CONF-8509152--1

UCRL-92975
PREPRINT

Real Time Analysis Under EDS

Daniel Schneberk

This paper was prepared for submittal to
Information Crossroads of the 80's
Washington, D.C.
September 8-13, 1985

July 1985

Lawrence Livermore National Laboratory

UCRL--92975

DE85 016944

Real Time Analysis under EDS*

Dan Schneberk
AVLIS Program, Y-Division
Lawrence Livermore National Laboratory
U. S. A.

Abstract

This paper describes the analysis component of the Enrichment Diagnostic System (EDS) developed for the Atomic Vapor Laser Isotope Separation Program (AVLIS) at Lawrence Livermore National Laboratory (LLNL). Four different types of analysis are performed on data acquired through EDS, i) Absorption spectroscopy on laser-generated spectral lines, ii) mass spectrometer analysis, iii) general purpose waveform analysis, iv) separation performance calculations. The information produced from this data includes: measures of particle density and velocity, partial pressures of residual gases, and overall measures of isotope enrichment. The analysis component supports a variety of real-time modeling tasks, a means for broadcasting data to other nodes, and a great degree of flexibility for tailoring computations to the exact needs of the process.

A particular data base structure and program flow is common to all types of analysis. Key elements of the analysis component are: i) a fast access data base which can configure all types of analysis, ii) a selected set of analysis routines, iii) a general purpose data manipulation and graphics package for the results of real time analysis. Each of these components will be described below with an emphasis upon how each contributes to overall system capability.

*This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

The analysis software completes the EDS operations cycle which begins with data acquisition and concludes with broadcasting data to PMC-1000 and the Process Modeling node. Two goals define the majority of the work for the analysis component; 1) compute and display meaningful measures of process performance, 2) make all EDS analyzed data available for further analysis or control. To accomplish these objectives analysis software divides into two categories; i) data reduction programs which condense acquired vectors into sets of meaningful scalars, and ii) the routines which make up a spreadsheet-type code (ANPRT) which manipulates, plots, and/or sends data to other nodes in the network.

Data reduction programs are specific to the type of data being reduced. Each program has its own applications work area, and its necessarily different functionality. Four different programs are included in EDS at this time: i) a code for performing absorption spectroscopy on spectral lines from diagnostic lasers, ii) codes which perform gas composition calculations on spectra from residual gas analyzers, iii) general purpose waveform analysis on vectors from the extractor diagnostic, iv) calculations of U235 collected from a combination of different types of data. The results of the different programs are: particle densities, velocities, mole fractions, fraction ionized, evaporation rate, partial pressures of residual gases, stripping efficiency, swu rate.

The spreadsheet-type code (ANPRT), manages the reduced data from every diagnostic. Each applications program has its own columns for writing into the matrix, and its own time vector recording when data was taken and analyzed. Any further plotting or manipulation of the scalar data is performed by ANPRT which can work on separate columns or the matrix as a whole. ANPRT includes: a command interpreter which supports batch and interactive processing, macros and subroutines; a set of intrinsic functions tailored to the needs of the process, a full X-Y plot capability using LLNL Device Independent Graphics Library (DIGLIB), numerous data base functions, and a single command for broadcasting data to other nodes.

To perform this work the elements of the analysis component access the shared data structures in the EDS system and follow the general guidelines for program flow. Like other elements of EDS all analysis software depends upon diagnostic data bases for direction and data identification. The relationship between shared data structures and data base provides the groundwork for many features of the analysis system.

Data Base & Program Flow

EDS data bases do not contain the vector data. Every effort is made to avoid manipulating the data until absolutely necessary. Data Bases do contain the information required for identifying, analyzing and the tracking the data flow through the system. Data bases record a history of data acquisition events. They connect program sequences, vectors and scalars to particular instances in the process.

The data base for the diagnostic lasers is the best example. One particular relation[1] in this data base is exclusively devoted to laser data. Each instance of this relation identifies a different wavelength(frequency), and position for a laser shot (wavelength-position). Of the total items of this type in the data base, only a few are active at any one time, but many different wavelength-positions may be used in the course of process operation. Each item is comprised of four types of information; i) basic identifiers, all the labeling information and parameters for analyzing the data, ii) channel information, the connection between the hardware channels and the software locations of the vectors, iii) time information, the times at which that wavelength-position was used and possibly re-used, iv) analysis pointers, the pointers to the functions to be applied to those vectors and the locations for putting the scalars into the reduced parameters matrix.

Every program sequence depends upon information from the data base at each step, through local tables. Figure 1 presents a pictorial of the five-step EDS sequence; i) acquire vector data, ii) copy to applications workspace, iii) reduce vectors to sets of scalars and put into reduced parameters matrix, iv) manipulate and plot from ANPRT, v) broadcast to other nodes. Prior to sequence execution local tables are filled with the most recent settings for key parameters and pointers. To maintain system speed, no program makes data base accesses during a program sequence. However, all key parameters are in the data base with time stamps attached.

---

[1] This paper subscribes to the definitions and conceptual framework found in E.F. Codd's, ' A Relational Model of Data for Large Shared Data Banks', ACM Communications Vol. 13, (6), June 1970.

Figure 2 contains the pictorial for the laser diagnostics sequence. Labels and parameters for the active wavelength-position items go directly into the applications work area, identifying the vectors of data to the applications program. This includes all the information required to perform the analysis: crossection, wavelength, laser path length, and information for calculating scan lengths. The analysis pointers write a set of flags for directing what function(s) to apply to the vectors acquired, and direct the writing of names and labels to the matrix managed by ANPRT. Re-configuring lasers will result in new values written to local tables, with the episode of use for those parameters and labels recorded in the data base.

EDS Data Reduction Software

While the quantity of vector data handled in EDS is substantial the amount of data reduction code is not large. Fundamentally there are two types of vectors; line spectra, and waveforms. Similar data-analytic problems are common across different diagnostics. There is a base set of component tools which covers the variety of different analyses. With different instrumentation, it is entirely possible that the base set could be made much smaller.

Mass Spectrometer Analysis

Partial Pressures of the different gases in the tank are obtained from this diagnostic. A UTI Residual Gas Analyzer is combined with Lecroy front-end hardware to probe the presence of gases in the low mass ranges. Each sweep of the mass range (1-50 AMU) generates a set of line spectra, the peaks corresponding to the different mass units found in the tank. Figure 3 contains a representative trace of the spectra obtained.

Analysis involves three steps, i) resolving the baseline for the spectra, ii) finding the peak heights of the different mass units, iii) un-cracking the peak height intensities into partial pressures of gases.

Resolving the spectra to a baseline of zero guarantees that peak heights are handled correctly, and allows for easy handling of signal to noise problems. In this task, the first job is to find the baseline. There are two alternatives which combine both speed and accuracy; i) find an unoccupied mass range and compute a mean value, ii) find that value in the amplitude of the trace which is most common by discretizing the range of the trace. The second is more general purpose, and yielded good results for our application. Once the baseline is determined, normalize to zero by subtracting the baseline value from the trace.

There are a variety of techniques for peak detection. Our choice of techniques has been driven by the kinds of problems with input ramp and traces. The sweep over the mass range of 1-50 AMU was not always linear. Traces were sometimes noisy, and mass peaks overlapped enough to make peak heights difficult to determine. Our approach made the most of what we knew. Of the entire range only a few mass numbers were important. Peak detection was restricted to the small ranges of interest. Only peaks above the current signal-to-noise threshold were considered. Overlapping peaks were de-convolved by examining the non-adjacent side of the peak. The limits for the small ranges and the signal-to-noise are in the data base, and could be changed during the run if necessary.

Uncracking the peaks into partial pressures of the gases was performed according to research done by project scientists.

Absorption spectroscopy from Diagnostic Lasers

The data reduction code for this diagnostic computes measures of source performance from absorption broadened spectral lines and related data. A number of different laser wavelengths are mixed into a single shot through the vapor. One absorption broadened spectral trace and a variety of different types of data are obtained for each wavelength. Two different types of variables are computed for a diagnostic laser shot; (i) for each wavelength, state-specific particle densities, widths, (Full width at Half Maximum-FWHM), centers, and amplitudes, (ii) from a combination of these numbers, partition functions, and results derived from in-house models: vaporization rates, mole fractions, fraction ions, and measures of source efficiency.

An absorption broadened spectral line is the result of a particular interaction between laser light and the vapor propagating from the source. Under the conditions obtained in our process, the incidence of a spectral line at that frequency indicates an absorbing quantum transition for certain particles in the vapor.

The mathematical description of absorption broadened spectral lines is a direct application of the differential absorption law[2].

$$(1)\quad I(\nu)/I_0(\nu)) = \exp\,(-\,N_{sl} * L * B(\,\sigma(\nu))$$

or more simply,

$$(2)\quad -\ln(I(\nu)/I_0(\nu)) = N_{sl} * L * B(\,\sigma(\nu))$$

where:

$I(\nu)$ - the laser light exposed to the absorbing substance at frequency ν.
$I_0(\nu)$ - the laser light not exposed to the vapor at the same frequency ν.
N_{sl} - the number particles of the vapor in the lower energy state subject to the transition corresponding to the laser wavelength.
sl - the energy level of the lower state, in wave numbers.
$\sigma(\nu)$ - the absorption crossection at frequency ν.
B() - denotes a functional description of a broadening mechanism for that spectral line.

The conditions in our process ensure the predominance of velocity, or Doppler, broadening. Specific measures are taken to keep other broadening effects at a minimum. The shape of the spectral line reflects the intensity of the flow vectors from the source. The directional velocity of the flow, relative to the laser light, determines the shift in off-center resonance. The velocity distribution of the vapor is written into the broadened line shape. In our case the distribution is roughly Maxwellian, permitting the B() in equations (2) & (3) to be replaced by a gaussian as follows.

---

[2] Mitchell & Zemansky, 'Resonance Radiation & Excited Atoms' Cambridge University Press, Cambridge, England, 1934, contains a full analysis of absorption broadened spectral lines.

$$(3)\quad -\ln(I/I_0) = N_{sl} * L * (\sigma(T',\lambda) * \delta(T',\lambda) / \delta(T,\lambda)) * \exp(-4*\ln(2) * ((\nu - \nu_c)/\delta(T,\lambda)))^2$$

where:

$\sigma(T,\lambda)$ - is the crossection at center frequency, wavelength λ, and temperature T.

$\delta(T,\lambda)$ - is the FWHM (full width at half maximum) at temperature T, wavelength λ.

ν_c - the relative center frequency for the line.

T' - the norming temperature for the crossections[3].

The resultant equation is an un-normed gaussian with a number of linear multipliers. Of the many parameters on the right all but N_{sl}, ν_c, and $\delta(T,\lambda)$ are known. It is the object of the data reduction program to estimate these parameters.

Two different types of line spectra are involved in the analysis of diagnostic laser data; i) a frequency calibration trace from the Fabry-Perot cavity, ii) the absorption broadened spectral line. Both the absorption trace and the Fabry-Perot signal are swept with the same ramp function. While the broadened spectral line contains the particle density and velocity information, it is the Fabry-Perot data which provides the frequency scale for the absorption trace.

Each peak in the Fabry-Perot trace demarcates the position around the center frequency of the absorption line for that point in the sweep, see Figure 3(b). The separation between the peaks measures the distance traveled in frequency for the absorption scan, (ie, 300 mhz, 2ghz etc..). Analysis reduces to; i) resolve a baseline, ii) find the peaks, iii) apply the calibration factor to the distance between the peaks and make the frequency scale.

The best procedure involves Fast Fourier Transforms, but reliable results can be achieved with other techniques. In our work the mass spectrometer routines described above work well for both baseline finding and peak detection. In this case the ranges for finding peaks are determined by the intended scan length dialed into the instrumentation for sweeping the laser. The expected scan length and the calibration factor, are both in the data base and in local tables.

---

[3] Our project scientists chose 2015K as the temperature for reporting crossections.

The frequency scale determined, analyzing the absorption line involves three steps, i) resolve a baseline, ii) transform the data and estimate center, amplitude, and width, iii) compute particle density and mach numbers. Again, techniques described above can work here for baseline fitting. The third task is fundamentally a matter of arithmetic. The majority of the work is involved in computing center, amplitude and width.

While there are many alternatives for estimating parameters in gaussian-shaped data, our work has focused on two techniques; i) order statistic estimators, ii) least squares. The first involves estimating quartiles by numerically integrating data which has the form indicated in (3). The second involves an additional log transform of (3), then a polynomial fit of the result over the peak absorption region. The first technique has the advantage of speed. However, the second is more resilent to common data-analytic problems.

At this time our procedure involves a combination of the two approaches. While obtaining the quartiles involves more work, the total area under the curve described by (3) can be computed in the course of least squares procedures. The estimate of particle density computed with the area total is regularly compared with the estimate from least squares and has proven a better goodness-of-fit statistic than the traditional choices. Estimates of density and mach number take just less than a second per line on an A-700.

More meaningful measures of source performance; neutral densities, evaporation rate, mole fractions etc, are computed from sets of single-state densities. Provided two different transitions for the same substance can be mixed into the same shot, standard equations[4] can be applied to compute neutral densities for any substance in the E-Beam source. Combined with a model of vapor propagation, neutral densities can generate all the measures necessary for mass balance and real time film flow modeling.

The extra computations for these results is not large. Neutral densities are obtained from partition functions (a table of statistical uncertainties, and energy levels for the different energy states for a substance). Also, it has been our experience that modeling results can be approximated by polynomials, minimizing the need for lengthy numerical routines.

---

[4] We make the assumption of internal thermal equilibrium and employ the equations for relating energy-state populations which conform to a Boltzmann distribution.

In our system partition function routines exist in two places, in the data reduction programs, and in ANPRT. The applications program can perform the operations faster but without the same flexibility. ANPRT, running from a command file, has access to a variety of modeling results, and can estimate certain parts of equations if trouble develops with a certain laser. This capability has proven invaluable for a variety of untoward conditions, and has made real-time modeling an oft-used feature.

Extractor Waveform Analysis & Separator Performance Calculations

Calculations for both of these diagnostics is performed within a small interpreter with intrinsic functions for combining columns of the applications work area. Key scalars are written to the ANPRT matrix, but the amount of data reduction for this diagnostic is small, and still being developed.

At this point in time waveform analysis is oriented towards computing received current loads on certain parts of the extractor hardware. This can involve some algebra between waveforms taken at different times, but will always involve some quadrature integration of the ion current waveform itself.

Measures of U235 separation performance integrate data from a variety of different diagnostics. Recieved current loads are taken from the extractor diagnostic. Vaporization rates and source efficiencies are provided by the diagnostic lasers. Once this data is moved to the correct applications work area, the calculations are straightforward combinations of columns of data.

ANPRT Capability

The spreadsheet-type interpreter 'ANPRT' has five components, i) command interpreter with macros & argument passing, ii) a processor for equations and if-conditionals equipped with a set of intrinsic functions tailored to real-time modeling needs, iii) full X-Y plot capability to any viewport or device in the system, (all graphics plotting is performed with DIGLIB), iv) a variety of data base functions for manipulating the matrix.

The reduced parameters matrix contains the analyzed data from the data reduction programs and the names and labels from the data base. The central data structure is a real valued matrix, each row a different instance of a data reduction code, the columns a different computed quantity. In addition to the

separate time vectors written by each data reduction code (decimal time relative to start of the run), there is a six-integer time array, one for each case, providing an unambigous time reference for the data. There is a 4-character ascii identifier for each case for further ease in addressing case ranges. Each real-valued column has an associated 16-character name and 40 character label. Figure 4 is a pictorial of this structure. At present the real valued matrix is 500 by 150. Once 500 events have occured the data reduction programs 'wrap-around' the matrix starting at 1.

The reduced parameters matrix resides in shared EMA and can be accessed by many versions of ANPRT. Each applications program keeps track of the number of cases it has written into the matrix in separate EMA variables locked into memory. Each version of ANPRT has local pointers for that session which can access the EMA pointers for each diagnostic on command. In this way there can be many versions of ANPRT which can operate on the same matrix independently.

The command interpreter combines two types of interpretation in the same general structure. Commands begin with a major keyword which identifies the type of interpretation applied to the rest of the command. The major keywords 'CALC' or 'IF' can be followed by equations or if-conditional statements conforming to the rules in BASIC. All other key words are followed by sets of keywords and parameters. Continuation lines and comments are supported. Macros and subroutines are accessed with the 'TR' command for compatiblity with HP-1000 'CI' and 'FMGR' usage. Menus arise as prompts when not enough parameters are supplied for the keyword oriented syntax. In this way the range of different users are accomodated. Experienced users can suppress menus once commands are known, while novices have the benefit of on-line tutotrial help in the menus.

The processor for equations and if-conditionals features a number of built-in functions. This type of interpretation parses expressions into reverse-polish strings of function addresses and data addresses. The two subroutines which perform the work are driven by the tables of variable, scalar, and function names. As mentioned above the syntax for operators and arguments conforms to BASIC-1000 in all major respects. Arguments are columns in the matrix; or user-defined scalars.

Any column in the matrix can be addressed by 16-character name or by a 'VNNN' identifier where 'NNN' is the column number in the matrix, (name or 'SNN' for scalars). The intrinsic functions are selected specifically for the needs of the experiment. In addition to the standard set of functions[5], a complete set of routines is supplied to compute energy state populations for any component in the vapor. Intrinsic functions exist for computing internal temperature, total population, or any specific state population. Figure 5 contains a list of some of the intrinsic functions.

All graphics are implemented with DIGLIB and drivers exist for HP 26XX terminals and Raster technologies graphics devices. While the plotting is performed with one command, 'PLOT', the variety of graphics settings are manipulated with the 'SET' major keyword. To plot any variable(s) the 'PLOT' major keyword is issued followed by one or two variable names or 'VNNN' identifiers, i.e. 'PLOT V5 URATE'. If only one variable is supplied the X-axis is the index of the variable. Every plot is performed with the extant graphics settings for: axis scaling, line type, point style, device, grid display type, color, superpose options, titles, and hardcopy options. Until a new setting is supplied the current value remains in force. Figure 6 has a list of the graphics commands.

Data base operations are performed primarily through other keywords, with case pointers and variable pointers as parameters to those keywords. With these commands it is possible to delete columns, rows, restrict the case range for 'CALC' operation, enter data directly into the matrix, etc,.. With these commands it is possible to focus on any column or row in the matrix.

Curve fitting is also included in the package; polynomial fitting and general purpose linear regression. These commands can prove useful in real-time or in post-run analysis.

From the command interpreter, one row of data can be sent across DS-1000 to PMC with the 'SEND' command, i.e.,(SEND VARS V3, URATE, V4-V7 will send variables 3,4 to 7 and the variable 'urate' for the specified case). Once this command is issued a master program is initiated which ships real values to the slave on another machine. In the event of any problem both the master and slave terminate without affecting ANPRT operation. The user-specified names of the columns are the parameters for the 'SEND' command. Figure 7 contains an example of commands used during process operation.

---

[5] By 'standard functions' I mean the set found in the variety of different languages, SIN,COS,ABS,ACOS,ASIN etc..,.

Conclusion

The EDS Analysis system contains two types of software; data reduction programs and a spreadsheet-type code. They combine to form a system which generates reliable estimates of source performance, extractor performance, gas composition and overall process performance. Further, EDS analysis can support a variety of real time modeling tasks, and bridge the gap between EDS and PMC-1000. EDS is a general solution to the problems which will arise for real time systems that combine scalar value periodic data acquisition with high-speed vector oriented data acquisition and analysis.

EDS analysis is both flexible and powerful. At any one time the analysis of the vector data can be directed or re-directed from the data base. Different types of data reduction can be performed by changing settings in the data base. Once within ANPRT, new variables or more sophisticated measures of system operation can be computed interactively or through a command file. Indeed all the kinds of modeling performed in ANPRT can be performed on an automatic basis. More than one ANPRT session can operate at one time supporting a variety of analysis and monitoring needs during the experiment. The advantage of full data base capability and filtering controls the data entry into PMC on the basis of values computed from other moving quantitites, assuring the integrity data input to control algorithms. The variety of X-Y graphics to any viewport presents a picture of system response over the history of the process.

This effort takes exception to other approaches to real time systems that uniformly sacrifice flexibility for supposed 'speed of operation arguments'. EDS shows that a great deal of flexibility and power can be obtained in a real time system oriented to vector data, within the time intervals assumed by many commercial process monitoring and control systems.

In spite of the fact that EDS can work without operator intervention in fully automatic mode, it is not a desirable plant system. There are disadvantages to too much flexibility. Eventually EDS will be replaced by a number of devices attached to a much larger control system. EDS is an intermediary system which can take a process from experimentation to plant operation. The success of the AVLIS program underscores this point.

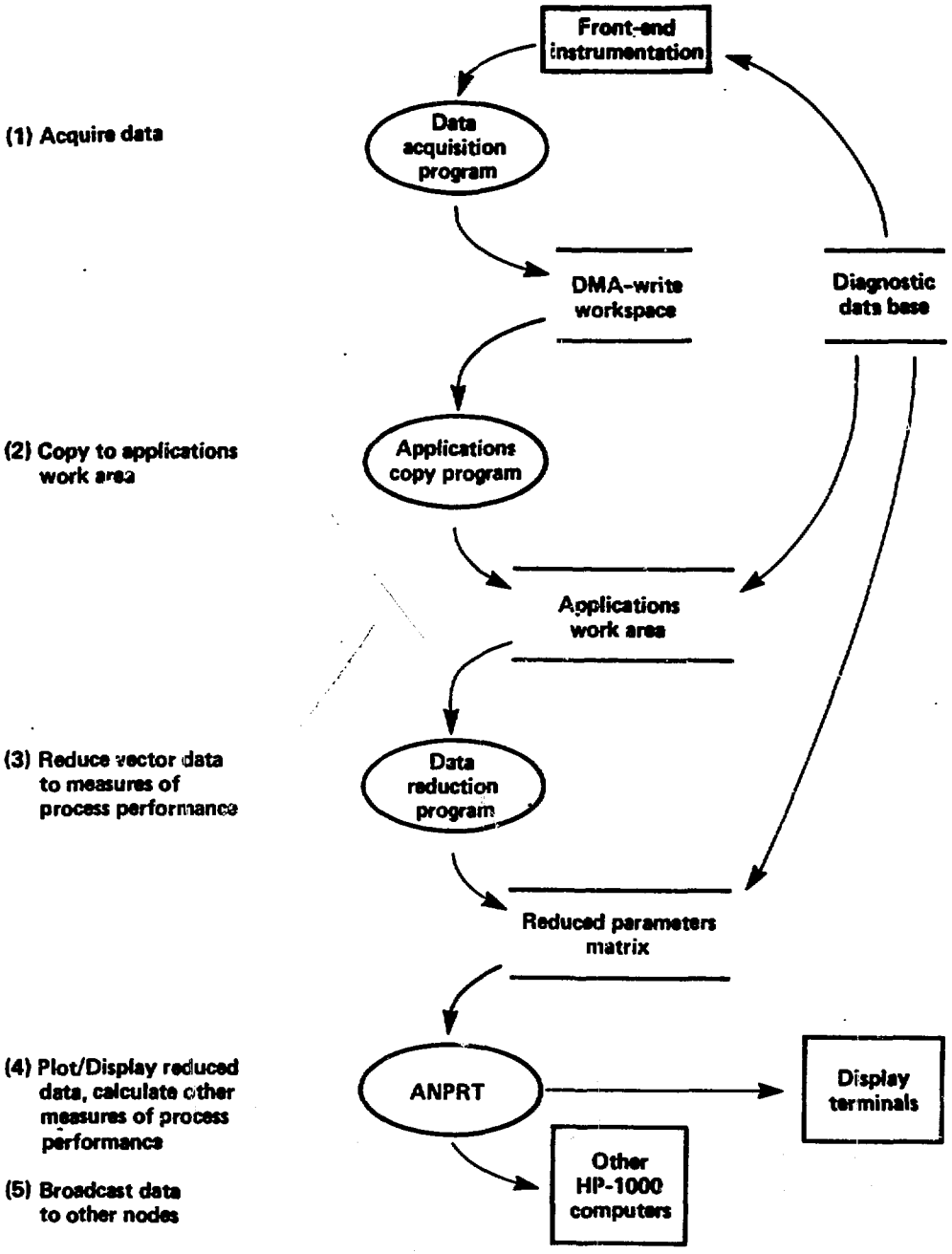

Figure 1. Generalized EDS sequence

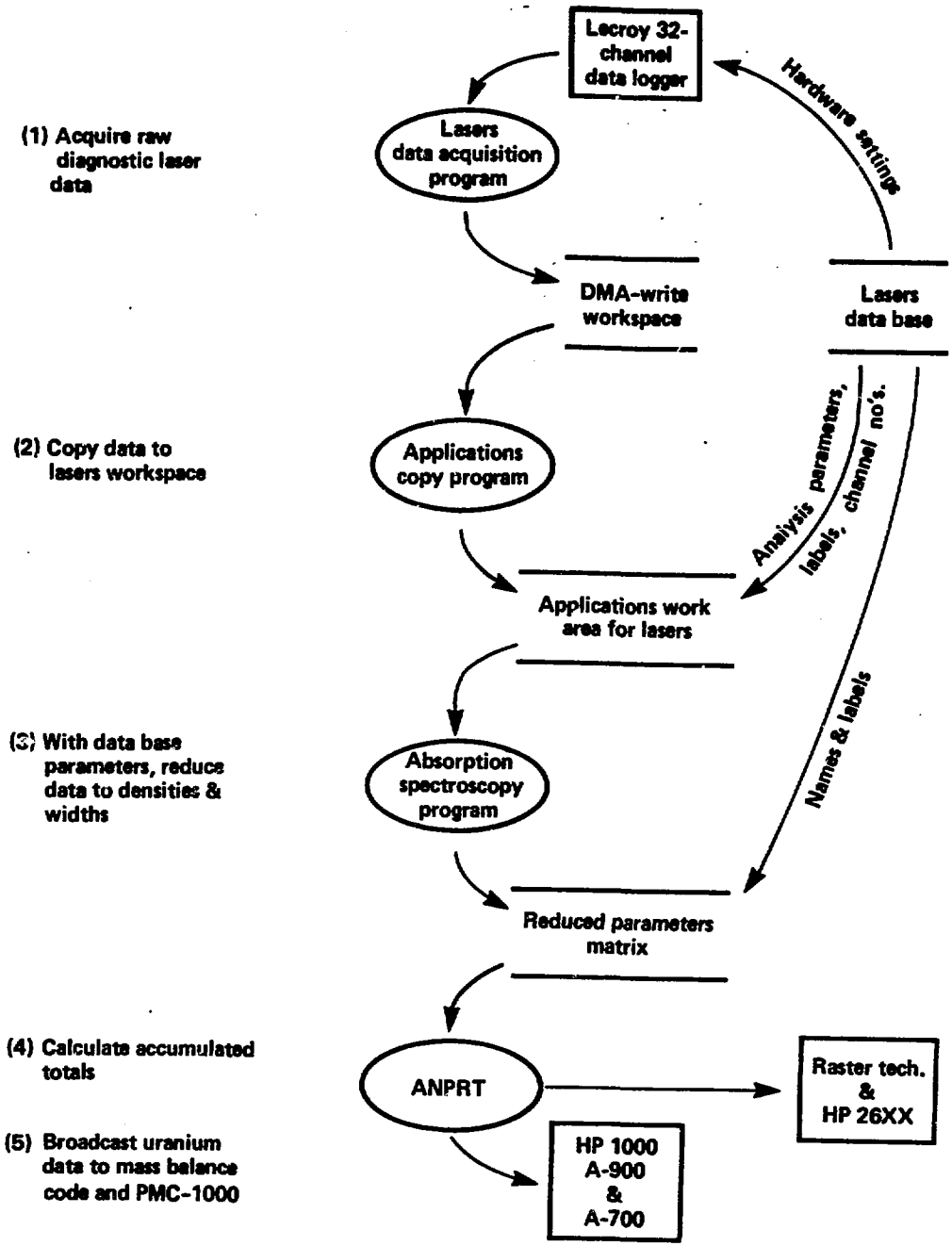

Figure 2. EDS sequence -lasers

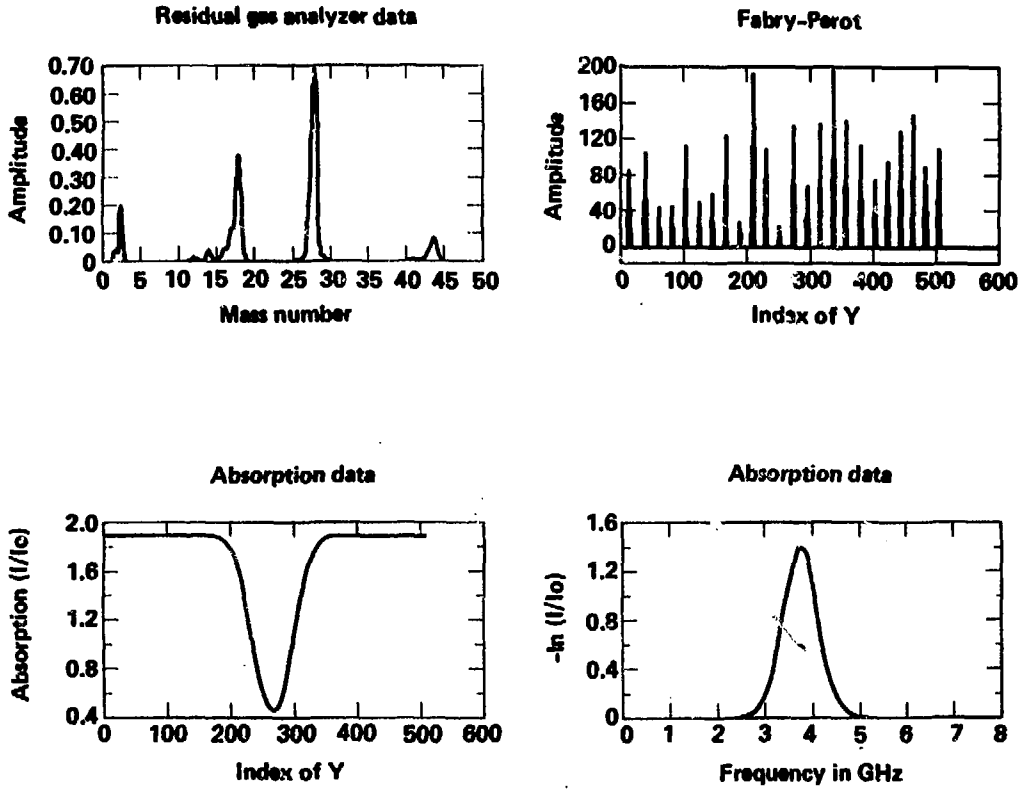

Figure 3. EDS vector data

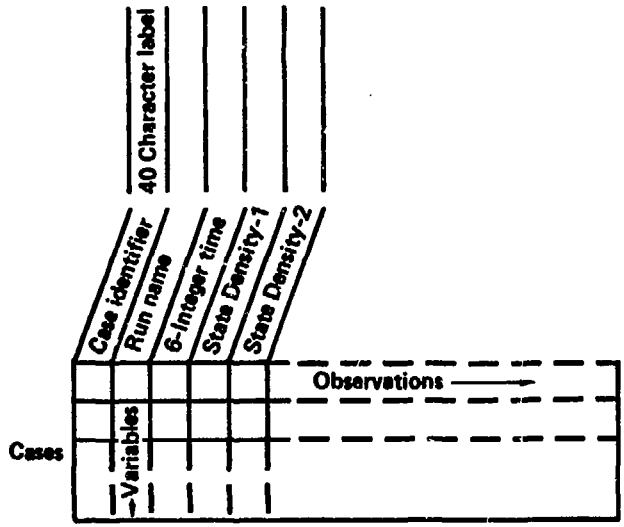

Figure 4. Reduced parameters matrix

Command File Example

** Command files are sent in the run string with the the call to execute ANPRT. The lines below are a sample for a command file used during a run, for the laser data. The '*' denotes a comment line.

```
*
* Tell ANPRT the diagnostic, REFRESH local tables, and set to
*  operate on last case
IAM VAPOR
REFRESH
SET CASES $,$
*
* Compute Uranium internal temperature from ground and 620
* state-specfic densities
CALC UTEMP = TENEUT(V4,V6,0,620)
*
* Calculate Neutral U-density,from ground and temperature
* and calculate URATE with UFACT function from modeling work
CALC UTOT = NTOT(V4,0,UTEMP)
CALC URATE = UTOT * UFACT
*
* Accumulate URATE from beginning of data
SET CASES 1,$
CALC U_ACCUM = RIEM(TIME,URATE)
*
* Send data to other nodes
SEND VARS UTEMP,UTOT,URATE,UACCUM
*
* Plot last 40 cases on viewport A
*  Auto scale data for X and Y Axis
*  Filter data with IF statement
*  Set Point style,and Line style and plot
SET CASES -40,$
SET VIEW A
SET AXIS AUTO BOTH
IF (URATE GT 0) AND (URATE LT 110)
SET POINT I
SET LINE A
PLOT TIME URATE
ENDIF
**
** all done
EX
```

FIGURE 7

ANPRT GRAPHICS SETTINGS

* All graphics setting are accessed with the 'SET' major keyword, combined with the appropriate minor keyword and option

Keyword	Description	Example
VIEW	Select Viewport, there are 8 available (all settings apply to the selected viewport)	SET VI A
POINT	Point Style Options, there are 12 styles	SET POINT A
LINE	Line Style Options there are 10 types	SET LI B
AXIS	Axis settings, can set for autoscaling or for explicit axis scale	SET AX Y 0 1 0.1 0.02 (set y-axis explicitly) SET AX A Y (set y-axis to autoscale)
LOG	Log axis option	SET LO A
PEN	Pen color options	SET PEN A
GRID	Turn on grid	SET GRID ON
SU	Superimpose Option	SET SU ON
DUMP	Automatic hardcopy	SET DUMP ON

FIGURE 6

ANPRT INTRINSIC FUNCTIONS

Function	Description	Example*
LOG	Computes Natural Log	V3 = LOG(V4 + 9)
EXP	Exponential Function	NEW1 = EXP(V23)
SIN	Sine Function (in radians)	V10 = SIN(V6)
COS	Cosine Function	V11 = COS(V21 * 0.44)
TAN	Tangent Function	V16 = TAN(theta)
SQRT	Square Root	SCAL1 = SQRT(4.56)
SUM	Calculates sum of V10 over specified case range	SCAL2 = SUM(V10)
ABS	Absolute value	V3 = ABS(V14)
ASIN	Arc-Sine function	V10 = ASIN(V13)
ACOS	Arc-Cosine function	V13 = ACOS(V104)
ATAN	Arc-tangent function	V12 = ATAN(V14)
INDX	Puts case index into V8	V8 = INDX
GRAB	Grabs variable V3 from case 1 to current range	V12 = GRAB(1,V3)
ERF	Error Function	V15 = ERF(V6)
SMTH	Performs 11-point quadratic smoothing	V16 = SMTH(V15)
TRAP	Trapezoid rule integration on V10 using V3	V15 = TRAP(V3,V10)
RIEM	Riemann sum on V11 using V3	V19 = RIEM(V3,V11)
TENEUT	Computes uranium internal temperature from two state specific densities	V10 = TENEUT(V3,V4,0,620)
NTOT	Computes Neutral Uranium density from one state specific density and a temperature	V12 = NTOT(V3,0,V10)
EXPOP	Computes an expected population from neutral density,temperature and an energy level	V15 = EXPOP(V12,V10,3801)

* - in ANPRT all of these commands would br preceded by the major keyword 'CALC'.

FIGURE 5