Polymers", *J. Chem. Doc.*, **10**, 243–247 (1970).
(11) H. Skolnik, "A Notation Symbol Index for Chemical Compounds", *J. Chem. Doc.*, **11**, 120–124 (1971).
(12) H. Skolnik, "A Chemical Fragment Notation Index", *J. Chem. Doc.*, **11**, 142–147 (1971).
(13) H. Skolnik, "The What and How of Computers for Chemical Information Systems", *J. Chem. Doc.*, **11**, 185–189 (1971).
(14) H. Skolnik, "A Computerized System for Storing, Retrieving, and Correlating NMR Data", *Appl. Spectrosc.*, **26**, 173–182 (1972).
(15) H. Skolnik and J. C. Snyder, "Input/Output Considerations for Large Data Bases", *J. Chem. Inf. Comput. Sci.*, **15**, 28–31 (1975).
(16) H. Skolnik, "Milestones in Chemical Information Science", *J. Chem. Inf. Comput. Sci.*, **16**, 187–193 (1976).
(17) E. G. Smith, "The Wiswesser Line-Formula Chemical Notation", McGraw-Hill, New York, N.Y., 1968.
(18) J. R. Ullmann, "The Credibility of Machine Intelligence", *Nature (London)*, **257**, 547–549 (1975).
(19) W. S. Woodward and T. L. Isenhour, "Computer Controlled Television Scan System for Direct Coding of Chemical Structure Models", *Anal. Chem.*, **46**, 422–426 (1974).

# Computer Control of Data Acquisition, Reduction, and Display in Rapid Scanning Liquid Chromatography†

MARK S. DENTON, BARBARA A. JEZL, W. R. HEINEMAN, and HARRY B. MARK, JR.*
Department of Chemistry, University of Cincinnati, Cincinnati, Ohio 45221

Received June 16, 1977

An outline of the on-line computer system and supporting software used in a preliminary study on utilizing a rapid scanning spectrometer as a liquid chromatography (RSS/LC) detector is given. Refinements of this system in data acquisition, reduction, and display that were made out of convenience or necessity are then described in more detail. The present system allows computer control of the RSS, variable CATing, automatic absorbance calibration, rapid plotting of absorbance ($A$) vs. wavelength vs. time (3D chromatograms) and $A$ vs. time (conventional chromatograms), and, finally, an automated integration routine for quantitative measurements.



## INTRODUCTION

In a preliminary paper, the feasibility and advantages of an oscillating mirror rapid scanning ultraviolet–visible spectrometer as a detector for liquid chromatography were demonstrated.[1] It was felt that the optimum spectrometric detector for liquid chromatography (LC) would monitor the entire UV–visible region throughout a chromatogram, thereby facilitating qualitative and quantitative measurements. The coupling of an oscillating mirror rapid scanning spectrometer (OMRSS) with a high-performance liquid chromatograph has proven quite successful in moving toward meeting this objective. The emphasis of this paper was on the actual experimental techniques used, the RSS/LC system, and the resulting chromatograms from a standpoint of hardware capability. This paper describes the computer system and supporting software required for practical application of this spectrometer as a routine detector.

The marriage of rapid scanning spectrometry and liquid chromatography led by necessity to computer interaction for two reasons: first of all, for on-line experimental control and storage of the massive amounts of data generated, and, secondly, for post-run data analysis and display. A description of the minimum capabilities needed for computer interaction in RSS/LC will be made, followed by an outline of what was added or refined out of necessity and convenience.

## PRELIMINARY STUDY

**Experimental.** A Raytheon 704 minicomputer with 16K memory and 16 bit word length was used for all RSS/LC studies. The peripheral devices include the following: two direct memory access (DMA) Raytheon magnetic tape drive units, a cartridge disc unit for all driver routines and LC programs, a high-speed paper tape reader, a teletype or Tektronix T-4002 graphic computer terminal, a Tektronix 601 memory oscilloscope, an 8 channel multiplexed ADC and 4 channel DAC, and X–Y recorder display units for plotting three-dimensional (3D) chromatograms.

A second generation Harrick rapid scan spectrometer (RSSB) was used for all of the preliminary studies. This instrument and the chromatographic system used were described previously.[1]

**Data Acquisition.** In the original system, the Raytheon 704 triggered a scan of the RSSB galvanometer mirror (GM) each second (while not the upper limit of repetition rate, this is more than sufficient for most LC applications). The computer merely caused a transistor to short the trigger input of the Harrick RSS signal processing module to ground using a +10-V pulse. Control of GM during the scan itself, however, was done by the Harrick processing module. The absorbance signal from this same module was then amplified 5× by standard circuitry and inputed to the ADC by the Raytheon. Initially a separate baseline spectrum was taken prior to each chromatogram. This greatly increases the number of runs necessary, and, to avoid this, the first 20–100 scans of a chromatogram (solvent background) were averaged and used as a baseline. This will obviously not work if the retention time of any compound is very small. The software routines were, therefore, altered so that an averaged baseline spectrum is taken only once for a particular solvent system and is then automatically subtracted from each spectrum. More will be said along these lines in later sections.

**Data Reduction.** The original software for data manipulation and display has been described previously.[2,3] The software for data reduction was limited to routines for averaging baseline spectra and subsequent subtraction before plotting spectra on a X–Y recorder.

**Data Display.** Examples of the first 3D liquid chromatograms can be found in Figures 2–5 of ref 1. Each of the 3D plots represent 3–5 h of work with the initial software. The procedure involved the use of two mag-tapes and several teletype commands. One mag-tape held the stored data while the averaged baseline was written on the second tape. This baseline was then subtracted from the data and stored in core ready for plotting. Usually 1000–1500 spectra were taken each run, and this procedure had to be repeated for each spectrum that was to be plotted.

† Symposium on Computer Assisted Chemical Research Design—Joint United States Japan Cooperative Science Program, Washington, D.C., Aug 1976.

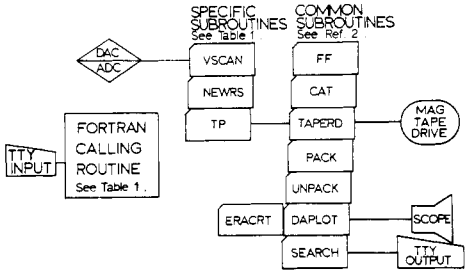


**Figure 1.** RSS/LC software block diagram.

The single wavelength plots shown in Figures 3–5 of ref 1 were an even more time consuming and difficult procedure. At the time, these were plotted out point-by-point by hand. There was also a considerable degree of uncertainty in these plots since peaks of lower retention times often obscured later peaks. For RSS/LC to become a routine analytical method rather than a novelty, it was obvious that the new software proposed in the preliminary paper had to become a reality.

## REFINEMENTS OF SYSTEM

This section will be limited for the most part to software changes as the hardware modifications of the electronic and optical systems have also been quite extensive.[4] There has been two basic hardware changes worthy of mention at this point. First, the RSS is completely compatible with any commercial or component high-performance liquid chromatograph. Second, since we are undertaking types of chromatography other than ion exchange, we now employ an 8-$\mu$L flow cell rather than the original 87-$\mu$L cell. The larger cell was completely adequate for the initial ion exchange separations of nucleotide derivatives[1] since the detector volume was less than 10% of the LC peaks of interest in all cases.[6] A smaller volume cell has become necessary for the high-performance, reverse-phase separations we are doing on dyes, vitamins, and polyaromatic hydrocarbons. Table I lists and describes each of the new routines and subroutines in the RSS/LC software. How these interact with the overall system and each other is illustrated in Figure 1.

**Data Acquisition.** The changes made in the software designed for collection of data (VCAT and VSCAN of Table I) were made, first of all, out of necessity as a changeover was made to a specially modified first generation RSS (RSSA)[4] and, secondly, out of a desire to have the capability of a variable CATing routine. With the RSSA, the Fortran calling routine (FCR), VCAT, calls upon VSCAN which, not only signals the galvanometer to start a scan, but also steps it through a preset number of positions (number of data points, NPT) set by a teletype command. VSCAN accomplishes this task by simultaneously outputting a variable sawtooth (−10 to +10 V) wave form on the DAC and acquiring one channel of spectral data on the ADC and storing it on magnetic tape. This procedure is repeated until the designated number of scans has been made. Each scan cycle is stored as a data buffer and the total number of cycles requested as a data file. Unique to VCAT is the ability to signal average, for example, 10 spectra at 10 spectra/s and store this as one CATed spectrum, rather than the usual storage of one unaveraged spectrum/s. Any wavelength region from 200 to 930 nm and any wavelength range up to 730 nm can be achieved by setting a variable parameter (variable sawtooth). By setting yet another variable parameter (32767 to 22 $\mu$s, time between points, TBP) along with the number of points, one can easily select the scan rate. The repetition rate is set by a combination of all of the above; NPT, TBP, the number of spectra CATed, and the number of spectra stored on mag-tape. For 1000 points of data, the maximum rep-rate is 12.8 spectra/s and the maximum scan rate is 45.5 spectra/s. At this time, these maxima are totally limited by the rate in which the data can be CATed and dumped onto mag-tape, not by the spectrometer itself.

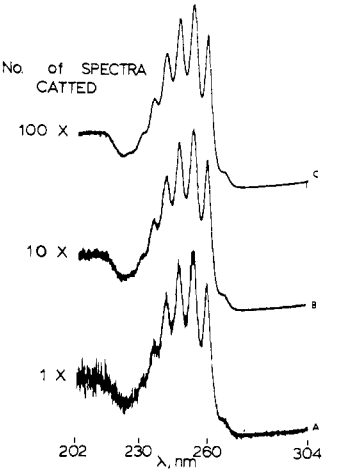


**Figure 2.** Benzene in hexane: effect of CATing. (A) 1 spectrum taken, VCAT not used; (B) 10 spectra, signal averaged; (C) 100 spectra, signal averaged.

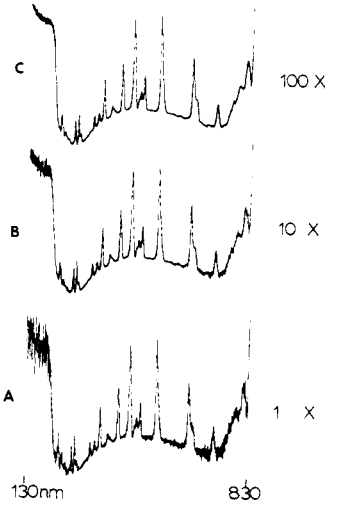


**Figure 3.** $Ho_2O_3$ in perchloric acid: wavelength standard. (A) 1 spectrum taken, VCAT not used; (B) 10 spectra, signal averaged; (C) 100 spectra, signal averaged.

The design of the new Fortran calling routine, VCAT, had several implications for our system. First, it takes more core space since it is a more complex routine and it uses double precision numbers. Second, the original triangular wave form was to be replaced by a sawtooth wave. Previously on the rising edge of the triangular wave, up to 500 data points could be taken, and another 500 points taken on the falling edge. The two resulting 500 point spectra are mirror images, but are not superimposable owing to hysteresis effects. Half of every spectra was unusable, and as a result, the resolution greatly suffered. The sawtooth has demonstrated its worth by thus increasing the resolution, and by simplifying the software used to retrieve data from a mag-tape. Finally, VCAT will enhance the signal-to-noise ratio and even make otherwise unusable data qualitatively and quantitatively meaningful. Especially note the shorter wavelengths of spectra in ref 1 and Figure 2A of this paper for examples of spectra done without this routine. Figure 2 generally demonstrates its utility on spectra of benzene in hexane. Figure 3 is a similar example which also illustrates the wide wavelength range obtainable thus far in RSS/LC only by this system.

**Data Reduction and Display.** Preceding each set of routine daily chromatographic runs, one first records spectra of air

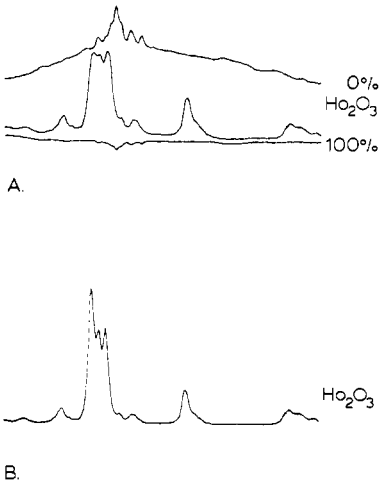


**Figure 4.** (A) Intensity spectrum of $Ho_2O_3$ filter (note distortion of spectrum at wavelengths where the Xe arc has strong emission peaks). (B) Absorbance spectrum illustrating automatic absorbance calibration routine (1 V = 1 au)

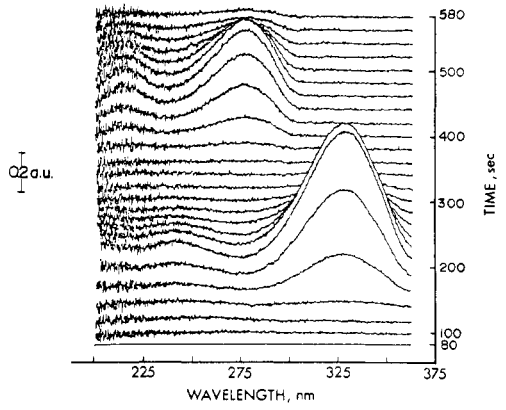


**Figure 5.** 3D chromatogram for separation of 4-thiouridine ($\lambda_{max}$ 245 and 331 nm) and cytidine·0.5$H_2SO_4$ ($\lambda_{max}$ 211 and 280 nm): Not signal averaged; Aminex A-4 cation-exchange resin, 10-cm column, 7-mm o.d. × 2.6 mm i.d.; solvent 1 M $NH_4Cl$/0.1 M HCl at 1.0 mL $min^{-1}$; repetition rate 1 spectrum/s; 940 data points/spectrum; baseline, 80 signal-averaged spectra.

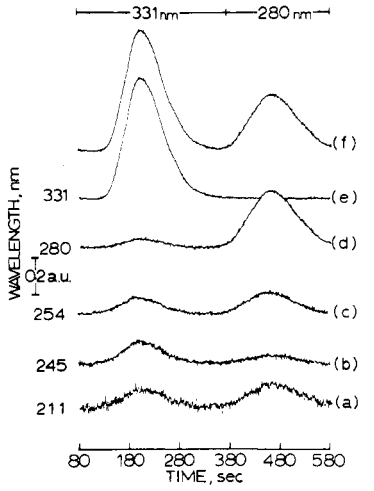


**Figure 6.** Single wavelength ($A$ vs. $t$) chromatograms from the 3D plot in Figure 5; 500 data points/spectrum: (a) 211 nm, (b) 245 nm, (c) 254 nm, (d) 280 nm, (e) 331 nm, (f) combination 331 and 280 nm.

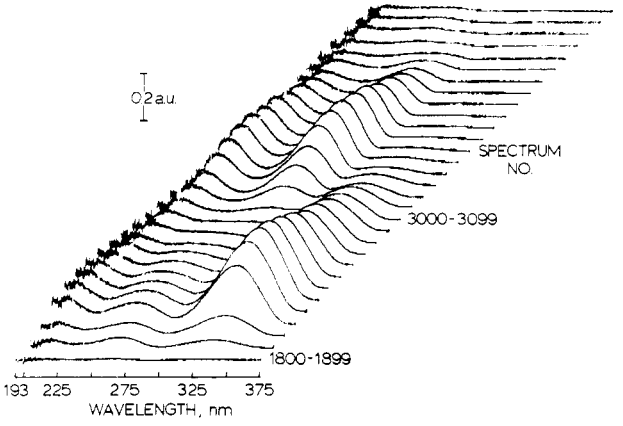


**Figure 7.** 3D chromatogram for separation of 4-thiouridine ($\lambda_{max}$ 245 and 331 nm) and cytidine·0.5$H_2SO_4$ ($\lambda_{max}$ 211 and 280 nm): Aminex A-4 cation-exchange resin; 4-cm column, 7 mm o.d. × 2.6 mm i.d.; solvent 1 M $NH_4Cl$/0.1 M HCl at 1.0 ml $min^{-1}$; repetition rate 10 spectra/s; 1000 data points/spectrum. A total of 100 spectra were signal averaged and baseline subtracted before plotting each spectrum; i.e., each spectrum represents 10 s of the chromatogram; baseline: 1200 signal-averaged spectra.

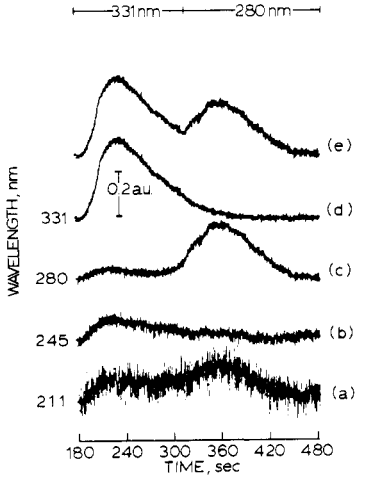


**Figure 8.** Single wavelength chromatograms of the raw 3D data.

vs. air (or solvent vs. solvent) as 100% intensity ($I$) or the baseline. Then, blocking the sample PMT, 0% or the Xe arc background emission is recorded. A wavelength standard (Hg emission lines or holmium oxide solution) is then recorded, and finally, the data, which must necessarily lie between 0 and 100% $I$, is stored. This procedure allows later conversion from intensity to absorbance spectra, thereby eliminating the necessity for an absorbance standard (e.g., $K_2CrO_4$). For the details of this routine see ref 5. The routine, NUHS, with its subroutine, NEWRS, uses this method to rapidly skim through intensity tapes and temporarily convert the spectra to absorbance vs. wavelength, baseline corrected plots on a storage scope; see Figure 4. Once the operator is certain that the run is of value, the intensity spectra can be converted and permanently stored on a separate mag-tape by the routine TABS and its subroutine TP.

The two Fortran-calling routines that allow rapid plotting of the 3D chromatograms are SKIM (absorbance vs. wavelength) and MPIC (absorbance vs. time). The utility of SKIM can be seen in Figure 5, a 3D chromatogram of two nucleotide derivatives. As opposed to the extremely time-consuming original plotting software, this plot represents an investment of only 20–30 min. Figure 6 demonstrates the power of MPIC to plot any number or combination of single wavelength spectra (conventional chromatograms). Spectrum f shows how the plotting of a combination of optimum single wavelengths at the $\lambda_{max}$ of each component enhances sensitivity in qualitative and quantitative measurements. As there are spectral differences between the two components, a degradation of chromatographic resolution (see Figure 7) could be tolerated

**Table I.** Routines in RSS/LC Software

INTENSITY TAPE: SAMPLE($I_S$)-REFERENCE($I_R$)

| ROUTINE | | FORTRAN(F) ASSEMBLER(A) FUNCTION |
|---|---|---|
| R2DAM | (F) | FORTRAN CALLING ROUTINE (F,C,R,) FOR PLOTTING A SINGLE λ FROM ABSORBANCE TAPES. ABS. VS. TIME. |
| R2DAS | (F) | F,C,R, FOR SKIMMING THROUGH AND PLOTTING ABS. TAPES. A VS. λ, |
| NUHS | (F) | F,C,R, FOR SKIMMING THROUGH INTENSITY TAPES (CHANGES TEMPORARILY TO ABS.). A VS. λ. FOR FAST VIEWING OF DATA JUST ACQUIRED TO DETERMINE VALUE OF RUN. |
| NEWRS | (F) | SUBROUTINE OF NUHS FOR CONVERTING I → A. |
| MPIC | (F) | F,C,R, FOR SKIMMING THROUGH AND PLOTTING SINGLE WAVELENGTHS FROM ABS. TAPES. A VS. T. CAN BE USED WITH I TAPES WHICH HAVE BEEN CORRECTED FOR BASELINE AND CONVERTED TO A TAPES. MULTI-RUNS WITH SINGLE BASELINE ENTRY POSSIBLE. |
| R2DAI | (F) | F,C,R, INTEGRATION ROUTINE FOR ABS. TAPES. |
| TABS | (F) | F,C,R, FOR CONVERTING I TAPES TAKEN DIRECTLY FROM THE RSS TO ABSORBANCE (BASELINE CORRECTED) TAPES. |
| TP | (A) | SUBROUTINE OF TABS FOR STORING CONVERTED I → A DATA ON MAG TAPES. |
| VCAT | (F) | DATA ACQUISITION: COMPUTER AVERAGED TRANSIENTS ROUTINE (F,C,R,) FOR INTENSITY DATA. |
| VSCAN | (A) | SUBROUTINE OF VCAT FOR OUTPUTTING A VARIABLE SAWTOOTH WAVE FORM ON THE DAC AND COLLECTING DATA ON THE ADC. |
| SKIM | (F) | R2DAS WITH BOTH BASELINE OPTIONS. (1) ADD BASELINE TO CORE. (2) 0.0 FOR BASELINE (BASELINE ALREADY SUBTRACTED FOR I → A CORRECTED SPECTRA). |
| INTG | (F) | R2DAI WITH BOTH BASELINE OPTIONS. ABS. TAPES OR CORRECTED I → A TAPES. MULTI-RUNS WITH SINGLE BASELINE ENTRY. |

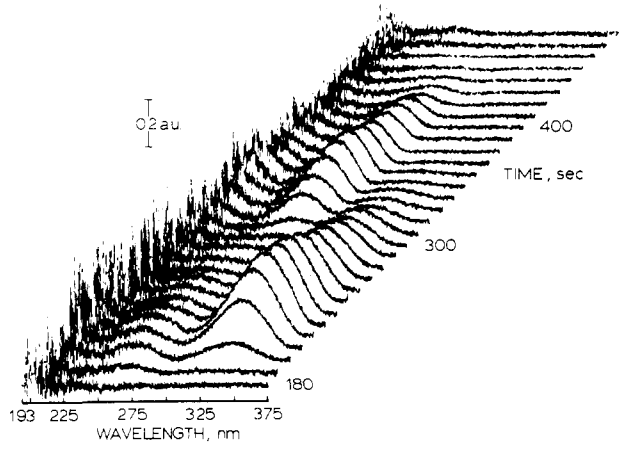


**Figure 9.** Same conditions as Figure 7 except spectra are plotted every 10 s without signal averaging.

since useful information could still be obtained from single wavelength plots (Figure 8). The tailing found here is a result of overloading the 4-cm column. Figure 9 is the same chromatogram run under particularly noisy conditions and plotted without signal averaging. It should be obvious from this example that the signal-to-noise ratio (S/N) can be greatly enhanced by using this technique. Even the high-frequency noise at low wavelengths in Figure 7 could be eliminated by filtering the ... RSS signal prior to the ADC. Figures 10 and 11 demonstrate the fact that, even with little or no chromatographic resolution, a 3D chromatogram gives, in one run, all the information needed to determine which single wavelengths are useful for qualitative identification of the components even where there is only a single spectral difference. Figures 10A and 10B further demonstrate the possibility of plotting the chromatogram at any slope, which facilitates the viewing of peaks obscured by more major peaks at higher wavelengths. Finally, Figure 11B illustrates how one can

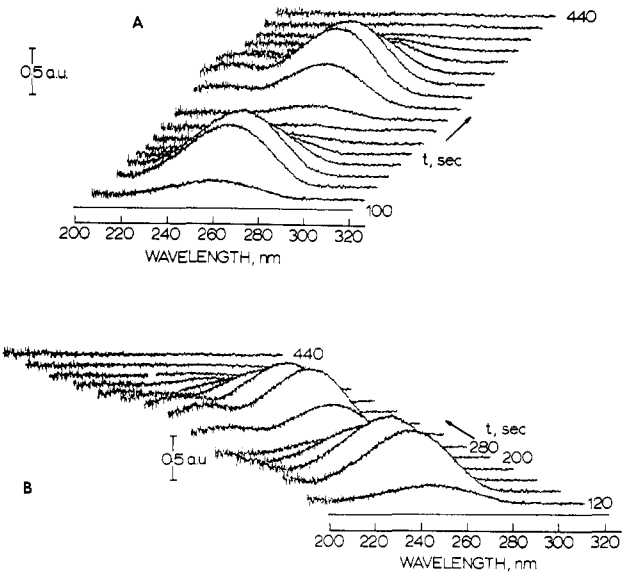


**Figure 10.** 3D chromatogram for separation of uracil ($\lambda_{max}$ 260 nm) and xanthine ($\lambda_{max}$ 231 and 260 nm): no signal averaged; Aminex A-4 cation-exchange resin; 10-cm column, 7 mm o.d. × 2.6 mm i.d.; solvent 1 M $NH_4Cl$/0.1 *M* HCl at 1.0 ml $min^{-1}$; repetition rate 1 spectrum/s; 940 data points/spectrum; baseline; 20 signal-averaged spectra. (A) Right 45° angle view; (B) left 27° angle view.

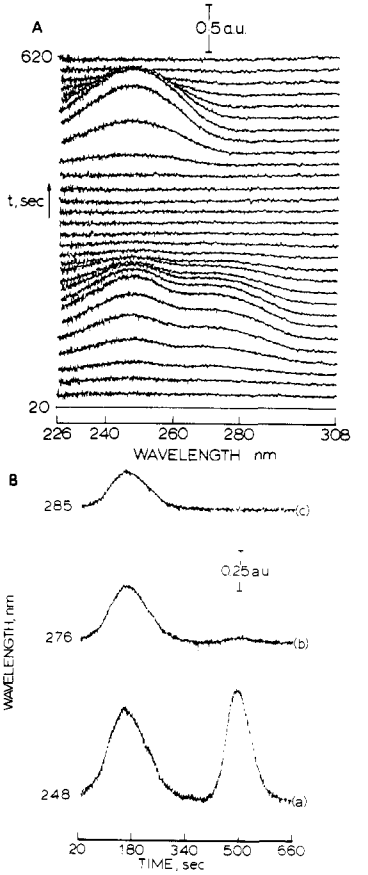


**Figure 11.** (A) 3D chromatogram for separation of guanine ($\lambda_{max}$ 248 and 276 nm) and hypoxanthine ($\lambda_{max}$ 248 nm): not signal averaged; Aminex A-4 cation-exchange resin; 10-cm column, 7 mm o.d. × 2.6 mm i.d.; solvent 1 M $NH_4Cl$/0.1 M HCl at 1.0 ml $min^{-1}$; repetition rate: 1 spectrum/s; 940 data points/spectrum; baseline, 20 signal-averaged spectra. (B) Single wavelength chromatograms taken from the 3D plot; ~640 data points/spectrum: (a) 248 nm, (b) 276 nm, and (c) 285 nm (hypoxanthine eliminated).

rapidly select the optimum wavelength to monitor a single component without the interference of another. This could have significant implications in many fields of analysis.

All of the software described above has been for qualitative determinations. INTG is a simple routine for doing quantitative measurements within these 3D chromatograms. The operator selects the wavelength and the time span of interest and the routine drops perpendicular lines to the baseline and integrates the peak.

## CONCLUSIONS

The use of a minicomputer to control data acquisition, reduction, and display in RSS/LC has come a long way from being a difficult and time-consuming task to a relatively simple and routine one. Although additional software ideas exist to further simplify the technique, a working system is at hand, and some of the unlimited applications available can now be studied in detail.[4] Investigations presently underway include separations of vitamins, dyes, and polyaromatic hydrocarbons.

## ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support provided by National Science Foundation Grants GP-35979, CHE76-04321 (H.B.M.), CHE74-02641 (W.R.H), and a U.C. Graduate Council Scholarship and Twitchell Fellowship (M.S.D.).

## REFERENCES AND NOTES

(1) M. S. Denton, T. P. DeAngelis, A. M. Yacynych, W. R. Heineman, and T. W. Gilbert, "Oscillating Mirror Rapid Scanning Ultraviolet-Visible Spectrometer as a Detector for Liquid Chromatography", *Anal. Chem.*, **48**, 20–4 (1976).
(2) A. M. Yacynych, Ph.D. Thesis, University of Cincinnati, Cincinnati, Ohio, 1975.
(3) H. B. Mark, Jr., R. M. Wilson, T. L. Miller, T. V. Atkinson, H. Wood, and A. M. Yacynych, "The On-Line Computer in New Problems in Spectroscopy: Applications to Rapid Scanning Spectroelectrochemical Experiments and Time Resolved Phosphorescence Studies" in "Information Chemistry; Computer Assisted Chemical Research Design", S. Fujiwara and H. B. Mark, Jr., Ed., University of Tokyo Press, Tokyo, 1975, pp 3–28.
(4) M. S. Denton and T. W. Gilbert, "Rapid Scanning Liquid Chromatography Detector: Instrumentation, Software and Applications", manuscript in preparation.
(5) A. M. Yacynych and H. B. Mark, Jr., "Automatic Absorbance Calibration Routine for a Computerized UV-Visible Rapid Scanning Spectrometer", *Chem. Instrum.*, in press.
(6) M. S. Denton and T. W. Gilbert, *J. Chromatogr.*, manuscript in preparation.



# Interactive Pattern Recognition in the Chemical Analysis Laboratory†

CHARLES L. WILKINS
Department of Chemistry, University of Nebraska—Lincoln, Lincoln, Nebraska 68588

Received June 16, 1977

A proposed interactive organic structure analysis system is described. This system based on the use of an MS-5076 ultra-high-resolution mass spectrometer equipped with both chemical ionization and electron impact sources is proposed to be used in the exploration of a variety of interactive pattern recognition studies aimed toward the development of methods intended to facilitate more data analysis on the laboratory minicomputers which are central elements of the system. An approach whereby a gas chromatograph–infrared interferometer would be linked to the mass spectrometer system and the combined information used as the source of structural inferences is discussed. In particular, the use of factor analysis, simplex pattern recognition, digital learning networks, and search methods are considered. It is suggested that information-theory-based-evaluation methods should be used in selection of the optimum solutions to problems encountered.

During the past few years we have been engaged in research directed toward implementing pattern recognition methodologies in an on-line fashion using laboratory computers. For that reason, we have primarily concentrated on development of methods which would lend themselves easily to adaptation to the laboratory framework. Effectively this has constrained somewhat the variety of pattern recognition techniques which we can realistically consider using. For the present, I will first focus on a proposed chemical analysis system we are in the process of developing and use that as an introduction to discussion of a number of our more recent research efforts. Figure 1 is a block diagram of the proposed analysis system hardware. This diagram contains both elements which are already installed in our laboratory and those which we hope to add in the reasonably near future. As is seen in the upper middle of the diagram, a central element of this structural analysis system is a high-resolution mass spectrometer, the AEI MS-5076 Ultra-High Resolution Mass Spectrometer, which is equipped with both electron impact and chemical ionization sources. In addition, the spectrometer is interfaced both to a gas chromatograph (for rapid mass spectrometric scans of mixture components after separation) as well as a data acquisition and control system which employs a Nova 2/10 computer equipped with 32K 16-bit words of main memory, a cathode ray tube terminal for system control, a 2.5 million word magnetic disk and a rapid electrostatic plotter-printer. This data acquisition system also services a Hitachi RMU-6D medium resolution mass spectrometer which is equipped with a field ionization/field desorption source and an electron impact source as well. Enclosed in the dotted lines in the diagram is the proposed addition to the analysis system. By adding this on-line interferometer-based infrared spectrometer, acquisition of infrared spectra of the same gas chromatographic effluents as are currently routed to the MS-5076 spectrometer would be possible. Because of the need for rapid data reduction, the proposed system would incorporate its own dedicated computer, a rapid plotter, floppy disks for intermediate data storage, and an operator console. As currently visualized, the system would be so constructed that stand-alone operation of the GC–IR system would also be possible. Finally, note the planned link between the mass spectrometer–computer control system and the infrared spectrometer computer. This would be a high-speed digital link which would allow use of the larger system peripherals as well as the transfer of reduced

† Presented at the Joint U.S.–Japan Seminar on "Computer-Assisted Chemical Research Design", Aug 16–20, 1976, Washington, D.C., sponsored by the National Science Foundation and the Japan Society for the Promotion of Science.