On-Line Computers in Research. High-Resolution Mass Spectrometry

RENGACHARI VENKATARAGHAVAN, RICHARD J. KLIMOWSKI, AND FRED W. MCLAFFERTY

Department of Chemistry, Cornell University, Ithaca, New York 14850

Received October 20, 1969

Experimentation in chemistry has been revolutionized in the last two decades through the direct application of sophisticated instrumentation, such as spectrometers, chromatographs, and electronic devices, for measuring, recording, and controlling (Figure 1). A new general development which promises to cause an equivalent impact is the incorporation of the high-speed digital computer *into the experimental system* (Figure 2). Such "on-line" operation involves the direct interfacing (coupling) of the computer to the instruments or other devices sensing the experimental outputs (results and conditions) to acquire and process the data, to return control signals to the experiment, and to correlate, display, and record the results in the form which is most meaningful to the experimenter. All of this is accomplished in "real time" (during or shortly after completion of the experiment). The data system not only relieves the experimenter of the tedium of observing, calculating, and recording the results, but it can yield impressive improvements in accuracy, speed, and other aspects of performance, often permitting experiments which otherwise would be impossible.[1]

Complete computer-interface systems designed to handle a variety of instruments simultaneously cost $200,000 or more, an expense that has seriously limited application in most chemical research laboratories. The recent development of the small but fast computer has made possible systems which are tailored for specific applications, such as gas chromatography, nuclear magnetic resonance spectroscopy, mass spectrometry, and X-ray crystallography, for $40,000–100,000. However adequate systems can be constructed for considerably less, as the computers alone cost only $8,000 to $25,000. Indeed, we feel that for many experiments the chemist should design, assemble, and/or modify his own hardware (equipment) and software (program instructions); the incentives are similar to those which have induced chemists to become familiar with the operation and maintenance of specialized instruments and the interpretation of various types of spectra.

A number of reports have appeared recently illustrating the awakening interest in this field. Perone and coworkers[2] employ a small digital computer on-line to a stationary electrode polarograph. The computer continuously monitors the experiment, performs real-time calculations on the output, and supplies feedback signals to set the interrupt potential and delay time between voltage sweeps. Pardue and his students[3] utilize a small digital computer for on-line processing of reaction rate data for quantitative analysis; preliminary reaction rate calculations are used to optimize data acquisition rates and other operating parameters. Several concurrently operating gas chromatographs can be connected on-line to a digital computer for data reduction to yield improved accuracy and resolution.[4] To illustrate the nature and merits of such systems we will describe some applications of computers to high-resolution mass spectrometry.[5–8]

Information Available from Modern Instrumental Techniques. A basic reason for the usefulness of spectroscopic and chromatographic techniques is that they yield two dimensions of information: the ordinate shows a characteristic response (for example, absorptivity or thermal conductivity) at each of a number of positions on the abscissa (for example, wavelength or retention time). The amount of information about a sample in its spectrum or chromatogram thus depends on the dynamic range over which the ordinate response can be measured and the number of positions (resolution increments) that are distinguishable on the abscissa. By this definition of information content high-resolution mass spectrometry far surpasses other molecular techniques such as infrared spectroscopy, nuclear magnetic resonance spectroscopy, and gas chromatography. Although sample requirements are <1 μg, a molecule of molecular weight 1000 will typically exhibit more than 100 peaks in its mass spectrum. These peaks represent the masses of the molecular and fragment ions formed from the sample by electron bombardment. The mass spectrum is a *line* spectrum; because the mass of an ion is defined exactly by its elemental composition, the essential limit of the narrowness of a peak is the resolving power of the instrument. By measuring the mass of an ion peak with an accuracy of 1 mmu (millimass unit) there are approximately 10^6 mass locations at which each of the peaks of a mass spectrum can be located, and the abundance of each of these can be measured over a dynamic range of $>10^4$; this compares to $<10^2$ resolution increments and $<10^3$ dynamic range for most other methods.

(1) Applications of computers in chemical experimentation have been discussed recently by J. W. Frazer, *Anal. Chem.*, **40** (8), 26A (1968); G. Lauer and R. A. Osteryoung, *ibid.*, **40** (10), 30A (1968).

(2) S. P. Perone, D. O. Jones, and W. F. Gutknecht, *ibid.*, **41**, 1154 (1969).

(3) G. E. Jones and H. L. Pardue, *ibid.*, **41**, 1618 (1969).

(4) A. W. Westerberg, *ibid.*, **41**, 1595, 1770 (1969).

(5) J. H. Beynon, "Mass Spectrometry and its Application to Organic Chemistry," Elsevier Publishing Co., Amsterdam, 1960.

(6) K. Biemann, *Advan. Mass Spectrom.*, **4**, 139 (1968).

(7) F. W. McLafferty, *Science*, **151**, 641 (1966).

(8) Computer–mass spectrometry systems are available from Perkin-Elmer, Inc., Norwalk, Conn.; Picker-Nuclear Corp., White Plains, N. Y.; and Varian Associates, Palo Alto, Calif.

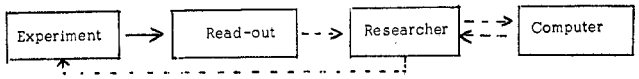

Figure 1. Schematic representation of the use of instrumental techniques to sense the conditions and results of the experiments. The researcher observes these, computes more meaningful results where necessary, and uses the results to modify the experiment.

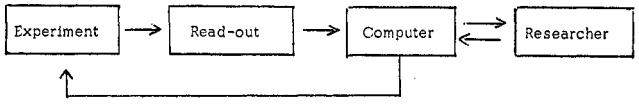

Figure 2. The on-line computer can assume the routine functions of data acquisition, data reduction, and control of the experiment. The research man makes high-level decisions and communicates these to the computer for execution.

One of the most serious handicaps in utilizing this tremendous amount of information is the great effort necessary to measure the spectral data and from this to calculate the exact masses and abundances. The elemental composition of each ion peak can be assigned uniquely if the mass can be determined with sufficient accuracy (Table I). Accuracy of mass measurement is paramount for larger molecules, as the number of possible elemental combinations which will yield a particular nominal mass (±0.5 mass unit) goes up exponentially with molecular weight. For example, chemically logical (including isotopic) combinations of C, H, N, and O yield 12 common ions of nominal mass 43, but 500 ions of nominal mass 430.

Exact mass measurement in most instruments is done by "peak matching." Images of the unknown peak and a selected reference peak are displayed alternately on an oscilloscope by a rapid scan over a small range of the ion accelerating voltage; the mass difference of these peaks is determined by the exact differential voltage necessary to make the peaks coincide. Although accuracies of 1–10 ppm are attainable by this method, it requires 1–5 min/peak. One seldom determines more than a few peaks per spectrum by this method.

Another serious problem is operation of a complex mass spectrometer to obtain a maximum amount of useful information with a minimum amount of sample. The operator must control many instrumental parameters simultaneously, such as rate of sample volatilization, ion focusing, and slit settings to maximize resolution or sensitivity. Special measurements, such as for metastable ions utilizing a defocused electrostatic analyzer,[9] require additional adjustments, as will be discussed later. The operator must control these parameters from experience or quick calculations based on information from instrument meters and the spectral recorder; such feedback obviously has severe limitations in speed and accuracy.

Photoplate-Recorded Mass Spectra

In a mass spectrometer with Mattauch–Herzog geometry all of the ions over a wide mass range (for example, m/e 25–700) can be focused simultaneously on a photographic plate. In this way the complete spectrum can be recorded in minutes. However, manual measurement of the distances between the centers of the ion lines of a complex spectrum requires several hours with an accurate comparator. To reduce the time, tedium, and human errors associated with these measurements an automatic comparator–microdensitometer was designed and constructed.[10] In its operation a light beam defined by a narrow vertical slit is focused on the photoplate to produce an image with the same orientation as the ion lines. The photoplate spectrum is moved through the light beam, and the transmission of light is measured by a photomultiplier at 0.25-μm intervals along the 330-mm length of the mass axis. These analog transmisson (T) values are transformed with an analog-to-digital (A/D) converter to digital values containing 3 significant decimal figures. To conserve storage space, only $(1 - T)$ values above a preset threshold are saved, along with a value from a shaft position encoder on the drive screw to define the displacement of the photoplate. Only 12 min is required to measure and threshold the 10^6 data points and store the significant information.

Table I
Common Ions of Nominal Mass 43

Composition	Typical structure	Exact mass[a]
CHNO	–CONH–	43.0058
C_2H_3O	CH_3CO–	43.0184
CH_3N_2	$CH_3N{=}N$–	43.0296
C_2H_5N	$-CH_2NHCH_2-$	43.0421
C_3H_7	$(CH_3)_2CH$–	43.0547
C_3H_5D		43.0532
$^{12}C_2{}^{13}CH_6$		43.0502

[a] Sum of the exact nuclidic masses of the atoms based on $^{12}C = 12.00000$.

Improvement in Mass Measuring Accuracy. Not only does this system shorten dramatically the time of data acquisition, but the results are much more detailed. These data provide information concerning the distribution of ions in each individual line; experimentally this corresponds closely to a Gaussian function (Figure 3). To find the center of an ion line, the computer data-reduction program smooths the corresponding transmission data, fits the results by a least-squares procedure to a Gaussian function, and calculates the center for this function. The observed precision of this dynamic measurement method is ±0.25 μm, which is far superior to the precision of >1 μm found for the static mechanical or optical analog methods used in manual operation of the comparator. Optimization of this method has reduced average mass measuring errors from ±2 or 3 to ±0.5 mmu.[11]

Improvement in Resolving Power. The ability to resolve ion peaks of nearly the same mass is one of the most important specifications for a high-resolution mass spectrometer. If two or more ion lines overlap on the

(9) M. Barber and R. M. Elliott, ASTM E-14 Conference on Mass Spectrometry, Montreal, 1964, p 150; T. W. Shannon, T. E. Mead, C. G. Warner, and F. W. McLafferty, *Anal. Chem.*, 39, 1748 (1967).

(10) R. Venkataraghavan, F. W. McLafferty, and J. W. Amy, *ibid.*, 39, 178 (1967).

(11) R. D. Board, Ph.D. Thesis, Purdue University, 1969.

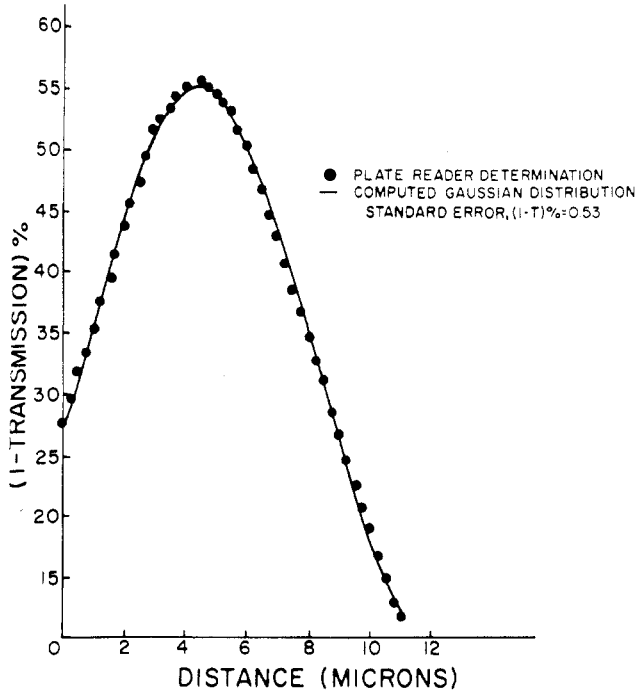

Figure 3. Experimental points from the automatic comparator-microdensitometer for a single ion line in a photoplate mass spectrum. The line is a best-fit Gaussian distribution generated by the computer.

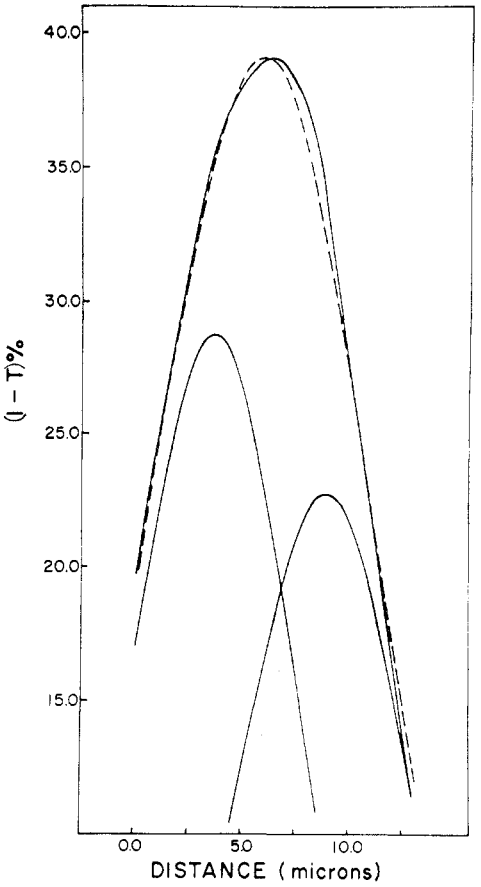

Figure 4. Deconvolution of data from overlapping ion lines ($C_{10}{}^{13}CH_9$ and $C_{11}H_{10}$, $\Delta m/e$ = 0.0045 amu). Upper curves: solid line, experimental data from comparator reading of photoplate mass spectrum; dashed line, best-fit Gaussian distribution generated by the computer. Lower curves: component ion distribution curves determined by the computer, whose measured mass difference is within 0.0004 amu of the true value. The deconvolution routine was automatically activated because the peak width at half-height exceeded specifications.

photoplate, the resulting transmission profile will represent the sum of their individual Gaussian distributions. If the computer program cannot fit the transmission data for a particular ion line to a Gaussian distribution within predetermined limits, or if it finds the width at half-height too large (see Figure 4), the program deconvolutes the peak into the component Gaussian functions by an iterative routine. Deconvolution is imperative for the comparator system, as the data are convoluted by the slit width of *ca.* 3 μm used as a compromise of resolution, sensitivity, and edge diffraction errors. Despite this, the ability of the overall photoplate system to distinguish overlapping peaks is substantially improved over the performance of the mass spectrometer in the conventional magnetic scanning mode of operation employing a $\sim$1-μm *ion* exit slit. Of even more importance is the resulting improvement in mass-measuring accuracy. A common source of error is the overlap of ^{13}C isotope peaks, as the mass difference between ^{13}C and ^{12}CH is only 4.5 mmu. A small contribution of this kind shifts the apparent centroid of the peak; even when the contribution is sufficient to cause a resolved doublet, the positions of these maxima are displaced from the true peak positions.[10,11]

Data Reduction and Display. After the basic data have been acquired by the computer, it can perform a variety of more conventional reduction operations. The computer program determines the exact positions and areas for each ion line in the spectrum, locates the lines due to perfluorokerosene which is added as a mass reference, interpolates between these using a higher order polynomial to assign exact masses to the positions of the peaks from the sample, and computes the possible elemental compositions corresponding to each mass. The data can be displayed in a variety of forms for interpretation, such as a topographical element map (Figure 5).[12] The computer can also check the data against a file of reference spectra for possible identification, or it can attempt to interpret the spectrum.[13-15]

On-Line Photoplate Data Acquisition System. Biemann and his coworkers[6] have coupled a photoplate comparator on-line to a medium-sized computer (IBM 1800). Transmission values at 0.5-μm intervals are digitized and stored in core memory. When a memory section of 1000 locations is filled, the computer starts to process these data, and the incoming data are stored in a new section. Data processing during the scan in-

(12) R. Venkataraghavan and F. W. McLafferty, *Anal. Chem.*, **39**, 278 (1967).

(13) R. Venkataraghavan, F. W. McLafferty, and G. E. Van Lear, *Org. Mass Spectrom.*, **2**, 1 (1969); A. M. Duffield, A. V. Robertson, C. Djerassi, B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum, and J. Lederberg, *J. Amer. Chem. Soc.*, **91**, 2977 (1969); and references cited therein.

(14) M. Senn, R. Venkataraghavan, and F. W. McLafferty, *ibid.*, **88**, 5593 (1966).

(15) K. Biemann, C. Cone, B. R. Webster, and G. P. Arsenault, *ibid.*, **88**, 5598 (1966).

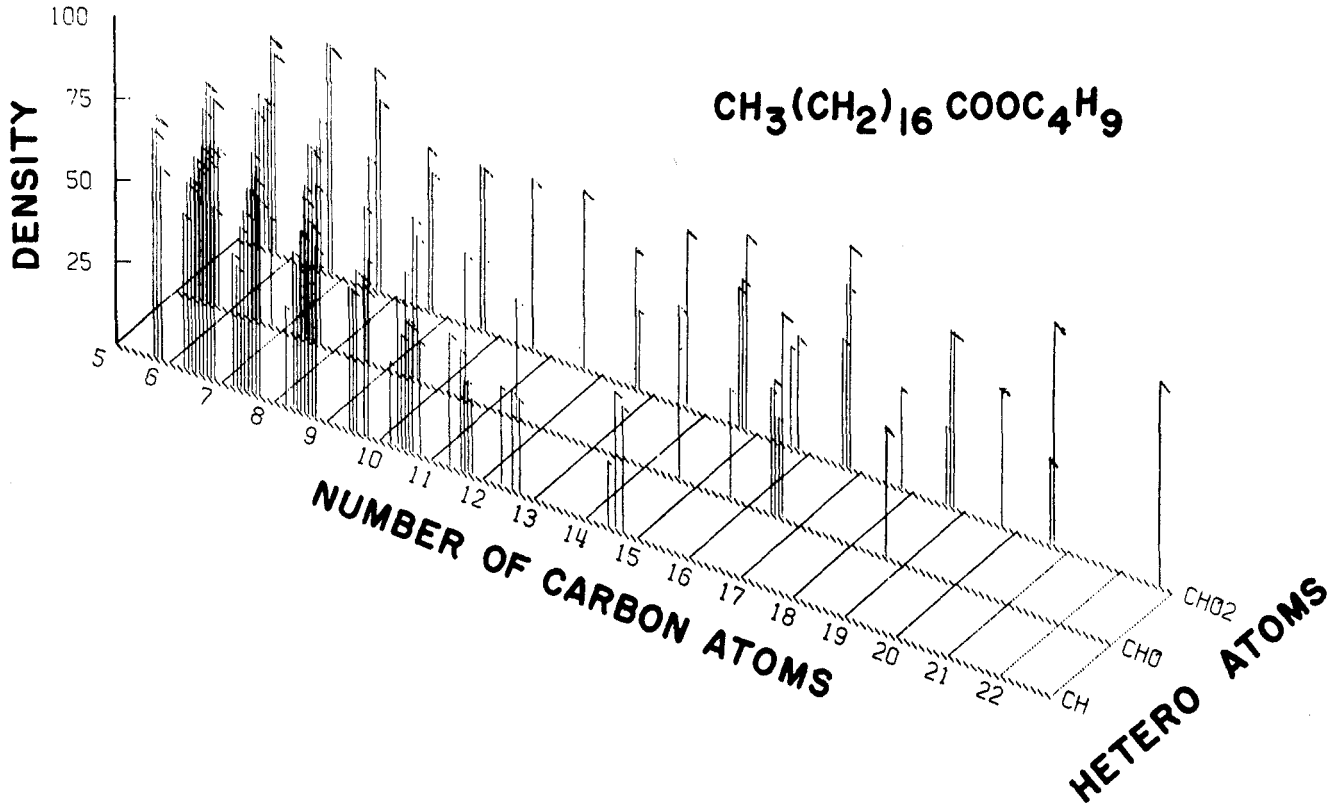

Figure 5. Topographical element map of butyl stearate. Note that the highest mass peak in the CHO row is C_{18}, indicating that four carbon atoms must be lost before an oxygen atom is lost.

cludes the calculations of background, peak centers, and simple deconvolution. As soon as the scan is completed the computer uses the stored peak data to determine the elemental compositions of all of the ions in the spectrum. Thus the reading and complete processing of the photoplate spectrum require an average elapsed time of only 20 min.

However, the delay caused by photoplate development does not make it feasible to utilize the data while the sample is still in the mass spectrometer. To achieve this, systems have been developed in which the output from the electron multiplier during magnetic scanning is processed on-line by a digital computer.

Real-Time Computing Systems

A further major reason for incorporation of a computer into the experimental system is to be able to utilize the results to optimize the experiment while it is in progress. For this either the operator can make adjustments based on the results (man in the loop), or the computer itself can be programmed to send appropriate control information to the apparatus (closed loop).

The design factors for an on-line computer system for high-resolution mass spectrometry are similar in many ways to those encountered in systems for a wide variety of chemical experiments, despite the fact that some requirements, such as data rate and data storage, may vary widely. It is imperative that the experimenter himself becomes closely involved in the design; in most successful systems the computer "looks at" the process and resulting data in the same way as the experimenter would. Although it may be necessary to call on experts for many technical problems of software and hardware, the success of the overall system will depend directly on how well the goals and other problems of the experiment are met.

Time-Shared *vs.* Dedicated Computers. It is usually possible to justify a larger computer and a greater number of specialized personnel if the same computer can serve several projects. This can be done by running each experiment at a separate time, with manual connection and disconnection of the computer for each run; different interface hardware as well as program software may be necessary for each. For greater convenience and flexibility a computer can service a number of experiments on a nearly simultaneous basis by "time sharing"; however the time-sharing operation itself requires substantial memory capacity and sophisticated software. Problems arise also if the computer and the experiment are separated; a sufficiently capable terminal for communication with the computer (input–output devices) must be located at the experiment, and transfer of high frequency analog data over distances >25 m is not advisable due to possible signal distortion. Transfer of digital signals often requires expensive special equipment. Therefore in designing an on-line system the small dedicated computer should be given serious consideration.

Interface Design. The interface makes possible the transmission of data from the experiment to the computer, and *vice versa*. To accomplish the former, the interface must accept one or more analog signals from the experiment and make these suitable for computer processing. The interface may include special hard-

ware devices which carry out simple high-rate operations to reduce the load on or increase the efficiency of the central processing unit, thus reducing the size requirements of the computer. The major components generally found in the interface are the signal-conditioning equipment, analog multiplexer, A/D converter, and time-base generator. The signal-conditioning equipment converts the analog output of the instrument into a form that is acceptable (in voltage level and range, amperage, etc.) to the A/D converter or that is better for computer processing (filtered to improve the signal-to-noise ratio, for example). The multiplexer can receive analog signals from a variety of sources for sequential introduction to the A/D.

The major criteria for the selection of an A/D converter are the data rate and the dynamic range of the information obtainable from the experiment. For example, with most gas chromatographs a sampling rate of 10 Hz should be sufficient to describe gc peaks adequately, even for studies of peak skewing or deconvolution, although a dynamic range of 10^6 may be necessary. On the other hand, temperature-jump experiments on a microsecond time scale require sampling rates exceeding the rates of the fastest A/D converter available commercially. In the design of a general purpose interface suitable for various instruments it is advantageous to incorporate an A/D converter with sufficient flexibility to handle data of different rates and magnitudes.

Information on the abscissal position (indicative of frequency, mass, retention time, etc.) must also be supplied to the computer with each ordinate measurement. This can come directly from the instrument being monitored; for the photoplate reader described above this was done by triggering the A/D converter at constant distance intervals sensed by a shaft encoder on the drive screw (4000 pulses/revolution). The instrument abscissa can also be driven at a reproducible rate with the triggering supplied by a time-base pulse generator. Such generators are of various precision and frequency. Crystal clocks are available with a precision of $1/10^8$. RC (resistor–capacitor) clocks are less precise, but much less expensive. The computer itself can be used to generate the time base if the demand for its attention from the experiment is not excessive.

Feedback Control. Communication from the computer back to the experiment must also go through the interface; a wide variety of experimental functions, such as heating, sample or reagent flow, pressure, slit width, position, or emergency shutdown can be placed under feedback control. Most of the special devices to perform these functions can be operated by an electrical analog signal. This is supplied by simple on–off relays or a digital-to-analog converter (D/A) in the interface, using a multiplexer if there are a number of control devices in the experiment.

Specifications of the Computer. The selection of the central processing unit (CPU) and its associated peripheral devices for a real-time data acquisition and reduction system depends on the experiment itself. The CPU of modern low-cost digital computers has a fast processor and a high-speed (0.7–2.0 μsec) random-access core; their size and speed basically determine the rate of data acquisition and the complexity of the computations that can be performed. Increasing the size of the CPU is relatively expensive, so that alternatives should be sought. Although the CPU is ideal for data storage, magnetic tape can store large volumes of data at high speeds more economically. Paper tape is much slower, but requires an even smaller investment. Tape data can be processed later by the CPU, but this processing must mainly be in a sequential fashion. If it is necessary, for example, for feedback control to compare data taken at noncontiguous times, the computer will have to carry out a time-consuming search of the tape. An excellent alternative bulk storage device is the magnetic disk (or drum). Although it is more expensive than tape for the storage of large volumes of data, all of the data stored on disk can be accessed randomly in a relatively short time (*ca.* 20 msec). Further, by using a disk to store part of the computer program for transfer ("program swapping") at appropriate times to the CPU, the processing capacity of the CPU can often be greatly increased. In essence the software can treat the disk as slow core for programming and storage purposes. Thus the optimum balance of investment in CPU, disk, and tape in the computer system depends on the rate and volume of data and how they are to be processed, stored, and retrieved.

The most common data output device is the teletype. Although relatively inexpensive, it is slow (10 characters/sec) compared to high-speed printers (300–1200 *lines*/min). For many experiments the cathode ray tube provides an elegant output device, but sophisticated software and hardware are often required to drive it.

A High-Resolution Mass Spectrometry System Incorporating a Small On-Line Computer

A number of types of systems coupling high-resolution mass spectrometers to computers have been developed.[8,16] A pioneering system is that of Burlingame and coworkers[17] which uses the powerful XDS-930 and XDS-Sigma-7 computers. The computer compresses the digitized raw data by deletion of all intensities below a preset threshold, displays the resulting data on a cathode ray tube to allow operator interaction with the system, and stores the accepted data for later computer processing. The system developed in our laboratory, described below, is similar in principle, but employs a much smaller computer which necessitates more complex hardware interfacing.[18]

(16) Reviewed by A. L. Burlingame, *Advan. Mass Spectrom.*, **4**, 15 (1968); W. J. McMurray, S. R. Lipsky, and B. N. Green, *ibid.*, **4**, 77 (1968); H. C. Bowen, E. Clayton, D. J. Shields, and H. M. Stanier, *ibid.*, **4**, 257 (1968).

(17) A. L. Burlingame, D. H. Smith, and R. V. Olsen, *Anal. Chem.*, **40**, 13 (1968); A. L. Burlingame, D. H. Smith, F. Walls, and R. V. Olsen, Proceedings of the 17th Annual Conference on Mass Spectrometry, Dallas, Texas, May 1969, p 28.

This system produces elemental composition information accurately and rapidly and controls the mass spectrometer for acquisition of metastable ion data. This also illustrates a system approach which should be very attractive for many other chemical experimentation problems.

Hardware Characteristics. A block diagram of the system is shown in Figure 6. The interface has an A/D converter which can convert the analog signal to a digital word of 12 bits (range of 1 to 4096) 20,000 times/sec (20 kHz), matching the 12-bit word length of the computer. The signal-conditioning equipment includes a logarithmic amplifier so that the larger dynamic range of the mass spectrometer (*ca.* 1/20,000) will not be reduced by the A/D, and a filter to reduce the noise level of the signal. A 20-kHz crystal clock (precision of $1/10^8$) triggers the A/D and drives a counter with a 24-bit buffer which registers the elapsed scan time. A divider switch on the clock makes lower conversion rates possible.

The interface also includes hardware logic for detection and thresholding. In the usual high-resolution mass spectrum the peaks are so far apart and narrow that significant data constitute only a few per cent of the total of the approximately 10^6 measurements emerging from the A/D converter. By allowing the transmission of only those peaks exceeding a preset threshold value, the data storage and processing requirements in the computer are reduced correspondingly. Although the CPU could discard the base-line information, to do this it would have to determine whether each datum is above or below noise level, a decision requiring 20 μsec once every 50 μsec (at a 20-kHz digitization rate). Thus the CPU can do other calculations until the interface signals that significant data have been gathered from the mass spectrometer. Further CPU operation economies are made possible by the interface because the computer does not have to trigger the A/D or update the clock register every 50 μsec, or record an abscissal (time) value with each ordinate (ion abundance) value. The latter values are taken at equivalent intervals, so that the time is only recorded once in a sequence of abundance values. The interface also incorporates a 10-bit digital-to-analog converter which enables the computer to send control signals back to the mass spectrometer.

The computer used[19] has 4096 12-bit words of core memory, the automatic multiply–divide option, teletype, and a 32,000 word random-access disk. The incorporation of the disk makes it possible, despite the small core memory, to store the relatively large amount of data from the mass scan and to process it in a continuous fashion. The disk also stores most of the

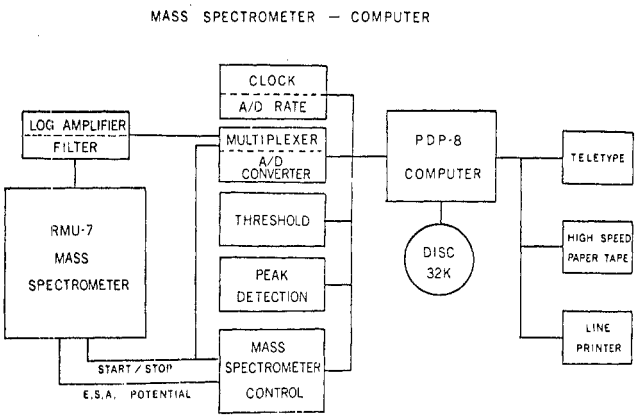

Figure 6. On-line data-system for high-resolution mass spectrometry utilizing a small (4096 core words) computer.

software, so that sections are brought into core memory as needed during the operation. The execution of different phases of the program thus is automatic and continuous. A high-speed paper tape reader-punch is convenient for development and maintenance work, but is not necessary for routine operation.

System Operation. When the start-of-scan button on the mass spectrometer is pushed, this is sensed by the computer and the clock is started. The clock in turn triggers the A/D converter continuously at the preselected rate. The threshold logic compares the digitized value of the ion current to the preset threshold requirement. When the value goes above this level, the computer is interrupted so that the succeeding values can be transmitted to it and stored. The peak detection logic of the interface checks each value and terminates collection when four consecutive subthreshold values are transmitted. The logic then generates an interrupt to read the clock value, which has been accumulating in a clock register of 24-bit capacity (14 min at 20 kHz). The value read thus corresponds to the elapsed scan time at which the last datum point was collected. Figure 7 shows typical values generated by a large peak.

The data are stored sequentially in buffer tables in the core memory. When each table is filled it is transferred as a data block to the disk as soon as there is sufficient time (*ca.* 35 msec) between peaks. This overlapping operation of data storage continues until the scan is stopped by pushing the stop button on the mass spectrometer. This signal is sensed by the computer to initiate the next data reduction step.

Software Characteristics. Effective utilization of hardware in an experimental system depends to a large degree on the organization of the software. High-level languages such as Fortran, Focal, Basic, etc., which have been developed for off-line applications are less suitable for on-line systems because the object codes generated by their compilers are not very efficient. When the processing time or number of core locations is critical, it is better to write the programs in a lower level language such as the Assembler language. Application programs for this system are written in Macro-8 along modular lines to use the full potential of the disk.

(18) R. Venkataraghavan, J. W. Amy, R. D. Board, R. D. Brown, R. J. Klimowski, and F. W. McLafferty, Proceedings of the 16th Annual Conference on Mass Spectrometry and Allied Topics, Pittsburgh, Pa., May 1968, p 114; R. J. Klimowski, Ph.D. Thesis, Cornell University, 1969; R. J. Klimowski, R. Venkataraghavan, F. W. McLafferty, and E. B. Delaney, submitted for publication.

(19) PDP-8 computer, Digital Equipment Corporation, Maynard, Mass.

```
0147  0202  0226  0239  0259  0258  0235  0222
0258  0278  0310  0370  0447  0562  0671  0742
0774  0803  0831  0890  0947  0983  1066  1183
1306  1394  1459  1514  1571  1622  1655  1690
1730  1730  1698  1694  1722  1774  1802  1819
1831  1855  1887  1890  1862  1807  1738  1675
1626  1590  1555  1526  1498  1471  1446  1407
1367  1331  1295  1242  1170  1102  1035  0979
0946  0918  0898  0870  0818  0742  0658  0578
0523  0478  0410  0338  0258  0190  0142  0098
0062  0026  0002

00059476
```

Figure 7. Typical output values from the on-line system describing the profile of a major peak, followed by a decimal value of the elapsed time.

The data reduction steps of the software are divided into five phases: (I) data acquisition; (II) peak center calculation; (III) reference peak identification; (IV) exact mass calculation; and (V) elemental composition calculation.

Before the scan is initiated, the operator answers questions posed by the computer concerning the sample, reference compound, scan speed and range, threshold, and instrument conditions. From this information the computer sets the desired operating parameters in the interface. Phase I acquires the data and stores the resultant compressed spectrum on the disk system as described above.

When the scan-stop button is pushed, the computer causes the phase II program to be "swapped" into core from the disk, *i.e.*, this program is transferred to the core locations of the phase I program. Phase II recalls the stored spectral data from the disk, rescales it from the logarithmic to a linear function, and determines the center of each ion peak and the area under its profile; these calculations usually require less than 60 sec for a typical spectrum. The last operation of the phase II program is to swap the phase III program into core. This phase establishes a time–mass relationship (dispersion curve) by automatically identifying the reference peaks using an extrapolation technique. In phase IV the time values of the sample peaks are converted to mass values. The exact mass of each peak is calculated using four reference lines with a Lagrange polynomial equation; this requires about 5 sec for the whole spectrum. In phase V all possible elemental compositions within the prescribed error tolerance and element limits are assigned to the individual ion peaks in the spectrum using an iterative routine employing valence rules.[10,20] An example of phase V output is shown in Figure 8 for the molecular ion region of caffeine.

Results

Resolution and accuracy are primary criteria of overall system performance. In Figure 9 the mass separation between the C_5H_5N and $C_5{}^{13}CH_6$ peaks is 1 part in 10,000, indicating that resolution comparable to that achieved in conventional operation of this mass spectrometer is possible with good mass measuring accuracy. The latter is heavily dependent on mass spectrometer performance, but in general accuracy limits of ±3 mmu are readily achievable in single-scan measurements at 10,000 resolution. We find that the major part of this error is due to random fluctuations in the signal, as shown by the error convergence in Figure 10 achieved on multiple scans. At 10,000 resolution the averaging of ten scans improves the accuracy of mass measurement to ±1 mmu. Similar results have been reported by Burlingame, *et al.*,[17] using a different mass spectrometer and on-line data acquisition system. Thus the computer of this system can also act as a powerful signal-averaging device.

REL.AB.	MASS	CALC.	ERR.	C12/C13	H	N	O
1.1	193.0711	193.0725	-1.41	8/0	9	4	2
	193.0773	NO HIT					
.0	193.0816	193.0805	+1.09	8/1	10	3	2
100.0	194.0811	194.0803	+ .74	8/0	10	4	2
5.4	195.0829	195.0836	+ .75	7/1	10	4	2

CAFFEINE $C_8H_{10}N_4O_2$ mw = 194

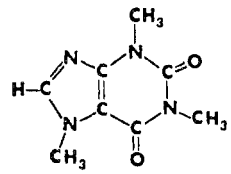

Figure 8. Partial listing of computer-generated elemental composition in vicinity of molecular ion for caffeine. NO HIT signifies that no composition is found within the specified error tolerance.

Time Requirements. One of the most dramatic advantages of the on-line system as compared to peak-matching or photoplate recording is the rapid availability of the reduced data. Output of results begins in 1–2 min after completion of the scan. Because of the slow teletype output, the amount of data determines the total time required, usually 10–20 min. We have recently installed a moderately priced ($10,000) high-speed printer. For a compound such as caffeine (Figure 8) this reduces the time required after completion of the scan *to approximately 2 min.*

Closed-Loop Feedback Control. The mass spectrometer–computer system described above can also be used to measure metastable ions by the Barber–Elliott–Major technique.[9,21] Lowering the energy of the electrostatic analyzer to m_2/m_1 of its normal value allows passage of ions of mass m_2 formed in the first field-free drift region by the reaction $m_1 \rightarrow m_2$. This also eliminates the normal energy ions, so that metastable ions can be detected over a dynamic range of 10^4–10^5.[21] By controlling the electrostatic analyzer potential, the computer can thus measure selected metastable transitions during the magnetic scan. Alternatively, all possible values of m_2/m_1 can be scanned repeatedly during a slow magnetic scan to record all of the meta-

(20) D. M. Desiderio and K. Biemann, Proceedings of the 12th Annual Conference on Mass Spectrometry, Montreal, June 1964, p 433.

(21) F. W. McLafferty, J. Okamoto, H. Tsuyama, Y. Nakajima, T. Noda, and H. W. Major, *Org. Mass Spectrom.*, 2, 751 (1969).

REL.AB.	MASS	CALC.	ERR.	C12/C13	H	N	O
6.5	78.0345	78.0343	+ .14	5/0	4	1	0
5.2	78.0456	78.0469	-1.24	6/0	6	0	0
100.0	79.0425	79.0421	+ .33	5/0	5	1	0
.6	79.0503	79.0502	+ .14	5/1	6	0	0
1.7	80.0466	80.0454	+1.14	4/1	5	1	0

Figure 9. Partial computer listing of elemental compositions for a mixture of benzene and pyridine. The mass separation at m/e 79 ($M/\Delta M$) is 1/10,000, scan rate 60 sec/decade. The relative abundance for m/e 80 is low because no correction has been made for the peak area below the preset threshold value.

stable ions.[22] This system saves many hours per spectrum over the extremely laborious manual method.

Other possible instrument parameters for which feedback control would be valuable are the filament current, bombarding electron energy, accelerating voltage, source and exit slit widths, ion source temperature, and magnet current. However, sophisticated hardware and a complex network of software are required for such a fully automated data-acquisition system.

Laboratory Data System

The on-line system described above provides speed, accuracy, resolution, and limited feedback functions at a relatively reasonable price; very similar data-handling systems could be of substantial utility for a wide variety of instruments and experimental techniques. However, a more powerful computer is required for many desirable on-line operations with mass spectrometer systems, such as the feedback controls and deconvolution of overlapping peaks described above. Much faster processing of data is advantageous when the mass spectrometer is directly coupled to a gas chromatograph, and direct identification of components by comparison against standard reference spectra stored in the computer can be envisaged. Computer interpretation of mass spectra[13-15] also has much higher core requirements. We have recently proposed[23] the direct determination of amino acid sequences in the mixtures of peptides resulting from the degradation of proteins and other polypeptides. Analysis of more complex mixtures appears feasible with repeated high-resolution scans and defocused metastable determinations during sample vaporization; however the necessary on-line reduction of the large volume of data produced would require a much more powerful computer system.

As noted earlier, justification of a large computer usually necessitates that it serve more experiments. Such a multiple-use system, especially if it utilizes "time sharing," generally requires much more complex hardware and software, so that justification is questionable for tasks that can be done by dedicated small computers. However, additional applications of potential similar to those enumerated above for mass

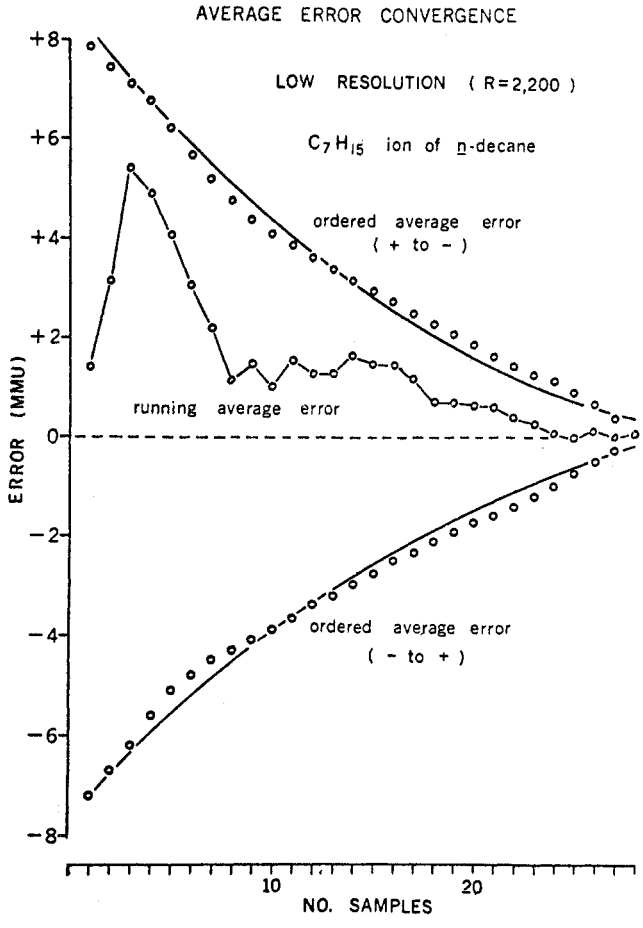

Figure 10. Improvement in mass-measuring accuracy resulting from multiscan averaging utilizing the on-line computer.

spectrometry can be cited for a wide variety of chemical research problems. We feel that the logical next step is to incorporate a number of such small dedicated computers handling a variety of such experimental devices into a larger on-line laboratory data system organized around a more flexible and powerful computer. The capability of the large computer for handling large volumes of data and more complex calculations (such as deconvolutions and Fourier transformations) would be available both on-line and off-line to any part of the system where necessary, but the small computers would handle primary data processing and experiment control. Worthy of special note is the potential of graphics-display devices such as the plotter and cathode ray tube (CRT) which can greatly facilitate man-in-the-loop type interaction of the investigator with the results of his experiment or of his theoretical calculations. For example, a larger computer can display molecular structures or potential energy surfaces on the CRT with rotation or modification of the image under direct control of the investigator for study of three-dimensional molecular interactions, interpretation of diffraction data, or location of potential minima.

(22) This computer system for metastables was developed by J. E. Coutant in this laboratory.

(23) G. E. Van Lear and F. W. McLafferty, *Ann. Rev. Biochem.*, 38, 289 (1969); F. W. McLafferty, R. Venkataraghavan, and P. Irving, *Biochem. Biophys. Res. Commun.*, in press.

We are indebted to D. H. Bessel, F. L. Hiltz, R. E. Hughes, R. A. Plane, D. A. Usher, and other members of the Cornell Chemistry Department staff who helped plan the laboratory data system that is described, and to the U. S. Public Health Service (National Institutes of Health Grant GM-16609) for generous financial support.