A. These lines are needed to calibrate the horizontal (x) axis in terms of time.

B. Analysis of the Curves after They Have Been Separated

A simple test is applied at first to determine which of the curves A-F correspond to the horizontal lines A, D, B, and E. The average y coordinate for each of these lines is then determined. The curves corresponding to C and F are treated as follows: The first N points (N was again usually set at about 15 to 20) at the right of the curve are used for a linear, least-squares extrapolation against which the next point on that curve is tested. If the next point is within a given limit of the extrapolated value, these N + 1 points are extrapolated to test the following point, etc. When the curve begins to bend, i.e., after several points in succession do not fit the extrapolated value, the average value of the linear part is determined. (The determination of the linear parts of curves C and F is important since, in this particular example, the Δy's of points on C and F from these linear parts of C and F are measured.)

*The subroutines were written by L. C. Just of the Applied Mathematics Division.

A. Arrangement of x, y Coordinates from CHLOE into Groups Corresponding to the Curves on the Picture

The first step is to get the data for a given picture from the magnetic tape into the computer. This was accomplished by using several tape-handling subroutines.*

On a particular vertical scan of a picture, CHLOE may occasionally not register a set of coordinates for a curve if the contrast is not high enough at that part of the curve. The program must, of course, allow for this. The method starts with the first point of the array (from the vertical scanning), which is from the right-hand side of Fig. 1, since successive scans moved to the left. Each succeeding point is tested to see if it belongs to a new curve or to one already started. After moving approximately one-fifth of the width of the picture, the process is stopped, and curves with less than N points are discarded (N was usually set at 15 to 20). In this way, the initial parts of curves A-F of Fig. 1 are obtained. (The method does not work if any of the curves cross or overlap.) Further points in the array are then tested by extrapolating the last N points belonging to each curve by a least-squares straight line "drawn" through these points. This process is repeated, always using the last N points as a basis for extrapolation and testing of new points until all points in the array have been considered. If no additional points are found for a curve during a change in x of about 5% of the picture width, that curve is considered to have ended and is not tested further. Curve G on the left-hand side of Fig. 1 is determined in a similar manner, starting with the last point in the array.

Lines H and I in Fig. 1 are determined from the horizontal scan data, which are set to cover only the area above curve A. These lines are needed to calibrate the horizontal (x) axis in terms of time.

B. Analysis of the Curves after They Have Been Separated

A simple test is applied at first to determine which of the curves A-F correspond to the horizontal lines A, D, E, and E. The average y coordinate for each of these lines is then determined. The curves corresponding to C and F are treated as follows: The first N points (N was again usually set at about 15 to 20) at the right of the curve are used for a linear, least-squares extrapolation against which the next point on that curve is tested. If the next point is within a given limit of the extrapolated value, these N + 1 points are extrapolated to test the following point, etc. When the curve begins to bend, i.e., after several points in succession do not fit the extrapolated value, the average value of the linear part is determined. (The determination of the linear parts of curves C and F is important since, in this particular example, the Δy's of points on C and F from these linear parts of C and F are measured.)

*The subroutines were written by L. C. Just of the Applied Mathematics Division.

On the basis of the average values of y thus obtained, the curves are separated into two sets. A, B, C, and G correspond to the traces from the upper beam of the oscilloscope, and D, E, and F to the traces from the lower beam of the oscilloscope. (At this point, the problem of analysis becomes essentially the same as if one had measured the coordinates of various points on the curves "by hand" and put them onto punched cards for analysis by a computer program, except that in this case we do not need to punch cards.)

In this example, the system being studied* is tested for linearity in a plot of

$$\log\log \frac{(A - B)(AMP) - (G - C_x)}{(A - B)(AMP) - (G - C)} \text{ vs } x,$$

where A, B, G, and C refer to the average y coordinates of curves A, B, G, and the linear part of C, respectively, C_x is the y-coordinate of C at x, and AMP is an amplification factor determined by an oscilloscope setting. A similar plot is made for D, E, and F (with the qualification that some information from A, B, C, and G has to be used to determine a value corresponding to G for A, B, and C).

These plots are made, and a least-squares straight line is determined for each. The slope of such a line is the desired product of the processes described. As a visual means of checking the amount of noise and the deviation from linearity of these plots, the computer is programmed to produce a plot on the off-line Calcomp x, y plotter.** Such a plot is shown in Fig. 3, which results from the analysis of Fig. 1. The open points represent the results from A, B, C, and G, and the black points represent D, E, and F. Both curves are plotted on the same graph since they represent the same experiment recorded under two different amplification factors on the two channels of the oscilloscope.

The program is written so that it will automatically handle variations of Fig. 1, e.g., if only one set of curves is present, or if a second line appears at the left of the picture.

---

*Curves C and F of Fig. 1 represent the formation of ozone (O_3) by the reaction of oxygen atoms with molecular oxygen in a system of argon, carbon monoxide, and oxygen that has been irradiated, slightly after the time corresponding to H on Fig. 1, with a 1.0-μsec pulse of electrons. (I is 30 μsec after H.) For a complete description of a similar chemical system, see: M. C. Sauer, Jr., and L. M. Dorfman, J. Am. Chem. Soc., 87, 3801 (1965).

**Information concerning the subroutines used for the Calcomp Magnetic Tape Plotting System No. 580, by which all the graphs shown in this report were plotted, may be obtained from C. G. LeVee and J. A. Ohde, Applied Mathematics Division, Argonne National Laboratory.

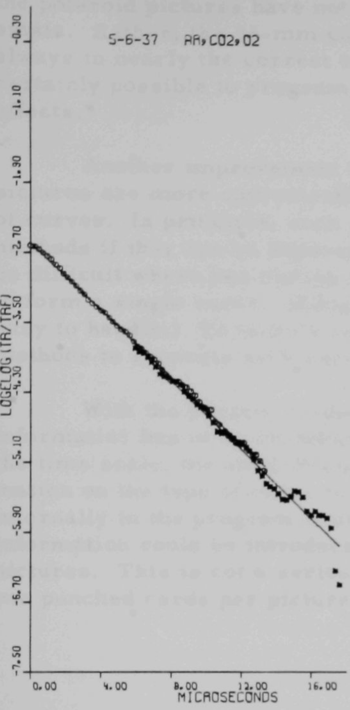

Fig. 3. First-order Test Plot for Curves in Fig. 1

Subroutines have also been written which have successfully analyzed similar curves which represent the decay of a transient species, i.e., where the curves corresponding to C and F start off at a low value of y and increase (with the opposite curvature to C and F in Fig. 1) to a higher value of y. Also, the results from a series (20, for example) of such curves have been used to obtain and automatically plot the optical absorption spectra of transients. Another subroutine is used to analyze curves that are similar, for example, to the upper trace (A, B, C, and G) in Fig. 1, except that the curve corresponding to C does not reach a constant value of y. Examples of these cases are shown in the appendix.

V. COMPARISON OF AUTOMATIC ANALYSIS WITH ANALYSIS "BY HAND"

The computer time necessary to analyze a typical picture such as Fig. 1 is only about 15 sec with a fast computer such as the CDC-3600. Analysis by hand, using only one-tenth (or fewer) of the points used by the computer, takes considerably longer, at least 15 min being required to produce the linear test plot and determine the slope. Since the computer uses many more points, the noise is automatically averaged in the result. Also, the larger number of points allows one to distinguish small amounts of curvature in the linear test plots from noise. When analysis is performed with fewer points, such curvature is not usually as easily distinguished from random errors or noise.

VI. SOME IMPROVEMENTS THAT CAN BE MADE IN THE PRESENT METHOD

The oscilloscope now being used does not produce perfectly parallel lines if the trace is triggered at various vertical positions. The traces have some "bow," the amount depending on the vertical position. The error induced by this effect is small for the purposes of the analyses described in this report, but in principle could be eliminated by using a "calibration pattern" of traces at various vertical positions. Also, small errors introduced by slight misalignment in the process of taking 35 mm pictures of

the polaroid pictures have not been corrected for by the program as it now exists. Rather, the 35-mm camera is adjusted so that the pictures are always in nearly the correct orientation, as previously described. It is certainly possible to program the computer to correct for both of these effects.*

Another improvement that can be made concerns the fact that many pictures are more conveniently taken in a manner that produces crossings of curves. In principle, such curves can probably be separated by computer methods if they can be separated by the human eye, but in practice this may be difficult where two curves approach each other at a small angle and merge to form a single curve. (Crossing at a sharper angle would be relatively easy to handle.) To make a curve-reading program generally applicable, methods to separate such curves are necessary.

With the present method, for each picture to be analyzed, some information has to be introduced to the computer on punched cards, i.e., the time scale, the amplification factor, an identifying number, and information on the type of curve to be analyzed. The latter could be determined internally in the program from the nature of the curves present. The other information could be introduced by marks of some kind made on the 35-mm pictures. This is not a serious problem, however, since usually only two punched cards per picture need to be used at present.

---

*A program to analyze pictures similar to those described here, which includes a correction for deviation from zero horizontal slope due to poor picture alignment, has been developed by A. H. Lent and J. W. Butler of the Applied Mathematics Division.

the polaroid pictures have not been corrected for by the program as it now exists. Rather, the 35-mm camera is adjusted so that the pictures are always in nearly the correct orientation, as previously described. It is certainly possible to program the computer to correct for both of these effects.*

Another improvement that can be made concerns the fact that many pictures are more conveniently taken in a manner that produces crossings of curves. In principle, such curves can probably be separated by computer methods if they can be separated by the human eye, but in practice this may be difficult where two curves approach each other at a small angle and merge to form a single curve. (Crossing at a sharper angle would be relatively easy to handle.) To make a curve-reading program generally applicable, methods to separate such curves are necessary.

With the present method, for each picture to be analyzed, some information has to be introduced to the computer on punched cards, i.e., the time scale, the amplification factor, an identifying number, and information on the type of curve to be analyzed. The latter could be determined internally in the program from the nature of the curves present. The other information could be introduced by marks of some kind made on the 35-mm pictures. This is not a serious problem, however, since usually only two punched cards per picture need to be used at present.

---

*A program to analyze pictures similar to those described here, which includes a correction for deviation from zero horizontal slope due to poor picture alignment, has been developed by Dr. H. Lass and J. W. Butler of the Applied Mathematics Division.

APPENDIX

Examples of Analyses of Different Types of Transient Curves

This appendix is for those who are familiar with typical oscilloscope pictures produced in pulse radiolysis experiments, and want to see examples of the analyses.

Figure 4 shows one of a set of transient decay curves analyzed by the method described in this report. Figure 5 shows the resulting first-order test plot, in which the completely black points represent the upper trace of Fig. 4. Figure 6 shows the spectrum (optical density vs wavelength) that was determined from a group of similar pictures in which the lower trace corresponded to a fixed wavelength (as a standard), and the upper trace corresponded to a variable wavelength. Figure 7 shows the same spectrum plotted as optical density vs energy in electron volts. The various sets of points correspond to spectra at different times along the decay curve. (In Figs. 6 and 7, the curves connecting the points were drawn in by hand, as a visual aid.)

Figure 8 shows a curve in which the transient at first increases in concentration and, after reaching a maximum, begins to decay. A subroutine has been written to correct this type of curve (on the basis of the decay rate measured sufficiently after the maximum that formation processes are negligible) to the hypothetical curve that would be produced if no decay had occurred. This is shown in Fig. 9, where the curve made up of x's represents the actual optical density read from Fig. 8, and the curve made up of circles represents the hypothetical corrected curve. The corrected curve is then analyzed by the same method as that described for Fig. 1, to obtain a test for linearity, and a resulting slope of the least-squares plot. Figure 10 shows the linear test plot of the corrected curve in Fig. 9.

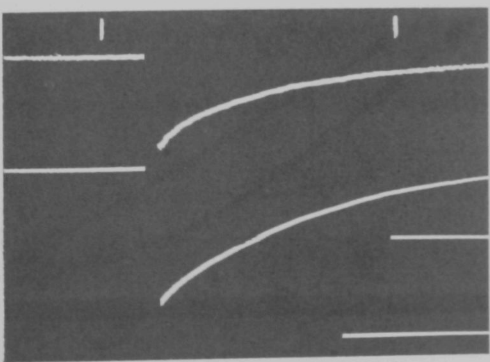

Fig. 4
Transient Decay Curves

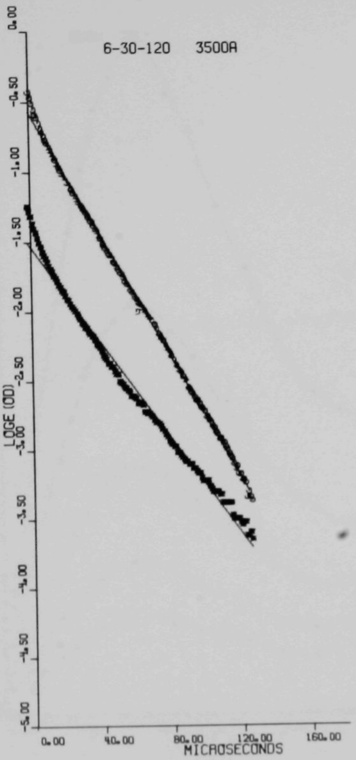

Fig. 5. First-order Test Plots for Curves in Fig. 4

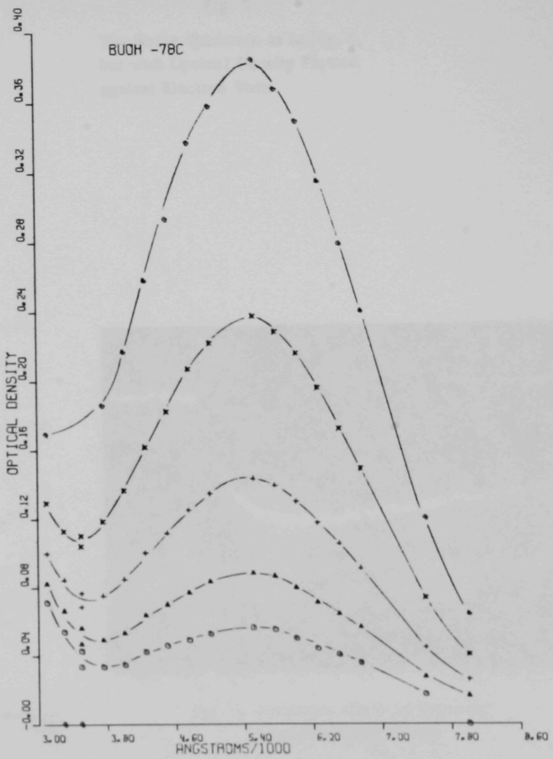

Fig. 6. Transient Spectrum Obtained from a Group of Pictures Similar to Fig. 4, Where Optical Density is Plotted against Angstroms

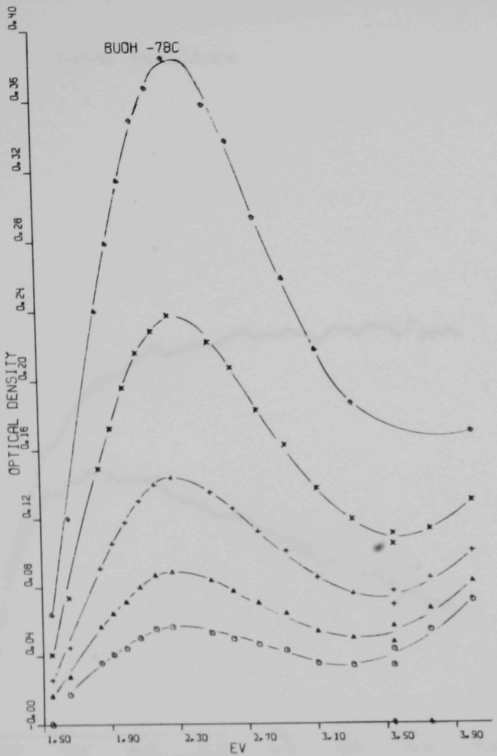

Fig. 7

The Same Spectrum as in Fig. 6, but with Optical Density Plotted against Electron Volts

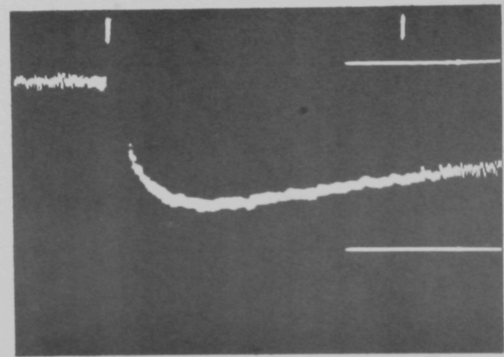

Fig. 8. Formation Curve of Transient Complicated by Decay

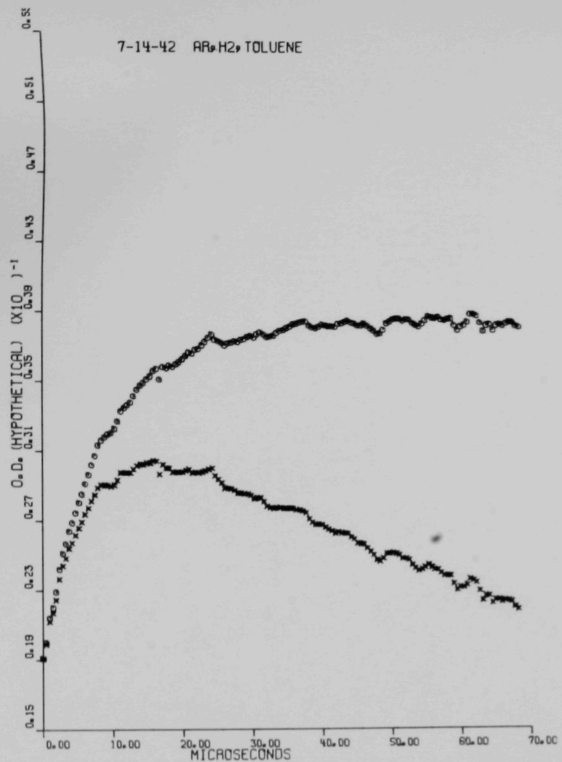

Fig. 9. Hypothetical Formation Curve Corresponding to Fig. 8, Where Correction for Decay Has Been Applied

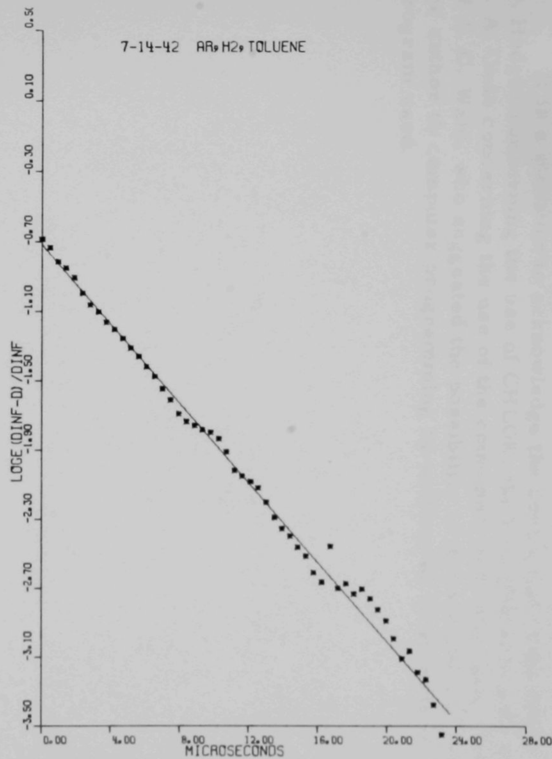

Fig. 10. First-order Test Plot for Hypothetical Curve in Fig. 9

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the advice and cooperation of D. Hodges concerning the use of CHLOE, the help given the author by J. A. Ohde concerning the use of the computer, and especially, the efforts of A. C. Wahl, who suggested the possibility of this method and introduced the author to computer programming by writing the first parts of the program used.