総 説

Gas Chromatographs in On-Line Computer

*Chi-Whe Koo

During the past 20 years, the computer revolution has dramatically changed our world, and it promises to bring about even greater changes in the years ahead. The general purpose digital computers being built today are much faster, smaller, more reliable, and can be produced at lower cost than earlier computers. The first big electronic computers were usually employed as super calculators to solve complex mathematical problems that had been impossible to attack before by human hand calculations. In recent years, computer programmers have begun using computers for nonnumerical operations, such as control systems, communications, and data handling and processing. In these operations, the computer system processes vast quantities of data at high speed.

Before discuss in detail the scientific applications of computer, define two basic concepts which are off-line and on-line. A system which is under the control of and connected to the central process unit of the computer is said to be on-line. Another words, the computer used by a scientist to directly collect data, sampling, and analyze data while an experiment is in process. In contrast, the computer used to analyze data that has previously been recorded is said to be used off-line. By means of the digital computers automation of the scientific laboratory can take several methods. Most of laboratory data are calculated by the computer and this is one kind of off-line system. Data can be accumulated by any of some other methods and at some later time, data reduction and computations can be performed by a central processing computer. Such a system by itself has very limit the versatility and dimension of laboratory automation. However it is very valuable when applied to lengthy mathematical calculations required for many theoretical and experimental problems. In this off-line case, does not have any relationship between laboratory instrument and computer, the computer is used for data analysis only.

The experimental scientist needs, in additions to computations, the control capabilities of the computer and the vers atilities it offers in the data acquisition functions. On-line system case, the computer is directly connected with laboratory instrument. And it can directly get experimental informations from the computers. These desirable features can best be obtained by a time-shared system. This time-shared systems can most easilly be applied where their a number of several routine control and data acquisition functions to be performed. These time-shared system often operate in a sequential mode. That is, the computer performs the data acquisition functions by sequencing from one unit to the next in an orderly predetermined program.

If the instrument is on-line system, that means instrument and computer are physically connected. Usually most of laboratory instrument's output signal is an analog type, another words voltage or resistance signals. And this analog signals are used to the strip

---

*IBM Korea Systems Engineer

chart recorder, oscilloscope, or some other method which can get informations. Unfortunately, computer cannot catch the analog signal directly, and inhere, it is needed some conversion from analog signal to integer signal which computer can understandable numerical value. The analog-to-digital converter can do this kind of job. This analog-to-digital converter works to converting from anolog signal to digitalize and send to the computer with this digitalized integer values. These digitalized informations can be stored in the computer's storage unit and also these informations can be analyzed by the programs.

Gas Chromatographs represents a significant financial investment by research, quality control, testing, and medical laboratories. Efficient operation of chromatographs to ensure accuracy and reproducible ability of results is a key objective of any laboratory. Another objective is improved test turn around time, providing maximum utilization of the instrument and resulting in increased productivity and profile. To accomplish these objectives, many laboratories are automating their chromatographs by coupling them with digital computers.

During the typical gas chromatograph run, the scientist set up the instrument and then injects a sample to altered start a run. The instrument controls the oven temperature and a detection procedures of an output signal. The output signal is recorded and analyzed, and a report on the sample is prepared. It can accurately measure chromatograph singal output throughout its wide dynamic range and can also control instrument functions amenable to external control, such as temperature programming, column switching, or back flushing. Usually manual processing of the gas chromatograph data is difficult and time consuming work. Digitizing, data smoothing, integrating areas, applying response factors, locating and identification of the peaks, and allocating areas in overlapped to the user's commands. The acquired data are stored until released, allowing operator to repeat or perform additional data processing. The following results are taken automatically by the computer system;

a) Base line correction
b) Peak identification
c) Deconvolution of fused peaks
d) Assign calibration and response factors
e) Area normalization composition
f) Internal standard composition
g) External standard composition
h) Calibration and response factor calculation

Digital filtering can be most readily adapted to the unique requirements of gas chromatograph. Four types of digital filtering are provided, bunching with on-line updating, moving average, convoluted least squares smoothing differentiation, and digital rejection based on peak width (figure 1.). The result is a system which is insensitive to noise and which does not distort the gas chromatograph peaks. This allows relatively noisy chromatograph to be attached directly to the system, and also assures that noise picked up by signal transmission does not degrade performance. peaks to individual components are tasks that must be done well or the potential of the chromatograph is wasted. Computer processing of chromatograph data can frequently reduce a half day job into a half hour job and add an extra accuracy.

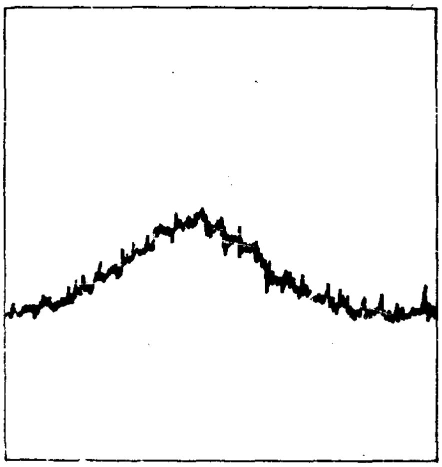

Fig. 1

Usually data acquisition and control system can handle inputs of up to 40 to 60 gas chromatographs independently and simultaneously. Time sharing of the computer allows each chromatographer to feel that he is the sole user of the system. The following tasks are performed automatically by the system;

a) Responds instantly to user's commands
b) Acquires data from the chromatographs in use
c) Smooths the data to reduce noise
d) Detects peaks
e) Integrates peaks
f) Measures retention time
g) Up to dates data acquisition parameters
h) Controls the gas chromatograph and accessory components

The acquired chromatograph data are automatically processed by the computer system at analysis end according.

Especially the moving average is a tenchnique of repetitively performing the some averaging procedure on a succession of data. Each repetition of the simple or weighted averaging procedure uses the same number of data, $2m+1$. For example the jth average will involve the data points Y_{j-m} through Y_{j+m} whereas the $(j+1)$th average uses Y_{j-m+1} through Y_{j+m+1}. the mathematical description of this process is

$$Y_j^* = \frac{\sum_{-m}^{m} C_i \; Y_{j+i}}{\sum_{-m}^{m} C_i}$$

where Y_j^* is the moving averaged data, C_i is a weighting factor associated with the ith term, Y is the referred data, j is the index of the control data point in the original data table, and i is the index with respect to j.

With this method, the function of the computer is to act as a digital filter to smooth noise fluctuations, hopefully without introducing distortion. Best results are obtained when the data are digitalized at high densities with the data points very close together, and the number of data points used is enough to include no more than one inflection.

First and second derivative logic of original gas chromatograph signal is used to detect and distinguish peaks, shoulders, and humps (figure 2). Special tests are provided to distinguish baseline shifts from chromatograph peaks. The sensitivity to peak detection can be adjusted automatically during the analysis to match the changing conditions. Thus, more slope sensitivity can be provided to detect the small, broad peaks eluting at the end of isothermal analysis.

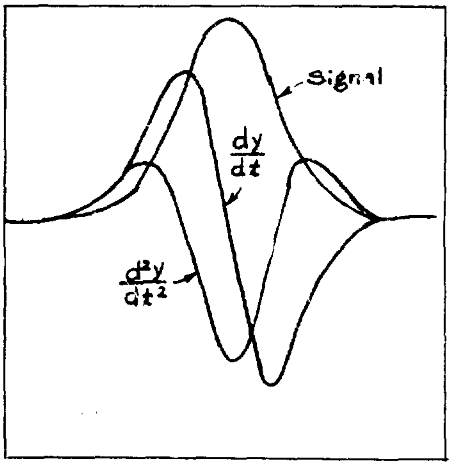

Fig. 2

Total peak area recovery is essential for high performance. Systems which give reproducible, but not total, area recovery are only satisfactory under repetitive applications. Total area recovery is dependent on sensitive peak detection and accurate baseline placement. Most systems use points of peak detection to start integration and define baseline. This decrease area recovery, especially when trying to recover small peaks in the presence of noise, since a finite slope is required for peak detection. This is solved by offsetting the points of start from the points of actual peak detection (figure 3). The

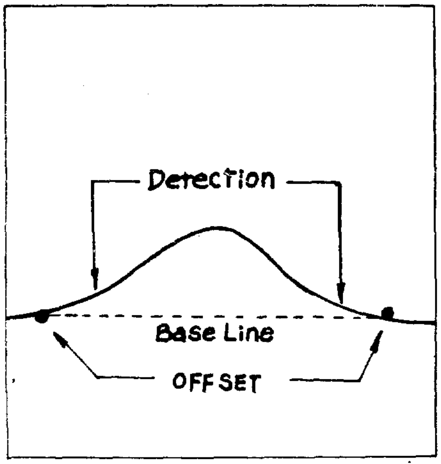

Fig. 3

offset is variable and can be changed automatically during the analysis.

Base line correction is the most important factor influencing area recovery. Baseline is reconized by slope and confirmed by the delay test (figure 4).

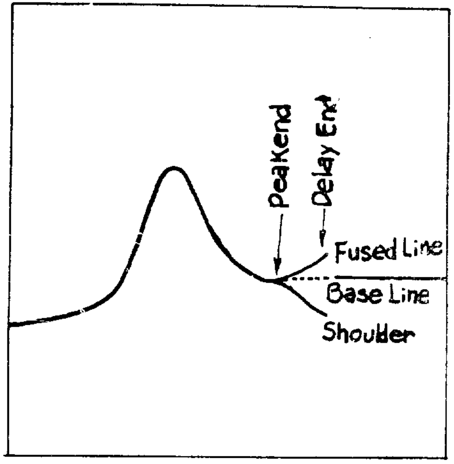

Fig. 4

The delay time is variable and can be programmed to change automatically during the analysis. Thus, a range of conditions can be selected, from having every vally point being interpreted as baseline, zero delay, to having a single base line level held under a group of fused peaks, long delay (figure 5). Both positive and negative baseline drift can be corrected. Groups of fused peaks are baseline corrected by drawing a line between the extreme base line points (figure 6). The principle of most negative baseline is used in cases where a vally lies below this line.

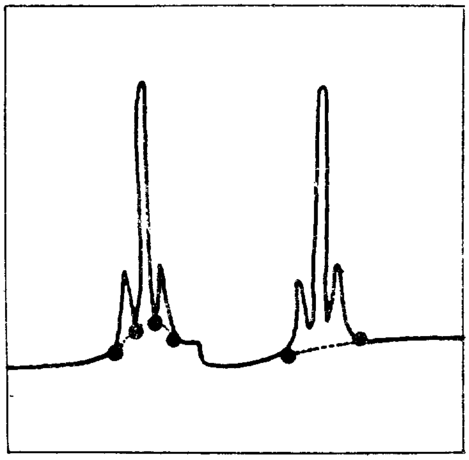

Fig. 5

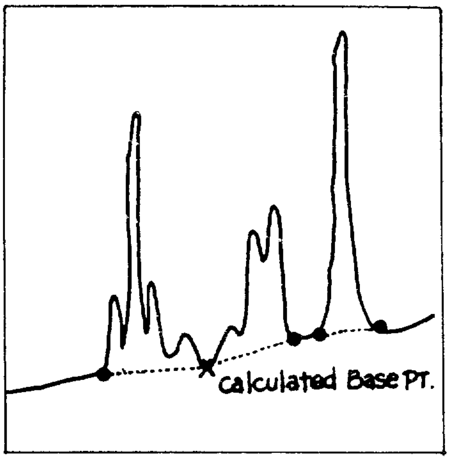

Fig. 6

Base line may also be fixed at any point during the analyses based on time and slope. This feature is useful to ensure that the analysis ends on baseline.

Usually three different methods of resolving fused peaks are used. These are the perpendicular, tangent, and skewed Gaussian methods. The perpendicular method simply drops a vertical line from the vally point to the baseline. This simplyfied technique is sufficiently accurate in many applications. The tangent mode is intended for cases where small peaks elute on the tail of larger peaks, greater than ten times difference. This is a common accurence in trace analysis or when peaks elute on a solvent tail (figure 7). The area below the tangent peak is assigned to the large peak. Performance data have shown this to be only reliable technique for resolving trace

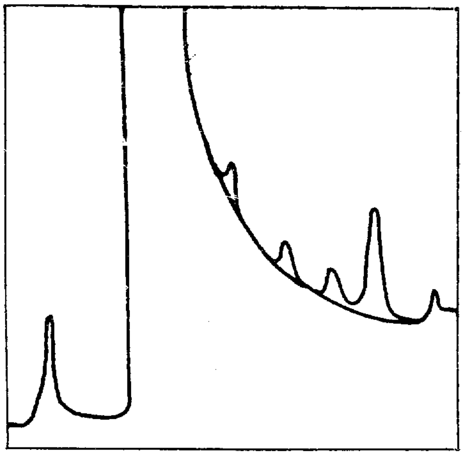

Fig. 7

component peaks with high accuracy and precision.

The skewed Gaussian technique is intended to deconvolute peaks which are more nearly the same size, bigger than ten times difference, and where the perpendicular is not sufficient. It assumes that the individual peaks can be approximated by a skewed Gaussian peak. The hight and width of adjacent fused peaks, corrected for the contribution of each peak on the other, are used to calculate the true area of each peak (figure 8). Peak skewedness is defined by the peak factors stored in the data table. These factors are obtained by analyzing known blends containing the fused peaks. These values of the skewed peak factors are then applied to the unknown samples.

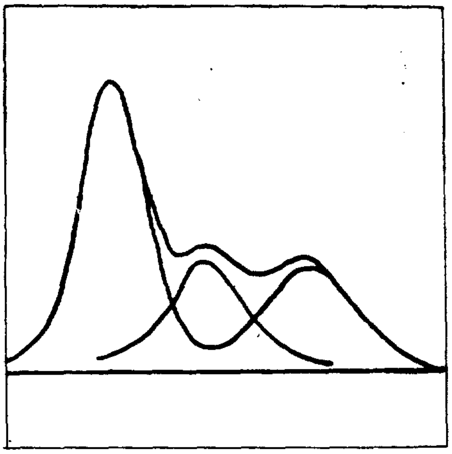

Fig. 8

In many applications, peak shapes are intermediate between that of a

Gaussian functions;

$$Y=Y_0\left\{\mathrm{EXP}\left(-\ln 2\left[\frac{2(X-X_0)}{\Delta X_{\frac{1}{2}}}\right]^2\right)\right\}$$

Cauchy function;

$$Y=Y_0\Big/\left\{1+\left[\frac{2(X-X_0)}{\Delta X_{\frac{1}{2}}}\right]^2\right\}$$

where Y_0 is the peak height

X_0 is the peak position

$X_{\frac{1}{2}}$ is the peak width at half height

One method of dealing with these shape of peak equations is to use a linear combination of the functions in above equations at the ratio of $f:(1-f)$;

$$Y=Yo\left(f\,\mathrm{EXP}\left\{-\ln 2\left[\frac{2(X-X_0)}{\Delta X_{\frac{1}{2}}}\right]^2\right\}\right.$$
$$\left.+(1-f)\Big/\left\{1+\left[\frac{2(X-X_0)}{\Delta X_{\frac{1}{2}}}\right]^2\right\}\right)$$

where $0\leqslant f\leqslant 1$

When asymmetrical peaks are encountered,

$$Y=1/\{\ln(1+bX)/b\}$$

where b is the function to its original symmetrical form. for $b=0$, then expression is indeterminated

$$\lim_{b\to 0}\{\ln(1+bX)/b\}=X$$

as b increases, the peak because more highly skewed, and negative values of b produce negatively skewed peak shapes. Peak identification is achieved by means of retention time. Since column aging, temperature, and flow rate variations cause poor absolute retention time reproducibility, reference peaks are used to calculate relative retentions.

This computer interfaced gas chromatograph systems allows to define as many reference peak as desired. Since a reference peak is defind as the largest peak within a time window, the only restriction is that the reference peak be distinctive. The multiple reference peaks are essential for automatic and unambigous identification of long temperature programmed analyses.

Foreground processes include signal scanning, digital filtering differentiation, peak detection, integration, and peak storage. Run time method parameters which define how foreground tasks are to be accomplished allow the scientist to program an entire analysis. Foreground parameters can be updated automatically as a function of retention times and peak conditions. Background processes have no real time deadlines for their completion. These processes are run in available time blocks, such as result calculations, applications, and report generation (figure 9).

If sufficient care is used in programming and system designing, this computer automation has great

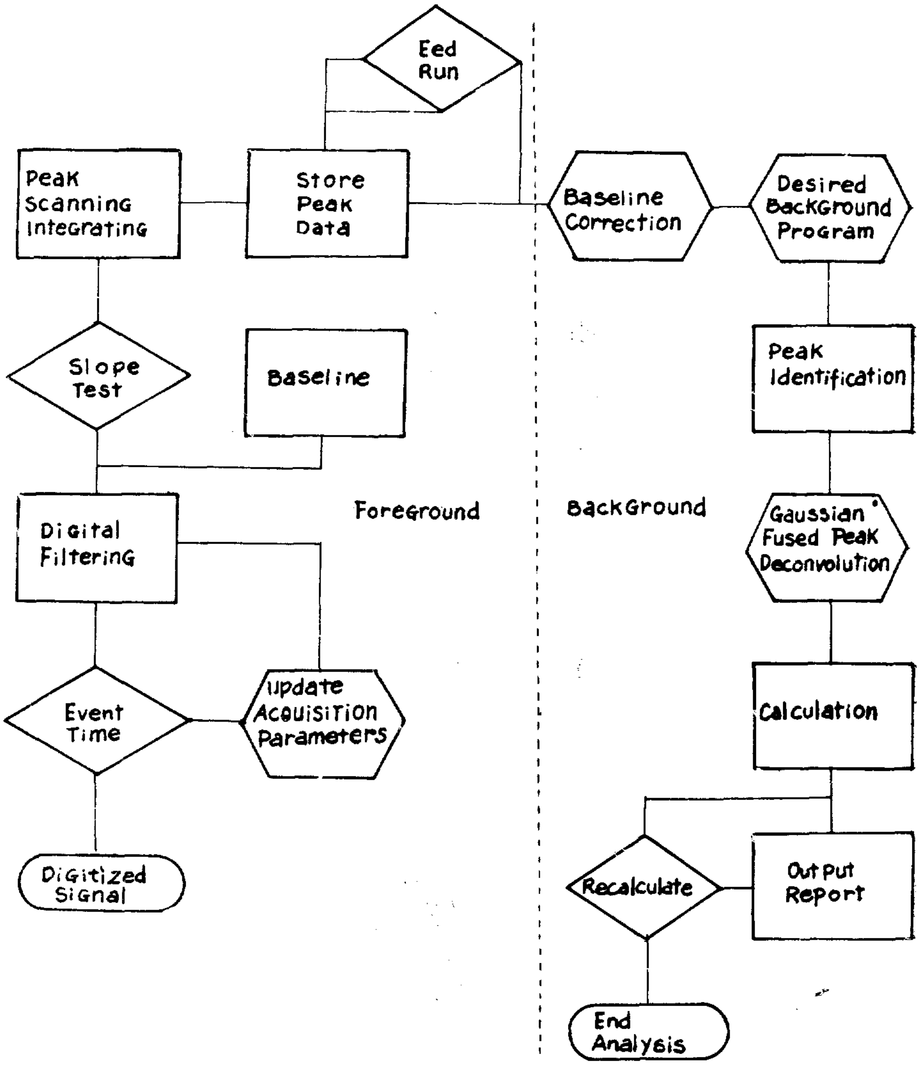

Fig. 9

accuracy and versatility in signal analysis. Although accuracy is obvious performed on the digitized data can be accomplished so that original instrumental accuracy is not lost. This is not generally true in usual instrumental system. This digital computer techniques readily lend themselves to mathematical treatment where by experimental data are smoothed and reproduced to the original capability of the instruments. Another words, the quality of the analysis is not limited by the accuracy of the output devices, for example a strip chart recorder. Accuracy is also improved, as much as human error is removed from the performance of the repetitive routine operations. When the problem is correctly programmed for the computer, all functions and data acquisitions are performed exactly the same every time which will always be maximum for that system. The computer will not changes the order of experimental procedure of forget as a scientist often does in his laboratory.

Another benifit of computer automation is the opening of new experimental possibilities. For example, computer can set data signal from the instrument with analog-to-digital converter. After get the data, computer can process these data and from this processing result, computer can make a decision to the instrument. But computer's output signal is integer value form. These integer values are converted to analog signal by digital-to-analog converter. And this analog signal, another word, control signal goes to instrument, and control the instrumental functions. The ability to data acquisition with on-line computations and fully control the instrument or experiment allows the experiment allows the experimenter to "closed-loop" operation (figure 10). That is the direction of an experiment or analytical measurement can be changed at any time based on all the information obtained up to that instant. The technique of "closed-loop" operations will be a major area of investigation in the near future research laboratory instrumental systems.

Statistical data analysis and handling is another good aspects for computerized system. The increasing usage of instruments and the more sophisticated data provided by these instruments requires more sophisticated analysis of the data. Analyzing the data frequently demands reconsideration of the most fundamental aspects of the phenomena in question. Some times this is more difficult than getting the data. That is why "EVEN A DIAMOND MUST BE POLISHED".

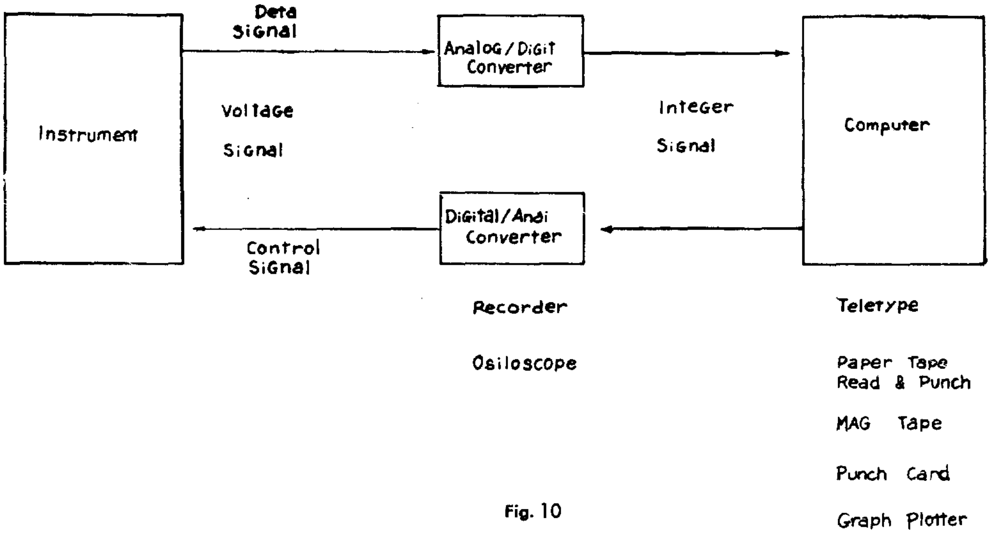

Fig. 10