An Off-Line System for Data Acquisition and Analysis[1]

RAYMOND C. MASTER, NDI Chem Systems Division,
Nuclear Data, Inc., P.O. Box 451, Palatine, Illinois 60067

ABSTRACT

A generalized solution to the problem of laboratory data acquisition and data reduction will be discussed. The Nuclear Data 7700 System represents a new approach to laboratory data acquisition and reduction in that it is capable of interfacing with, providing complete data analysis and generating comprehensive output reports for many instruments in the analytical laboratory. Featuring both off-line data reduction and digital magnetic tape cartridge program storage, it was designed to offer an economical and minimum obsolescence solution to the data reduction problem. The system allows the full capability of a small general purpose computer to be employed with a variety of computer programs to perform data analysis for a variety of instruments. Besides software availability, providing complete data reduction and comprehensive output reports, for the more important analytical instruments such as Gas Chromatographs, IR and Atomic Absorption Spectrophotometers, Physical Testing Instruments, etc., the flexibility of this approach makes computerized data reduction practical for such instruments as Thermal Analyzers, Titrators, etc. The gas chromatography data handling aspects of the system are discussed in detail.

INTRODUCTION

In recent years, considerable interest has developed in the area of automating the data output of many types of analytical instruments. The various options range from the use of dedicated integrators and digitizers, through the small dedicated computers per instrument or minor multiplexed system, up to a major on-line laboratory computer installation. Other possibilities include the use of time share facilities or in-house batch processing of data. Properly selected, all data reduction schemes mentioned can serve a need and act to ease the burden of data reduction in the laboratory. Each laboratory must develop a data reduction approach most suitable to their needs, both present and future, weighing such factors as performance and economics. Generally speaking, a system capable of receiving raw data from a multiplicity of analytical instruments is still very expensive. The typical analytical research or service laboratory utilizes a large variety of analytical instruments but not a great number of any one type. The recognition of the fact that these laboratories would find both the dedicated computer path and the large on line computer system approach unjustifiable from a cost per instrument basis, led to the development, by Nuclear Data, of the 7700 System; a system capable of utilizing the powers of a small general purpose computer with a variety of computer programs for many types of analytical instruments.

A block diagram for the 7700 System is shown in Figure 1. The system consists of a central processing station plus sufficient data acquisition stations to provide the instrument interfacing coverage required in the laboratory. The primary data acquisition method, and an essential element in the 7700 System, is a Nuclear Data AR-04 Analog Instrumentation Magnetic Tape Recorder. This recorder was designed specifically for handling the low data rates characteristic of analytical chemical instruments, and includes the features and controls required for interfacing onto such instruments. To eliminate the problems of magnetic tape handling, the recorder utilizes a specially designed reel to reel cassette.

An auxiliary method of data acquisition for long experiments with low data rates (such as in Gel Permeation Chromatography) is via an ND Printing Digitizer. Besides providing the general capability of digitizing analog data and generating a punched paper tape record, this unit is also equipped with the circuitry to sense and record related events such as elution and injection marks for Gel Permeation Analysis.

The recorded experimental data is analyzed at the central processing section of the 7700 System. For magnetic tape data processing the cassette containing the experimental data is installed on the reproducer and the proper cartridge installed in the program tape reader. Facilities are also available for direct on-line inputting of data to the system. The data tape will be rewound and aligned at the beginning of the tape, and run phase begun all automatically under computer control. When the program is loaded, the operator is informed as to which program has been loaded and commanded to answer pertinent questions regarding the analysis. All operator interfacing with the system is done via the keyboard in a conversational mode. All system interfacing electronics are optimized (e.g., digitizing rate, etc.) automatically, for the type of analysis by the computer program.

All computer software has been designed in a consistent and logical manner to minimize both the amount of operator intervention and the amount of learning time required to operate the system. For each analytical instrument a software package was designed for the greatest flexibility possible allowing, for any given analysis, the maximum amount of program customizing by the operator in a conversation mode via the keyboard. However, for more routine analysis, the amount of operator interfacing with the keyboard is held to a minimum.

To illustrate the software power possible with a small general purpose computer in the proper hardware configuration, the program for gas chromatography (GC) analysis will be covered in detail followed by a brief description of the analysis of data from other analytical instruments, including IR spectrophotometers, Gel Permeation Chromatographs and physical testing instruments. More complete information on the system hardware details is given in Appendix I. More detailed information on the hardware and software will be made available upon request.

GC data is recorded on the AR-04 magnetic tape recorder. The recorder is connected to the electrometer output (flame ionization detectors) or detector output (thermal conductivity detectors) in order to use the largest signal available and allow continual use of the attenuator, if desired. It is possible to use the attenuator output signal (strip chart recorder signal) since the recorder allows inputs down to 0 to 1 mv, full scale, but this mode of recording is not recommended. The time of injection is noted on the tape by using the event marker circuit for which a front panel switch and remote terminals are provided.

---

[1] One of seven papers to be published from the Symposium "Computer Systems and Applications in the Oil and Fat Industry" presented at the American Oil Chemist' Society World Congress, Chicago, September-October 1970.

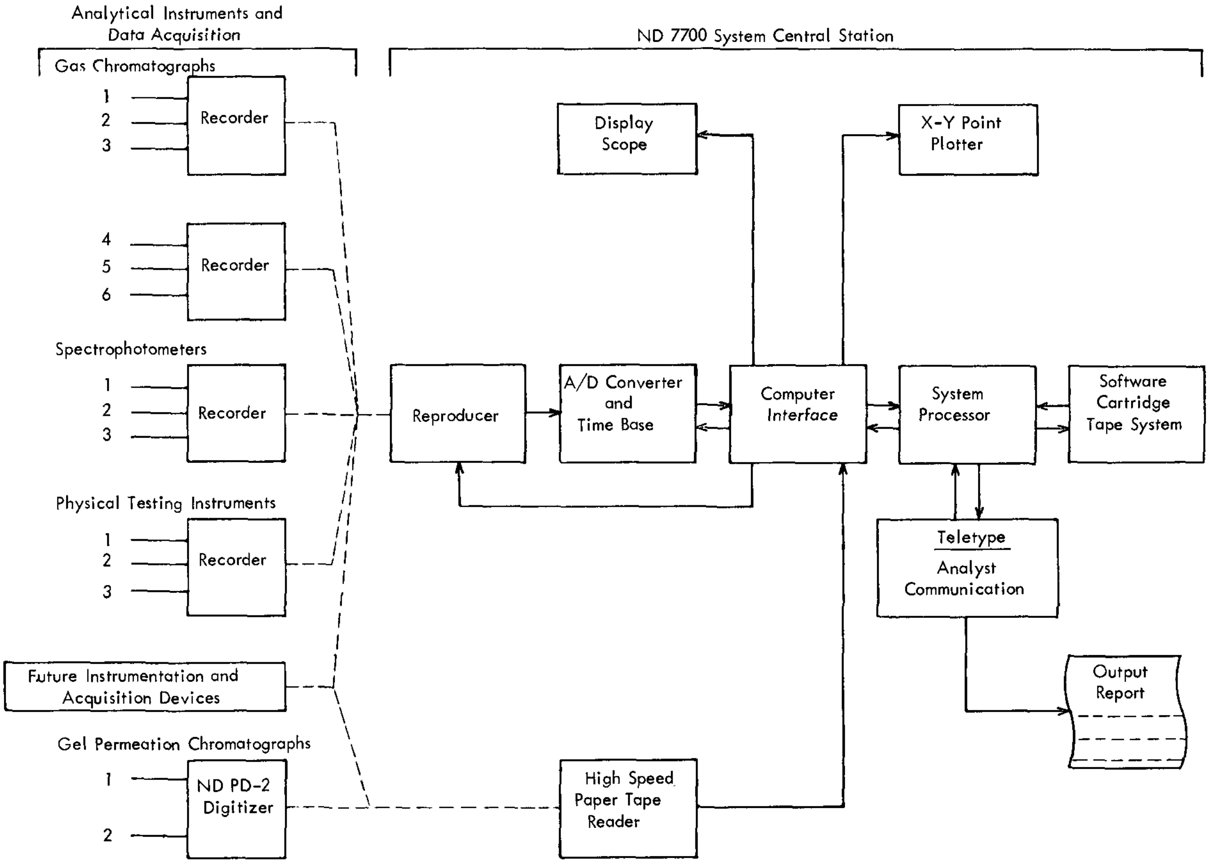

FIG. 1. Typical laboratory installation and ND 7700 system block diagram.

Because of the diversity of requirements in GC data analysis, the GC software package is one of the most formidable programs in the 7700 System software library. Besides providing the analyst with the most basic capabilities of determining the baseline corrected areas of all peaks, including overlapping peaks, the standard GC program also includes many other features usable in performing both routine and non-routine analysis. Included among these features are:

1. Storage of compound identification dictionaries containing compound names, peak weighting factors and relative retention time windows.
2. Application of a time window which can be employed to analyze only a region of interest within the chromatogram.
3. Selection of unknown peak weighting factors as zero or unity or, on a peak by peak basis during the analyst period, the application of particular values for each unknown peak.
4. The use of internal standards for normal quantization or trace analysis.
5. Selection of the overlapping peak area allocation method best suited to a particular analysis. One method can be selected for the entire chromatogram, or the analyst may choose the most appropriate method for each multiple peak independently.
6. Peak grouping, which is usable at any section of a chromatogram. This feature may be applied for such purposes as lumping all heavies in an analysis.

A complete initial dialog for the GC program is shown in Table I. The amount of keyboard conversation which the analyst or operator has with the initial dialogue is dependent upon the use of the special features for a particular analysis. In a "STANDARD PARAMETERS" analysis the operator identifies the dictionary list number and action (use, calibrate, enter, or print list), the approximate length of run in minutes and the time window of the reference peak. All further dialogue is aborted. If a dictionary is not usable for a particular analysis, the operator calls list zero and the output report will contain raw areas as well as per cent normalized areas, instead of per cent concentration values. To minimize the requirement of frequent updating of dictionaries, relative retention times are stored for each compound, with one peak having a relative retention time of unity. When using a list the operator identifies the approximate location of this reference peak. Upon complete of the initial dialogue, the data taking phase of the analysis begins. The operator may then enter a report heading of any form or length desired while this and other analysis phases are in progress. The Teletype output of each question in the initial dialogue can be terminated by the operator initiating the input response. The operator is notified of all invalid or inconsistent responses and is allowed to reenter that response, or in the case of an inconsistent response, allowed to reenter the earlier responses. The operator can also reenter all responses by a "NO" entry to the last dialogue question.

Since all recorded analytical data stays on tape until the data cassette is reused, it is possible to rerun the same data reduction to confirm the results or to rerun the data varying some of the parameters of analysis.

The features and uses of the GC program are further illustrated in the examples below, but first it is necessary to define the basic peak finding and data handling procedures employed.

BASIS PEAK DEFINITION AND DATA HANDLING FOR GC PROGRAM

Since the digitized chromatogram values must be sorted "on the fly" it was necessary to design a data taking routine capable of properly rejecting noise peaks and store only

TABLE I

Gas Chromatography Program Dialogue

```
GC Cartridge Loaded.
Please Enter List Number and Action.
N = Normal Run
C = Calibrate List
E = Enter List
L = Print List
IN
Length of Run in Minutes = 18
Reference Peak Window Start Time in Minutes = 7.5
Reference Peak Window Stop Time = 8.5
Standard Parameters?
No
Time Window Desired?
Yes
Window Start Time in Minutes = 3.1
Window Start Time = 16.6
Unknown Peak Weighting Factors.
0 = 0.00
1 = 1.00
? = Entry of True Factors via Teletype on a
Peak by Peak Basis during Analysis Period.
?
Select Peak Allocation Method.
B = Baseline Drop
C = Parabolic Curve Fit
N = No separation of Overlapping Peaks
M = Multiple Choice during Analysis via Teletype
S = Standard Computer Logic Choice Between B and C
M
Min. Slope as % Full Scale/Min. = 0.5%
Will Internal Standards be Used?
Yes
Peak Line Number = 3
Value of Int. Standard = 3.78G
What Recording Mode is Used?
S = Standard P-W-M
A = Autorange P-W-M
S
Are all responses O.K.? Yes
Enter Heading as Desired. Terminate with "EOT".
```

TABLE II

Gas Chromatography Area Run,
Normal Usage or Prior to Dictionary Creation

Peak	Retention time	Absolute	% Area
+ 1	+ 3.26	+0.1341286E+07	+ 6.45
+ 2	+ 4.90	+0.2536640E+07	+ 12.20
+ 3	+ 6.47	+0.3294950E+07	+ 15.84
+ 4	+ 8.09	+0.3438363E+07	+ 16.53
+ 5	+ 9.77	+0.2133080E+07	+ 10.26
+ 6	+ 11.29	+0.4033070E+07	+ 19.39
+ 7	+ 13.59	+0.4022130E+07	+ 19.34

valid describing values of the chromatogram. Besides taking advantage of the digital averaging capabilities of the A/D converter, the peak finding routine employs special filtering techniques and slope sensitivity criterion. A multiple pass baseline correction routine is employed, which insures that all valleys between peaks are on or above the baseline, and the stored values of the chromatogram are examined to determine the presence of overlapping peaks. The standard area allocation method for resolving multiple peaks consist of a combination curve fitting and baseline drop technique. Each multiple peak is tested with regard to a criterion designed to determine which method will yield area resolution with minimum error. Once the baseline corrected areas are obtained, a search for the chosen dictionary is performed to begin the compound identification sequence.

The data input rate from the A/D converter for GC analysis is one digital value every 1.5 msec (equivalent to 24 msec, real time). Each A/D read represents a digital averaging of four A/D conversions in the case of nonautoranged data and seven conversions for autoranged data. Each incoming value is compared to the running average of the preceeding 32 values. A peak start is defined as when 16 consecutive points are all greater than their respective preceeding average plus a set constant. If one envisions a linear peak start from a flat baseline region, the running average of 32 previous samples at any point in time will lage the current value by a constant; the greater the slope the greater the difference between the current point and the running average. Actually, in the linear case, the running average has the value which the input value had 16 samples earlier. By insisting that samples exceed their running average by a set constant, a slope criterion is imposed. This is the value entered from the keyboard during the initial dialogue phase. In the case of the wide dynamic range autoranged data the slope criterion also contains a per cent of value term to allow detection of very low slope peaks at baseline and yet avoid peak splitting due to minor irregularities at higher signal levels.

Once the peak start criterion is met, the starting time and amplitude parameters (first of the 16 points) are stored and the peak integration process starts (the area is actually also included from first of the 16 points) along with a sample interrogation for maximum amplitude (retention time value) and test for peak termination. The peak termination criterion is similar to the peak start criterion in that a peak termination is defined as when 16 consecutive points are all less than the running average of the preceding 32 points, less the slope constant. The describing parameters of the peak are stored and the program proceeds to detect the next peak. The standard program has a 68 peak limit. The peak detection criterion in itself imposes (for a symmetrical peak) a minimum peak width of approximately 1 sec, real time.

The baseline correction program insures that all valleys between peaks are on or above the true baseline. Starting with a line drawn between the first peak amplitude and last peak termination amplitude, a search is made for the start or termination amplitude most below that line. Holding the end points fixed, the lowest point is raised to the baseline; elevating all other stored amplitudes by a proportional amount and adding the trapezoidal area increments to the stored peak areas. A search is then made for the next lowest start or termination amplitude and the correction process proceeds until no values are below the baseline.

The allocation of areas of overlapping peaks including an automatic regrouping criterion, as well as other features and powers of the GC program are covered in the sections below.

ANALYSIS MODES AVAILABLE IN THE GC PROGRAM

In instances where it is not practical or feasible to enter and work with a dictionary, the operator calls for list zero and an output report is produced giving peak number, retention time, absolute area and per cent normalized area as is shown in Table II.

This mode of operation is also usable prior to a dictionary entry in that from this output it is possible to calculate the relative retention times and the response factors for each compound.

A dictionary (or list) entry is initiated by an "E" response, following the selected list number, to the first dialogue question. The operator then may begin the entry under a computer generated heading. This is illustrated in Table III. Each dictionary can contain up to 34 compounds with compound names up to 30 characters in length. Each GC program cartridge can contain seven dictionaries with additional requirements met by other GC program cartridges.

For each type of analysis for which a dictionary is generated, a retention time reference peak must be selected

TABLE III

Gas Chromatography Dictionary Entry and Listing

```
GC Cartridge Loaded.
Please Enter List Number and Action.
4E
Enter list as shown by heading.

                                                         Response
                  Relative                                 Factor
     Compound  Retention time   % Window                   Group

01      A          .401           5.0                    262.65a
02      B          .6             5.1                      1.22
03      Xb
        C          .8             5                         .87
04      D          1              5.6                       .911
05      E          1.21           7                        1.15
06      F          1.391          8.11                      .731+
07      G          1.68           8                         .720
08
Please Enter List Number and Action.
4L

01      A       +  .4010       + 5.0000                 + 2.6500
02      B       +  .6000       + 5.1000                 + 1.2200
03      C       +  .8000       + 5.0000                 +  .8700
04      D       + 1.0000       + 5.6000                 +  .9110
05      E       + 1.2100       + 7.0000                 + 1.1500
06      F       + 1.3910       + 8.1100                 +  .7310+
07      G       +  .0000       +  .0000                 +  .0000
08

Please Enter List Number and Action.
4N
Length of Run in Minutes = 18
```

[a] Rubout before ALT MODE termination allows value to be retyped.

[b] Rubout before ALT MODE termination generates CRLF and four spaces for name retyping.

TABLE IV

Gas Chromatography % Concentration Output Report

```
7/14/69 CJL
GC#3, Col #1, 10ML/MIN

                                                           % or
Peak          Compound                 Retention time     Weight

+  1      Acetone                        +  1.65         +  5.49
+  2      I-Butanol                      +  2.48         + 11.34
+  3      Ethyl acrylate                 +  3.62         +  8.44
+  4      Methyl isobutyl ketone         +  4.51         + 10.56
+  5      Unknown                        +  5.02         +  0.31
+  6      Toluene                        +  5.72         + 10.14
+  7      Octane                         +  6.71         +  9.87
+  8      Unknown                        +  7.69         +  0.12
+  9      Cellosolve acetate             + 11.04         + 22.70
+ 10      Cyclohexanone                  + 14.78         +  5.89
+ 11      Methyl carbitol                + 18.14         + 10.02
+ 12      Diisobutyl ketone              + 19.58         +  3.81
+ 13      Unknown                        + 21.70         +  1.27
```

to which the other peak retention times are normalized. The relative retention time concept is employed in dictionary storage to avoid frequent updating requirements due to column temperature changes, etc. In subsequent employment of the dictionary, the operator identifies the real time window in which the reference peak will be found for that particular run. With this in mind, the retention time reference peak should be a peak which is readily identified for this type of analysis.

In addition to the relative retention time value, a per cent of value window is entered for each peak in the dictionary as well as the calculated response (weighting) factor for each peak.

A special feature of this dictionary storage concept is that of the Grouping Factor in the last dictionary column. Whereas a dictionary line is normally terminated with the keyboard space bar it may be terminated with a "+". This termination says that if ever an unknown peak(s) is found between the peak so terminated and the next identified peak, it automatically be grouped with the "+" terminated peak. This is especially useful in chromatograms where, for instance, there would be a number of unknown peaks at the end of the run. Employing this feature they could all be grouped and reported together (e.g., last list entry HEAVIES with "+" termination).

During the dictionary entry procedure any entry may be cancelled and reentered before termination. At the termination of a list entry the computer dialog is reinitiated allowing the operator to now proceed in the analysis and use this dictionary or to enter another dictionary or to confirm the proper storage of the recently entered dictionary, as is also illustrated in Table III. Dictionary content examination is possible at the beginning of each run. If a particular dictionary is unused, the operator will be so notified.

During usage of a dictionary, the exact time location of a peak found within the keyboard entered reference peak window is used to convert the chromatogram times from absolute to relative wherein the dictionary can be overlayed on the time normalized chromatogram for a peak identification search. Since it is possible to construct a dictionary with overlapping time windows for neighboring peaks and therefore face situations where a given peak is found to fall in two dictionary time windows, a multiple pass identification search is required. In the first pass, all peaks whose relative retention times lie in only one dictionary time window are identified. The respective dictionary time windows are eliminated for the second pass, thus eliminating one window from most, if not all, overlapping window pairs and enabling positive identification of additional peaks on the second pass. A typical output report with dictionary usage is shown in Table IV, which gives peak number, compound name, retention time and % concentration based on the response factors stored for the individual compounds. The per cent concentration values given in the output report are calculated from the areas measured and the response (weighting) factors stored in the dictionary for each identified peak.

The system software operation with an internal standard employed is similar to that above except that during the initial dialogue it is required to enter the dictionary peak number to be used as the internal standard as well as the value of the internal standard. The output report values will then be absolute values in whatever units the internal standard value was entered. This is illustrated in Table V.

If a particular peak in the list will be used as an internal standard for quantization, its response factor is often stored as unity with the other response factors normalized to that of the internal standard's, but this normalization is not necessary, allowing the response factors to be stored in some absolute or more meaningful units.

The GC software package provides the operator or analyst considerable power with regard to quantitating unknown peaks. In the "standard parameters" mode of operation all unidentified or unknown peaks in the chromatogram are assigned a response factor of unity (1.0). However, during the preanalysis or initial dialogue phase, the operator may elect one of three methods of quantitating unknown peaks, whether operating with or without internal standards. These are the following: (a) all unknown peaks to be assigned a response factor of zero (0.00); (b) all unknown peaks to be assigned a response factor of unity (1.00); (c) after the initial data reduction phase and before the output report is generated, notification is to be given as to the detection of each unknown peak in terms of its retention time position and the operator is allowed to enter

TABLE V

Gas Chromatography Analysis Illustrating Peak by Peak Entry of Unknown Peak Response Factors and Internal Standard Usage

```
Will Internal Standards be used?
Yes
Peak Line Number = 3
Value of Int. Standard = 20.7
Enter Heading as desired. Terminate with "EOT".
Reference Pk Limits Incorrect. Re-Enter.
10.0M
11.0M
7/21/69 GC#3 JSG

Unknown Peak at +  1.646  0
Unknown Peak at +  2.611  0
Unknown Peak at +  9.856  0

                                                % or
Peak       Compound      Retention time        Weight

+ 1        Unknown         +  1.646          +   .000
+ 2        Unknown         +  2.611          +   .000
+ 3        C 16            +  4.868          + 21.259
+ 4        C 18            +  9.146          + 20.813
+ 5        Unknown         +  9.856          +   .000
+ 6        C 18:1          + 10.541          + 20.700
+ 7        C 18:2          + 13.424          + 21.704
+ 8        C 18:3          + 18.205          + 20.951

Total + 105.42
```

TABLE VI

Gas Chromatography Output Illustrating Multiple Choice of Peak Allocation During Analysis Period

```
Double Peaks at +  2.58 and +  2.87.
S
Double Peaks at +  4.31 and +  4.90.
S
Double Peaks at +  4.90 and +  5.63.
S
Double Peaks at +  5.63 and +  6.18.
S
Double Peaks at + 12.10 and + 12.48.
S
Double Peaks at + 15.65 and + 15.80.
N
Double Peaks at + 15.65 and + 16.34.
N
Double Peaks at + 15.65 and + 16.88.
N
Double Peaks at + 19.25 and + 20.10.
S

                                                              % or
Peak          Compound                   Retention time      Weight

+  1   Hydrogen                            +  2.167       +   .796
+  2   Unknown                             +  2.342       +   .000
+  3   Propane                             +  2.584       +  6.129
+  4   Propylene                           +  2.872       + 10.731
+  5   Isobutane                           +  3.480       +  9.512
+  6   N-Butane                            +  4.310       +  2.335
+  7   1-Butene and isobutylene            +  4.903       +  4.690
+  8   trans-2-Butene                      +  5.633       +  2.295
+  9   cis-2-Butene                        +  6.183       +  1.609
+ 10   Isopentane                          +  6.962       +  4.199
+ 11   N-pentane                           +  8.284       +   .223
+ 12   Unknown                             + 10.546       +   .000
+ 13   Carbon dioxide                      + 11.746       +   .396
+ 14   Ethylene                            + 12.108       +  2.556
+ 15   Ethane                              + 12.486       +  3.576
+ 16   Reference time peak                 + 13.266       +   .000
+ 17   Unknown                             + 13.800       +   .000
+ 18   Unknown                             + 14.499       +   .000
+ 19   Total C5 olefins and C6 plus        + 15.657       +  2.668
+ 20   Oxygen and argon                    + 18.742       +  1.036
+ 21   Nitrogen                            + 19.255       + 39.393
+ 22   Methane                             + 20.100       +  7.855

Total +99.999
Enter G When Next Run Ready
```

a response factor of any value for each unknown peak, independently.

This third method is particularly valuable in that among other uses it can be employed as a tool to reject noise peaks which were not rejected by prior program sections in that they satisfied the chromatogram peak model employed. This mode of operation is also illustrated in Table V. Also shown in this output report is one example of the built in capability for detection of invalid entries. If a peak is not found in the reference peak window entered during the initial dialog phase, the operator will be so notified after the initial data reduction phase.

The standard area allocation method for resolving or appropriating the areas of multiple or overlapping peaks consist of a combination curve-fitting (parabolic skim) and perpendicular drop, to corrected baseline, technique. Each multiple peak is tested with regard to a criterion design to determine which method will yield area resolution with minimum error. However, if requested during intial dialogue, the operator can select from the five methods of overlapping peak area allocation methods below: 1. Perpendicular drop to baseline for all overlapping peaks (B). 2. Parabolic curve fit for all overlapping peaks (C). 3. No separation of overlapping peaks (N). 4. The standard computer logic choice between (B) and (C), (S). 5. Multiple choice of individual methods for each overlapping peak (M).

The operation of multiple choice method is illustrated in Table VI. After the initial analysis and baseline correction phases and before the output report is generated, the operator is notified of the retention time locations of all overlapping peaks and is allowed to use the method of his choosing for that pair by entering B, C, N or S. When calling for no separation for a particular peak pair, all area is thereafter associated with the lowest retention time peak. These methods of area allocation are not limited to a doublet but are applicable to any number of overlapping peaks in a group, giving the operator complete power of individual selection of area allocation methods for each multiple peak in the chromatogram. Two peaks are considered to be overlapping if the time difference between the termination of the first peak and the commencement of the second peak is less than 385 msec, real time. Peak starts and termination are determined by the slope sensitivity employed for the analysis. This criterion applies at all signal levels. The second possibility is, if the level of peak termination is greater than 1.125 times the level of the peak start. If a peak terminates at a high level, a new peak start is forced, capturing all remaining chromatogram area and forcing compliance in the first criterion.

This second criterion is very important since certain irregularities or noise on a peak can cause a peak termination without a subsequent peak start. However, assuming such noise did also cause a new peak start, it still may not meet the first criterion, and without the second criterion it would not be flagged as a multiple peak. If three or more peaks are involved in the overlapping group, each peak terminator is referenced back to the last good baseline value (starting level of first peak in the group) applying higher powers of the 1.125 multiple for each additional peak in the group which allows for the possibility of the existence of a positively increasing baseline. The program also contains an automatic peak regrouping criterion which is employed where a multiple peak is detected and yet fails to meet the chromatographic peak model utilized in the software (e.g., second peak's maximum intensity time position too near the first peak's termination time or second peak's maximum intensity not sufficiently greater than first peak's termination amplitude but is greatly separated in time).

IR SPECTROPHOTOMETRY, GEL PERMEATION CHROMATOGRAPHY AND PHYSICAL TESTING DATA ANALYSIS

The system software is designed to present a consistent format of conversational dialogue to the operator for all types of analysis. The system software library is in a state of continual development according to a plan to write a comprehensive computer program for each analytical instrument for which there is compatability with the basic systems concept. Besides software additions for other important analytical instruments, the flexibility of this approach makes computerized data reduction practical for such instruments as thermal analyzers, titrators, etc. The salient feature of the other major software packages written to date are described briefly below.

The IR software package is designed to perform a smooth on the data, locate all peaks and valleys and provides a printout of all peaks below a preset rejection level in per cent transmittance, their respective wavenumber values and average integrated background on both sides of each listed peak. For each analysis the operator identifies the scan function number to be employed, rejection level and whether the input is in transmittance or absorbance. The scan relationships are entered as up to four section linear functions having different linearities in terms of wavelength or wavenumber versus time or encoder pulses and stored on the program cartridge. If the operator is unsure of the contents of a scan function, the function can be listed prior to use, or a new function can be entered. A rejection level may be set to any value between zero and 100% transmittance. In the standard parameters mode, for qualitative analysis, a rejection level of 60% is employed to provide a listing and punched paper tape output of the peaks suitable for inputting to a larger machine equipped with a search file.

Gel permeation chromatography (GPC) provides the chemist a reliable method of obtaining molecular weight distribution. This analytical instrument also represents an excellent example of the advantages of a computer system in the laboratory. The data output from the chromatograph consists of a recording of polymer concentration versus time; marked at fixed increments of elution volume. Data reduction generally consists of the calculation of molecular weight or molecular size averages and the replotting of normalized data in various forms, such as weight fraction versus molecular size, which allows better direct comparison of samples. The manual computational labor involved generally forces the laboratory to accept much less than would otherwise be desired in terms of data reduction. The availability of a computer system removes this restriction.

Data from the GPC is acquired via a Nuclear Data Printing Digitizer, a single channel or two input multiplexed unit. The printing digitizer is equipped with a standard Teletype to generate both a printed listing of the original data and a punched paper data tape record. The digitizing rate is set compatible with the column flow rate to provide at least 10 values per elution dump. For a 1 ml/min flow rate the unit is set to digitize at 15 sec intervals providing a nominal 20 values per elution dump. Each 5 ml fraction dump of the syphon is also noted on the data tape record as an elution dump occurring at a specific percentage of the time between data samples. The digitizer is also equipped to record the occurrence of the injections when running in the automatic injection mode.

The initial dialogue allows the operator to set the low and high elution count limits for baseline correction and separate low and high elution count limits over which the calculations are to be made. The latter values are determined by the usable portion of the calibration curve whereas the baseline correction is determined by observation of the strip chart recording. The program will handle both single runs and multiple runs, performing sequential analysis of the data from an overnite run. For each analysis the operator identifies the calibration curve (molecular weight or molecular size versus elution volume) to be employed. The calibration curve parameters are stored on the program tape and identified by number from the keyboard. The calibration curves are generated applying a least square fit of a fourth order polynominal to the values obtained by running standard polymers on the various columns. Once the coefficients of the calibration curves are determined, they are stored on the program tape.

The data reduction phase is initiated by the original data being read into a buffer region in the computer and displayed, followed by a display of the baseline corrected data. The next step in the program consists of a transformation of the data from the time scale axis to an elution volume axis such that any variation in flow rate is corrected for, yielding a baseline corrected concentration curve linear with elution volume with exactly 10 data points, in equal increments of elution volume, per elution dump (e.g., data point every 0.5 ml). This data is used to calculate the values presented in the output report including molecular size and weight averages, polydispersity figures, standard deviations and a skewness figure of the distribution. In addition the standard program provides output plots on an X-Y plotter of both the normalized weight fraction versus log of molecular size (or weight) and the integral or cumulative function.

The physical testing software package is designed to perform the calculations for the analysis of physical properties of materials and to generate a complete formatted output report. Depending on the type of analysis, calculated values include tensile strength, elongation, modules, rupture energy, or lap shear strength, drum peel strength, or coefficient of friction.

Data is collected via the ND-AR04 Magnetic Tape Recorder which can be connected to three physical testing instruments in the lab. In this application the recorder is equipped with an automatic calibration circuit which is driven by the recorder event marker circuit to record a precision data pulse prior to the recording of a set of sample records. The calibration pulse is then used during data analysis to insure the integrity of the values reproduced on the absolute basis.

The general operating mode is to prepare all samples (up to eight replicates each of 15 samples) to be run for a single report in advance and to load the samples into the machine with the recorder in a continuous run mode. To avoid recording the noise signals generated by the load cell during sample changing, the machines are equipped with circuitry to produce the zero output level in the pen off mode.

In a typical application where a number of replicates of each sample is tested and the number of different samples tested for a given report may vary, the operator updates the calibration between sample sets by use of the event marker circuit.

Data analysis is initiated by loading the data tape cassette on the central system reproducer and the physical testing program cartridge in the magnetic tape program reader and keying "G" on the system keyboard. While the data tape is in the automatic rewind mode the operator is requested to identify the type of analysis and to enter the full scale load value and, depending upon the type of analysis, a second parameter such as elongation rate, sled weight, etc. To expedite a complete output report without further intervention, the operator then types under a computer-generated heading the sample identifications, number of replicates and replicate numbers to ignore for each sample. Upon completion of the initial dialogue and output report heading, the data analysis phase is begun under computer control in which all data related to the complete report is reproduced continuously at a 16 to 1 time compression.

A report heading is generated allowing the operator to enter salient report details such as date, report number, environmental conditions, etc., under which the output report list for each sample, the sample identification followed by a formatted tabulation of all values calculated for each replicate of the sample as well as statistical values for each parameter.

APPENDIX I-ADDITIONAL HARDWARE DETAILS

An essential element in the design of the ND 7700 system was the development of an instrumentation grade magnetic tape recorder for acquisition of the analog electrical signals available from the analytical instruments. The successfully achieved target was to design a tape recording system which could reproduce a recorded spectrum with a S/N ratio of 60 db (1000:1) over the signal bandwidth of DC to 10 Hz. The ND AR-04 recorder is a basic two track direct drive capstan unit and operates at a recording speed of 3.75 ips. The reproducer at the central operates at a playback speed of either 60 ips (primary playback speed for most experimental data) or 15 ips. Tests have shown, that without the use of shaft encoders on the analytical instrument, the tape system will reproduce the time scale to within better than 1 part in 500. The basic recording method, for signals within the 0 to 10 Hz bandwidth, is pulse width modulation. For recording signals in the 0 to 300 Hz bandwidth frequency modulation is utilized.

To provide wider dynamic range, as is required for many GC applications, autoranging is employed to extent the theoretical dynamic range to 128,000:1. Of course, as with any device, the actual dynamic range achievable depends upon signal magnitude availability in terms of the ratio of the largest signal to the noise threshold of the instrument of receiving device. The autoranging capability is achieved by a technique of multiple track recording wherein at the carrier frequency rate, the gain for optimum recording level is insured. The recorder information is reassembled in the A/D converter, at the central during data processing, to form a double precision number for transmission to the computer.

This series of recorders was designed specifically for interfacing onto the instruments in the analytical laboratory in that, besides the signal recording specification, the other features were also developed for this application. To eliminate the problems associated with magnetic tape handling, a cassette was developed to house the two 8 in. reels. The standard tape cassette allows up to 1 1/2 hr of recording time and is equipped with both end-of-tape sensors and leader tape on both ends of the tape, which because of the photoelectric sensing system on both the recorder and reproducer transports, insures the intactness of the cassette. For applications requiring longer recording times a cassette is available to provide up to 3 hr of recording time.

To allow sufficient spacing between experiments and to provide adequate time for the 60 ips reproducer to get within flutter specifications from a stopped condition, the recorder incorporates a leader time-out circuit which notifies the operator when a sample can be injected or scan initiated. Also provided is an event marker control which allows the operator to mark the beginning of experiment on the magnetic tape record. The AR-04 series recorders also include an instrument selection switch which allows the recorder to be centrally located and switched between three instruments, strip chart recorder terminals which provide the proper signal-level to a strip chart recorder allowing the maximum available signal to be used within the magnetic tape recorder and remote event marker terminals. A gain switch, with binary steps, provides for input signal ranges of from 0 to 1 mv to 0 to 1 v. Each recorder is equipped with erase facilities such that when a tape cassette is to be reused the old experimental data is erased as the new data is recorded.

The reproducer transport has the mechanical features of the recorder described above and all the necessary read electronics to demodulate the recorded signals and reconstruct the data spectrum. During the data reduction the reproducer is operated under computer control, wherein the data cassette is rewound and aligned at the beginning of tape and commanded to begin the run phase automatically. At this phase of data reduction the entire computer system is dedicated to this one analysis as though the data was originating from the analytical instrument except that the data is reproduced at a rate 16 times as fast as it was originally acquired. The data from the reproducer is fed to the 7700 system programmable analog to digital converter, designed specifically for this system. The A/D converter was designed to eliminate operator intervention; all conditions are optimized for each type of analysis and are constants of the computer program for that analysis. The following are standard features of the A/D converter.

1. A programmable crystal-controlled time base which provides sampling rates from 10,000 samples per second to one sample every 100 sec.

2. A basic 4096 channel (12 bit) A/D converter having a D.C. input sensitivity of 1.4 mv per channel and conversion time of approximately 200 μsec for a maximum amplitude input signal.

3. For digitizing the reproducer outputs from a data tape recorded in the autoranged mode the A/D converter is arranged, by computer command, to accept both the gain and data signals and provide a dynamic range of 18 bits with the same accuracy on the most significant 10 bits for values within that range.

4. Digital averaging is provided within the A/D converter. This feature is used in analysis where the rate at which the computer can accept data is considerably slower than the conversion time of the A/D converter. This minimizes the idle time for the A/D and provides additional noise rejection in an amount proportional to the square root of the number of values averaged (added).

With the averaging capability the range of values for both nonautoranged data (16 bits maximum) and autoranged data (22 bits maximum) is greater than the basic 12 bit word size of the computer and is read from the A/D converter 24 bit buffer in two 12 bit bytes.

Before the data read-in phase is initiated the operator installs the appropriate program cassette into the program tape transport and commands the beginning of the analysis by keying "G" on the system keyboard. This will start the program tape transport and load the program, or the first program phase, into the computer. The program tape unit consists of a magnetic tape transport and a computer interface section. The program tape transport is a read/write instrument used to store computer programs or any other data from the computer memory in binary form. All functions are performed under computer control. The program tape storage system greatly expands the program capabilities of the system processor which is a 4096 word, 12 bit machine. A standard ASR-33 Teletype is provided as the primary I/O device for the system. Besides providing the input/output decoding for peripheral instructions and data transfer, the computer interfacing also includes the necessary digital to analog converter capability for oscilloscope and X-Y Plotter displays.

[Received October 1, 1970]