Computer Programs for the Evaluation of Low Resolution Mass Spectrometry Data

V. Zitko

Biological Station,
St. Andrews, N.B., E0G 2X0

August 1981

Canadian Technical Report of Fisheries and Aquatic Sciences No. 1025

Government of Canada
Fisheries and Oceans

Gouvernement du Canada
Pêches et Océans

Canadian Technical Report of
Fisheries and Aquatic Sciences 1025

August 1981

COMPUTER PROGRAMS FOR THE EVALUATION OF LOW RESOLUTION MASS SPECTROMETRY DATA

by

V. Zitko

Fisheries and Environmental Sciences
Department of Fisheries and Oceans
Biological Station
St. Andrews, New Brunswick E0G 2X0

This is the one hundred and thirty-ninth Technical Report from
the Biological Station, St. Andrews, N.B.

Cat. No. Fs 97-6/1025 ISSN 0706-6457

Correct citation for this publication:

Zitko, V. 1981. Computer programs for the evaluation of low resolution mass spectrometry data. Can. Tech. Rep. Fish. Aquat. Sci. 1025, iii + 28 p.

ABSTRACT

Zitko, V. 1981. Computer programs for the evaluation of low resolution mass spectrometry data. Can. Tech. Rep. Fish. Aquat. Sci. 1025, iii + 28 p.

Listing and documentation of seven programs for mass spectrometry data manipulation are presented. The programs are written in FINNIGAN BASIC Revision I. Two additional programs for the assignment of elemental formulae and for the calculation of isotopic abundancies are documented. These programs are written in the BASIC version of the Tektronix Plot 50 system.

Key words: Mass spectra, reconstruction, isotope, abundance, PCB, PBB, organochlorine, pesticides, quantitation

RÉSUMÉ

Zitko, V. 1981. Computer programs for the evaluation of low resolution mass spectrometry data. Can. Tech. Rep. Fish. Aquat. Sci. 1025, iii + 28 p.

On présente une liste et l'exposé détaillé de sept programmes pour la manipulation des données de spectrométrie de masse. Les programmes sont rédigés en FINNIGAN BASIC Revision I. Deux programmes supplémentaires pour l'affectation des formules élémentaires et pour le calcul de l'abondance des isotopes sont détaillés. Ces programmes sont rédigés la version BASIC du système Tektronix Plot 50.

INTRODUCTION

Programs used routinely to evaluate gas chromatography-mass spectrometry (GCMS) data obtained on a Finnigan 1015 mass spectrometer and processed by a Finnigan Model 6100 Data System are presented. The programs are generally applicable to GCMS instrumentation with small (16K) central processors. Larger data systems allow the use of more complex software.

The programs assume that GCMS data are stored as nominal mass-intensity arrays, G(I), where I = nominal mass, G(I) = intensity, one array corresponding to one scan of the specified mass range (usually 50-500 m/z). A file consists of 200-300 scans (maximum 1000), and corresponds to one gas chromatographic analysis. In the Model 6100 Data System, files are stored on the disc and are read for processing by the DREAD statement. The central processor has space for only a few scans. Consequently, most of the programs use only two scans at a time. One is the scan of interest detected, for example, as a peak in the reconstructed gas chromatogram or limited mass reconstructed gas chromatogram. The other scan is that of the "baseline," immediately preceding the peak. Their difference is the "background-corrected" mass spectrum. Reconstructed gas chromatogram is an array R(I), where I = scan number, R(I) = summed intensity of all ions in scan I. In a limited mass reconstructed gas chromatogram, R(I) is the intensity of a specified ion in scan I.

A program detecting peaks and backgrounds in the reconstructed gas chromatogram can be written. However, in practice, it is better to examine the chromatogram manually, because of its complexity, and to let the computer perform only the subsequent, more routine tasks, requiring less judgment.

In addition to programs dealing with mass spectral data retrieval and processing, two programs aiding the interpretation of mass spectra are presented. For given masses, these programs derive molecular (or ion) formulae and calculate their isotope distributions. These programs also assume nominal masses and unit resolution, but could be modified easily for more accurate masses and a higher resolution.

PROGRAM MSPRTC

Program MSPRTC retrieves and prints background-corrected mass spectra. The presented version prints "major ions" with intensities ≥1% and "minor ions" with intensities ≥3 standard deviations from the background mean intensity. Such detailed presentation is needed only in some applications. In most cases, "minor ions" are not needed and their output may be eliminated by placing RETURN in line 650. The minimum intensity of "major ions" is determined by line 620. For many applications a threshold of 5% may be sufficient (V < 5).

Scan numbers of mass spectra to be retrieved, followed by their respective backgrounds, are entered as DATA in line 20-44 and "0" is used to indicate termination of the data.

When running the program, the information in lines 3-7 is printed and the program stops to give the user a chance to enter or modify scan numbers and name of the file. The latter is done by typing 10 FILES abc (Return). Program is then restarted by the command GOTO 10 (Return). After this, user's input is still needed in line 12 to specify the desired mass range (such as 50,500).

The arrays M(1000) and G(1000) are set to 0's (lines 45 and 50) and the first DATA value is read (line 80). If it is 0, the program ends (line 85); otherwise, the next value (Y), assumed to be background scan number, is read.

In the data system mass spectra are stored as one-dimensional arrays of intensities with m/z values as indices. The spectrum and background arrays are read from the disc by commands DREAD #1 (lines 100, 110, respectively); #1 refers to the file; 1 may be replaced by an integer variable if the program is to retrieve spectra from more than one file. In such case a loop for this variable must be added. In practice, however, the work is most frequently done on only one file at a time.

The background correction (subtraction of the "background" array Y from the spectrum array X) is accomplished in line 120.

The rest of the program (lines 500-850) is a subroutine, normalizing (intensity of the most intense ion = 100) and printing the mass spectra.

First the heading is printed, SM (scan number of maximum), SB (scan number of background) in line 500, and the absolute intensity of the base ion (line 590). This value is useful to assess the overall intensity of the spectrum. The base ion is found by the K loop in lines 530-580. The spectrum is then normalized and printed by the K loop in lines 600-640. The printout is limited to the integer part of the normalized intensity (line 610).

Level of background "noise" is determined in lines 670-780. For the averaging, intensities larger than 50 (absolute count) and negative intensities are excluded (lines 690, 700). The latter may occur as the result of the background subtraction. Mean and standard deviations are calculated (lines 740-760) and printed (lines 770, 780).

Masses and intensities of "minor" ions are printed in lines 790-830. Ions with intensities less than three times standard deviation (line 800) are not considered. Intensities are then normalized again (line 801). Ions with intensities above 1% are excluded (line 805) since these have been printed out already. The remaining masses, and intensities multiplied by 100 (line 810), are printed.

The background-corrected spectrum is normalized twice during the execution of the program (lines 610 and 801). This is to save memory. Execution of the program is slowed down by this only negligibly.

Plots of mass spectra are easier to scan visually than prints, but the printer used on our system is more reliable in unattended operation than the plotter. The program may be modified easily to plot rather than print mass spectra.

PROGRAM MSPRTC

```
1 DIM M(1000),G(1000)
3 PRINT "PROGRAM MSPRTC"
5 PRINT "ENTER 20-44 DATA,MAX.FOLLOWED BY BACKGR."
7 PRINT "10 FILES --- GOTO 10"
8 STOP
10 FILES LSWA2
11 PRINT "ENTER MASS RANGE"
12 INPUT K1,K2
20 DATA 126,123,136,133,0
45 MAT M=ZER
50 MAT G=ZER
80 READ X
85 IF X=0 THEN 150
90 READ Y
100 DREAD #1,X,M
110 DREAD #1,Y,G
120 MAT M=M-G
130 GOSUB 500
135 GOTO 80
150 END
500 PRINT "SM=";X;"SB=";Y
520 LET C=D=0
530 FOR K=K1 TO K2
540 LET C=M(K)
550 IF C<D THEN 580
560 LET D=C
580 NEXT K
590 PRINT D
600 FOR K=K1 TO K2
610 LET V=INT(100*M(K)/D)
620 IF V<1 THEN 640
630 PRINT K;V,
640 NEXT K
650 PRINT
660 PRINT "MINOR PEAKS,NORMALIZED *100"
670 LET S=S1=0
680 FOR K=K1 TO K2
690 IF M(K)<0 THEN 730
700 IF M(K)>50 THEN 730
710 LET S=S+M(K)
720 LET S1=S1+M(K)↑2
730 NEXT K
740 LET N=K2-K1+1
750 LET S=S/N
760 LET S1=((S1-N*(S↑2))/(N-1))↑.5
770 PRINT "BASELINE=";S
780 PRINT "DEVIATION=";S1
790 FOR K=K1 TO K2
800 IF M(K)<3*S1 THEN 830
801 LET V=INT(100*M(K)/D)
805 IF V>=1 THEN 830
810 LET V=INT(10000*M(K)/D)
820 PRINT K;V,
830 NEXT K
840 PRINT
850 RETURN

*
```

PROGRAM OUT4

This program retrieves and prints background-corrected mass spectra and keeps records of limited mass reconstructed gas chromatograms (LMRGC) leading to a given set of scan numbers. The spectrum retrieval and printout algorithm is the same as in MSPRTC. In the presented version, the output is limited to "major" ions with normalized intensity ≥5% (line 2110) in a mass range determined by lines 2020 and 2090.

The examination of data begins usually by retrieving mass spectra of major components appearing as peaks in the reconstructed gas chromatogram. MSPRTC is suitable for this. The next step is to locate compounds whose spectra contain specified ions. This is done by reconstructing chromatograms of those ions only (limited mass reconstructed gas chromatogram as opposed to reconstructed gas chromatogram summing up intensities of all ions). The limited mass reconstructed gas chromatograms may contain several peaks and it may not be immediately obvious which one corresponds to the searched-for compound. At the same time it may be interesting to look at mass spectra of the other peaks as well.

Routinely, searches according the specified masses are performed for at least 10-15 compounds. Because of the large number of retrieved mass spectra, the reasons for obtaining a particular spectrum may be lost easily. Program OUT4 keeps track of limited mass reconstructed gas chromatograms leading to the retrieval of spectra.

The program uses the standard sign on procedure (lines 1-9), allowing for data entry and file specification. The execution, stopped in line 9 is resumed by the command GOTO 10.

Data must be entered as specified in line 4, in the following format: mass of the limited mass reconstructed gas chromatogram (in the listing abbreviated to RGC), up to 5 maximum and background scan numbers, "1" terminating the scan number set for a given mass, and "0" terminating the whole set of data. The size of the input is limited by user's work area. For the Finnigan 6100 Data system, software revision I, the maximum possible data array is A(60,12), line 500.

The array A is initialized (line 510) and the data are read by the I,J loops (lines 520-585). N (line 580) counts rows of A actually used for data. Lines 590 and 595 are to assure the user that the program is running as expected.

For easier orientation, the data are sorted according to increasing masses (lines 600-710) and the user is informed (line 720). The sorted data are printed to serve as an index of the retrieved mass spectra (lines 730-780).

The program then proceeds to retrieve and print mass spectra by the I loop (lines 800-880) for limited mass reconstructed gas chromatograms with a nested J loop for spectra retrieved by a given mass. The spectra are retrieved and printed by the routine in lines 1000-2150. This is the same routine as in MSPRTC.

PROGRAM RTRV1

This program constructs mass spectra in given scans from ions whose intensity has a maximum within a specified window around the scans. The program is based on an algorithm developed by J. E. Evans (Finnigan Corporation).

This type of spectrum retrieval complements the information provided by background-corrected mass spectra as retrieved, for example, by MSPRTC. On pure compounds, both programs yield very similar mass spectra. For mixtures of compounds, comparative examination of spectra provided by these programs may give a better insight into the composition of the mixture or at least may alert the user to the possibility that a mixture of compounds is present.

The program uses the standard sign on procedure (lines 1-9). Following data entry and file specification, the execution is resumed by GOTO 10.

The array P(500) in line 15 is a "scan number register." Its use will become obvious in the description of the algorithm.

User's input is required in lines 510 (scan range) and 530 (window). The latter specifies which ions will be included in the mass spectrum. For W = 0, only ions with maximum in the specified scan S will be included. For W = -1(1), ions with maxima in scans S-1 (S+1) or S will be included.

The spectrum construction algorithm is given in lines 1000-1170. The obtained mass spectrum is then normalized (lines 1200-1270) and printed (lines 1290-1340). In the presented version, only ions with normalized relative intensity ≥5% are printed (line 1310). This sensitivity should be adjusted according to needs.

In the spectrum construction algorithm, the arrays M,G, and P are initialized (lines 1000-1020). In the Y loop (lines 1030-1110), 5 spectra (S-2 to S+2) are read successively from the disc into the array M (line 1040). In the nested I loop (lines 1050-1100), the intensities in the array M are compared with those in the array G within the specified mass range (Z to Z1). Higher values are placed into G (line 1080), recording their scan numbers in P.

The array M is then initialized (line 1120) and the spectrum is reconstructed according to the specified window (W) in lines 1130-1170. For example, if W = 1, according to 1140 the program execution continues to 1150. If the scan number, stored in P, is S, the intensity is placed into the array M (line 1160). If the scan number is not S, control is transferred to line 1154 and, if the scan number is S+1, the intensity is placed in M.

PROGRAM RTRVS1

This program retrieves mass spectra composed of ions whose intensities have maxima within a specified window around a given scan. Record is kept of masses leading to the retrieval of the spectra. The algorithms used are those of RTRV1 and OUT4, respectively.

PROGRAM OUT4

```
1 DIM M(1000),G(1000)
2 PRINT "PROGRAM OUT4"
3 PRINT "PROGRAM RETRIEVES SPECTRA SORTED ACCORDING TO RGC"
4 PRINT "ENTER 100-199DATA RGC,SM,SB,1,RGC,SM,SB...1,0"
5 PRINT "ENTER MASS RANGE 2020,2090,FOR K=...TO ...."
6 PRINT "AVAILABLE SPACE 60 RGC,5(SM,SB)PER RGC"
8 PRINT "10 FILES ...../ GOTO 10"
9 STOP
10 FILES LSWB2
15 PRINT "READING BEGINS"
100 DATA 168,107,104,123,118,136,134,142,140,146,143,1
101 DATA 196,141,138,145,143,148,146,148,143,1
102 DATA 204,179,177,192,187,200,197,1
103 DATA 189,171,166,190,184,201,198,1,0
110 DATA 300,301,302,303,304,305,306,1,200,201,202,1,0
500 DIM A(60,12)
510 MAT A=ZER
515 LET N=0
520 FOR I=1 TO 60
530 FOR J=1 TO 12
540 READ A
541 LET A(I,J)=A
545 IF A(I,J)=0 THEN 590
550 IF A(I,J)=1 THEN 580
560 NEXT J
580 LET N=N+1
585 NEXT I
590 PRINT "DATA MATRIX CONSTRUCTED"
595 PRINT "N=";N
600 DIM R(12),S(12)
610 FOR K=1 TO N-1
620 FOR I=1 TO N-K
630 FOR J=1 TO 12
640 LET R(J)=A(I,J)
650 LET S(J)=A(I+1,J)
660 NEXT J
670 IF R(1)<S(1) THEN 700
680 FOR J=1 TO 12
685 LET A(I,J)=S(J)
690 LET A(I+1,J)=R(J)
695 NEXT J
700 NEXT I
710 NEXT K
720 PRINT "DATA SORTED"
730 FOR I=1 TO N
740 FOR J=1 TO 12
750 PRINT A(I,J);
760 NEXT J
770 PRINT
780 NEXT I
800 FOR I=1 TO N
810 PRINT TAB(10);"RGC";A(I,1)
820 FOR J=2 TO 12 STEP 2
830 IF A(I,J)=1 THEN 880
840 LET X=A(I,J)
850 LET Y=A(I,J+1)
860 GOSUB 1000
870 NEXT J
875 PRINT TAB(30);"----------"
880 NEXT I
900 END

*

1000 MAT M=ZER
1010 MAT G=ZER
1020 DREAD #1,X,M
1030 DREAD #1,Y,G
1040 MAT M=M-G
1050 GOSUB 2000
1060 RETURN
2000 PRINT "SM=";A(I,J),"SB=";A(I,J+1)
2010 LET C=D=0
2020 FOR K=70 TO 450
2030 LET C=M(K)
2040 IF C<D THEN 2070
2050 LET D=C
2070 NEXT K
2080 PRINT D
2090 FOR K=70 TO 450
2100 LET V=INT(100*M(K)/D)
2110 IF V<5 THEN 2130
2120 PRINT K;V,
2130 NEXT K
2140 PRINT
2150 RETURN

*
```

PROGRAM RTRV1

```
1 DIM M(1000),G(1000)
2 PRINT "PROGRAM RTRV1"
3 PRINT "RETRIEVES SPECTRA WITH NO BACKGROUND SUBTRACTION"
4 PRINT "SPECTRA ARE CONSTRUCTED FROM MAXIMIZING PEAKS"
5 PRINT "ENTER MAXIMA 20-80 DATA,END BY 0"
6 PRINT "DATA SPACE (100)"
7 PRINT"10 FILES.../GOTO 10"
9 STOP
10 FILES LSWB1
15 DIM P(500)
20 DATA 157,161,172,184,187,193,200,211,224
21 DATA 231,236,243,252,260
22 DATA 0
500 PRINT "ENTER THE MASS RANGE"
510 INPUT Z,Z1
520 PRINT "ENTER THE WINDOW (-1,0,+1)"
530 INPUT W
540 READ S
550 IF S=0 THEN 700
590 GOSUB 1000
600 GOTO 540
700 END
1000 MAT M=ZER
1010 MAT G=ZER
1020 MAT P=ZER
1030 FOR Y=S-2 TO S+2
1040 DREAD #1,Y,M
1050 FOR I=Z TO Z1
1060 IF M(I)>G(I) THEN 1080
1070 GOTO 1100
1080 LET G(I)=M(I)
1090 LET P(I)=Y
1100 NEXT I
1110 NEXT Y
1120 MAT M=ZER
1130 FOR I=Z TO Z1
1140 IF W<>-1 THEN 1150
1142 IF P(I)=S-1 THEN 1160
1150 IF P(I)=S THEN 1160
1152 IF W<>1 THEN 1170
1154 IF P(I)<>S+1 THEN 1170
1160 LET M(I)=G(I)
1170 NEXT I
1200 PRINT "SCAN";S
1210 LET C=D=0
1220 FOR K=Z TO Z1
1230 LET C=M(K)
1240 IF C<D THEN 1270
1250 LET D=C
1270 NEXT K
1280 PRINT D
1290 FOR K=Z TO Z1
1300 LET V=INT(100*M(K)/D)
1310 IF V<5 THEN 1330
1320 PRINT K;V,
1330 NEXT K
1340 PRINT
1350 RETURN

*
```

PROGRAM RTRVS1

```
1 DIM M(1000),G(1000)
2 PRINT "PROGRAM RTRVS1"
3 PRINT "RETRIEVES SPECTRA WITH NO BACKGROUND SUBTRACTION"
4 PRINT "SPECTRA ARE CONSTRUCTED FROM MAXIMIZING PEAKS"
5 PRINT "ENTER 20-80DATA RGCLM,SM,SM,...(MAX 9),1,END BYO"
6 PRINT "DATA SPACE IS (40,10)"
7 PRINT "10 FILES.../GOTO 10"
9 STOP
10 FILES LSWA2
15 DIM P(500)
16 DIM A(40,10),R(10),S(10)
20 DATA 284,139,186,207,1,109,54,200,1,194,158,171,182,1,0
100 MAT M=ZER
110 LET N=0
200 FOR I=1 TO 40
210 FOR J=1 TO 10
220 READ A
230 LET A(I,J)=A
240 IF A(I,J)=0 THEN 285
245 IF A(I,J)=1 THEN 275
250 NEXT J
275 LET N=N+1
280 NEXT I
285 PRINT "DATA MATRIX CONSTRUCTED"
290 PRINT "N=";N
300 FOR K=1 TO N-1
310 FOR I=1 TO N-K
320 FOR J=1 TO 10
330 LET R(J)=A(I,J)
340 LET S(J)=A(I+1,J)
350 NEXT J
360 IF R(1)<S(1) THEN 390
370 FOR J=1 TO 10
375 LET A(I,J)=S(J)
380 LET A(I+1,J)=R(J)
385 NEXT J
390 NEXT I
395 NEXT K
400 FOR I=1 TO N
410 FOR J=1 TO 10
420 IF A(I,J)=1 THEN450
430 PRINT A(I,J);
440 NEXT J
450 PRINT
460 NEXT I
500 PRINT "ENTER THE MASS RANGE"
510 INPUT Z,Z1
520 PRINT "ENTER THE WINDOW (-1,0,+1)"
530 INPUT W
600 FOR E=1 TO N
605 PRINT TAB(20);"RGCLM";A(E,1)
606 PRINT TAB(15);"----------------"
610 FOR F=2 TO 10
620 IF A(E,F)=1 THEN 670
630 IF A(E,F)=0 THEN 700
640 LET S=A(E,F)
650 GOSUB 1000
660 NEXT F
670 NEXT E
700 END

*

1000 MAT M=ZER
1010 MAT G=ZER
1020 MAT P=ZER
1030 FOR Y=S-2 TO S+2
1040 DREAD #1,Y,M
1050 FOR I=Z TO Z1
1060 IF M(I)>G(I) THEN 1080
1070 GOTO 1100
1080 LET G(I)=M(I)
1090 LET P(I)=Y
1100 NEXT I
1110 NEXT Y
1120 MAT M=ZER
1130 FOR I=Z TO Z1
1140 IF W<>-1 THEN 1150
1142 IF P(I)=S-1 THEN 1160
1150 IF P(I)=S THEN 1160
1152 IF W<>1 THEN 1170
1154 IF P(I)<>S+1 THEN 1170
1160 LET M(I)=G(I)
1170 NEXT I
1200 PRINT "SCAN";S
1210 LET C=D=0
1220 FOR K=Z TO Z1
1230 LET C=M(K)
1240 IF C<D THEN 1270
1250 LET D=C
1270 NEXT K
1280 PRINT D
1290 FOR K=Z TO Z1
1300 LET V=INT(100*M(K)/D)
1310 IF V<5 THEN 1330
1320 PRINT K;V,
1330 NEXT K
1340 PRINT
1350 RETURN

*
```

Lines 1-9 contain the standard sign on procedure. The format of data entry is mass used to retrieve the scan number, and up to nine scan numbers. These entries are separated by "1" and the data input is terminated by "0."

Because of a longer program, the data array A(40,10) is smaller than in OUT4, but can store the same amount of scan numbers since there is no need for scan numbers of background corrections.

The data are read by lines 110-280 and sorted according to increasing retrieval mass in lines 290-460. Following the sorting, the user is asked for mass range (lines 500-510) and window (lines 520-530). The E,F loops (lines 600-670) select the scan numbers. Spectra are retrieved, normalized, and printed by the subroutine in lines 1000-1350. A detailed description of the subroutine is presented with the program RTRV1.

PROGRAM RGEVA4

This program is particularly useful for visual inspection of isotopic clusters of organochlorine and organobromine compounds. The clusters are "plotted" rather than printed by the teletype. Recognition is easier for a plotted pattern than for a printed series of numbers.

The program plots intensities of ions in up to three mass ranges of background-corrected mass spectra. This is a convenient presentation for the confirmation of organochlorine and organobromine compounds. The mass ranges may be specified to include the molecular ion cluster and additional characteristic clusters (for example, those at M-35 and M-70 for organochlorine compounds). The program also keeps record of file examined and of masses leading to the retrieval of the spectra. As discussed in connection with the program OUT4, this is an important feature for data interpretation.

For a given retrieval mass, up to five maximum and background scan number pairs may be specified. Since the size of the data array is fixed as D(20,17) in line 13, blank spaces within used rows (the first 11 rows, line 16) must be filled with "0"'s. For example, line 30 DATA means retrieval mass 419, mass range 415-430, two additional mass ranges unspecified (filled with "0"'s), three maximum/background scan number pairs (151,148; 263,260; 269,265), and two unspecified pairs (four "0"'s). This line was entered because the reconstructed gas chromatogram of mass 419, obtained manually and inspected visually on the system's display, contained three maxima (151, 263, and 269). It was of interest to determine whether these maxima contained ions of organochlorine or organobromine compounds. The isotopic clusters would be detected by examining the mass range 415-430.

The sign on procedure (lines 1-10) is standard. the data are read in lines 50-57. The user is asked for the name of the file (lines 60-70) for easy identification of the printout.

Mass spectra are retrieved in lines 150-290, with the background correction performed in line 210. Mass ranges are read by the K loop in lines 230-280.

The subroutine in lines 500-720 normalizes the spectra within the specified mass ranges. The spectra are plotted by the L loop in lines 660-680 by printing the character "-" up to Q times. This variable is normalized to 60 for the base ion (lines 620-630). The value of the normalized intensity is printed next to the "drawn" intensity (line 700). This "plotting" algorithm was described by J. E. Evans of Finnigan Corporation.

PROGRAM COR7

The occurrence of mixtures complicates the interpretation of mass spectra since ions present in a scan may be generated from several compounds. The program COR7 assists in deciding whether this may be the case. The program retrieves the intensities of up to five ions from a maximum of 20 consecutive scans and calculates their correlation coefficients. If the correlation coefficients are ≥ 0.98, there is a good chance that the ions were generated from the same compound. However, the possibility that several compounds are eluted from the gas chromatograph exactly together cannot be excluded.

The scan range examined should be rather narrow, usually about 5-10 scans.

The program can also print the intensities of the ions or plot them for a visual pattern examination. It is often helpful to see the intensity values and their pattern, in addition to the correlation coefficients.

The sign on procedure (lines 1-9) is standard. Array A (20,5) is reserved for the ion intensities; arrays X(20) and Y(20) are used in the correlation coefficient subroutine; and array V(5) is for storage of the selected ion masses.

The masses are entered in the I loop (lines 110-160), followed by the scan range. If the entered range is more than 20 scans, the program asks for a new range (line 195). Spectra are read from the disc and selected intensities are placed into array A by the X loop (lines 200-250).

The printout of intensities is optional (lines 297-360), and the program proceeds to calculate the correlation coefficients. The subroutine in lines 600-700 is used. Lines 440-530 select, in turn, all pairs of ion intensities and place them into the X and Y arrays, respectively, for the calculation.

After completing the calculations, several self-explanatory options are offered by line 550.

The plotting subroutine (lines 1000-1310) is the same as in the program RGEVA4. The intensities of one ion at a time are plotted with the option to terminate the plotting routine after each ion.

PROGRAM RGEVA4

```
1 DIM M(1000),G(1000)
2 PRINT "ENTER 30-49 DATA THE FOLLOWING INFORMATION:RGCMASS,"
3 PRINT "UP TO 3 MASS RANGES(BALANCE 0'S),5 MAX.&BACKGR.(+0'S)"
4 PRINT "DATA MATRIX IS (20,17);ROWS=RGC'S,COLUMNS AS ABOVE"
5 PRINT "10 FILES----GOTO 10"
6 STOP
10 FILES 1T1402
13 DIM D(20,17),F$(6)
15 PRINT "NUMBER OF ROWS"
16 INPUT I1
30 DATA 419,415,430,0,0,0,0,151,148,263,260,269,265,0,0,0,0
31 DATA 405,400,415,0,0,0,0,212,207,220,217,0,0,0,0,0,0
32 DATA 457,450,465,0,0,0,0,268,266,292,289,276,275,0,0,0,0
33 DATA 306,300,318,0,0,0,0,253,251,271,269,276,275,0,0,0,0
34 DATA 403,400,415,0,0,0,0,274,271,278,276,280,276,0,0,0,0
35 DATA 471,469,480,0,0,0,0,303,302,366,365,0,0,0,0,0,0
36 DATA 318,316,326,0,0,0,0,304,302,0,0,0,0,0,0,0,0
37 DATA 366,362,372,0,0,0,0,344,341,0,0,0,0,0,0,0,0
38 DATA 492,490,506,0,0,0,0,191,188,291,289,351,349,0,0,0,0
43 DATA 262,260,275,0,0,0,0,230,227,300,298,310,307,315,312,0,0
44 DATA 409,400,415,0,0,0,0,256,251,279,275,0,0,0,0,0,0
45 DATA 385,380,395,0,0,0,0,132,130,174,171,210,207,228,226
46 DATA 239,236
47 DATA 376,372,388,0,0,0,0,239,238,253,250,290,288,300,298,0,0
48 DATA 408,406,420,0,0,0,0,252,250,279,276,304,301,315,312
49 DATA 321,318,419,415,430,0,0,0,0,151,148,263,260,269,265,0,0,0,0
50 FOR I=1 TO I1
52 FOR J=1 TO 17
53 READ D(I,J)
55 NEXT J
57 NEXT I
60 PRINT "FILE"
70 INPUT F$
85 PRINT
86 PRINT "XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX"
90 PRINT TAB(30),"FILE",F$
100 PRINT "XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX"
110 FOR I=1 TO 10
120 IF D(I,1)=0 THEN 310
130 PRINT TAB(20),"RGC";D(I,1)
140 PRINT TAB(19),"----------"
150 FOR J=8 TO 16 STEP 2
152 MAT M=ZER
154 MAT G=ZER
160 IF D(I,J)=0 THEN 300
170 LET Y=D(I,J)
180 DREAD #1,Y,M
190 LET Y=D(I,J+1)
200 DREAD #1,Y,G
210 MAT M=M-G
220 PRINT TAB(10),"MAX";D(I,J),"BACKGR.";D(I,J+1)
225 PRINT TAB(10),"################################"
230 FOR K=2 TO 6 STEP 2
240 IF D(I,K)=0 THEN 290
250 LET X=D(I,K)
260 LET X1=D(I,K+1)

*
```

PROGRAM RGEVA4 (cont'd)

```
270 GOSUB 500
275 PRINT TAB(30),". . . . . . . . . . . . "
280 NEXT K
285 PRINT TAB(30),"......................."
290 NEXT J
295 PRINT TAB(30),"+++++++++++++++++++++++"
300 NEXT I
310 END
500 LET A=B=0
510 FOR N=X TO X1
520 LET A=M(N)
530 IF A<B THEN560
540 LET B=A
550 LET Z=N
560 NEXT N
570 PRINT "BASE PEAK IS";Z
580 FOR N=X TOX1
590 LET Q=M(N)/B
600 LET Y=100*Q
610 LET Y=INT(Y)
620 LET Q=Q*60
630 LET Q=INT(Q)
640 PRINT
650 PRINT N;
660 FOR L=1 TO Q
670 PRINT "-";
680 NEXT L
690 IF Q<=0 THEN 710
700 PRINT TAB(Q+6),Y;
710 NEXT N
715 PRINT
720 RETURN

*
```

PROGRAM COR7

```
1 DIM M(1000),G(1000)
2 PRINT "PROGRAM COR7"
3 PRINT "PROGRAM CORRELATES THE OCCURRENCE OF IONS"
4 PRINT "SPACE AVAILABLE FOR 5 IONS AND 20 SCANS"
5 PRINT "10 FILES ...../GOTO 10"
9 STOP
10 FILES LSWA2
20 DIM A(20,5),X(20),Y(20)
30 DIM V(5)
100 LET N=0
105 PRINT "ENTER IONS, TERMINATE BY 0"
110 FOR I=1 TO 5
115 PRINT "ION";I
130 INPUT V(I)
140 IF V(I)=0 THEN 170
150 LET N=N+1
160 NEXT I
170 PRINT "SCAN RANGE,MAXIMUM 20"
180 INPUT K,L
185 LET K1=K
190 LET N1=L-K+1
195 IF N1>20 THEN 170
196 LET C=0
200 FOR X=K TO K+N1
210 MAT M=ZER
220 DREAD #1,X,M
221 LET C=C+1
225 FOR J=1 TO N
226 LET V=V(J)
230 LET A(C,J)=M(V)
240 NEXT J
250 NEXT X
251 PRINT "PRINTOUT: YES(1),NO(2)"
252 INPUT Q1
260 FOR I=1 TO N
270 PRINT V(I),
280 NEXT I
290 PRINT
295 PRINT "-------------------------------------------------------------"
297 IF Q1=2 THEN 440
300 FOR I=1 TO N1
310 FOR J=1 TO N
320 PRINT A(I,J),
330 NEXT J
340 PRINT
350 NEXT I
360 PRINT "-------------------------------------------------"
440 FOR J=1 TO N-1
445 FOR K=J+1 TO N
450 FOR I=1 TO N1
460 LET X(I)=A(I,J)
470 LET Y(I)=A(I,K)
480 NEXT I
490 GOSUB 600
500 PRINT R,
510 NEXT K
520 PRINT

*
```

PROGRAM COR7 (cont'd)

```
530 NEXT J
540 PRINT "-------------------------------------------------"
550 PRINT"ANOTHER SCAN RANGE(1),SET OF IONS(2),PLOT(3),END(4)"
551 INPUT Q2
552 IF Q2=1 THEN170
553 IF Q2=2 THEN100
554 IF Q2=3 THEN 1000
560 END
600 LET Z1=Z2=Z3=Z4=Z5=0
610 FOR I=1 TO N1
620 LET Z1=Z1+X(I)
630 LET Z2=Z2+X(I)↑2
640 LET Z3=Z3+Y(I)
650 LET Z4=Z4+Y(I)↑2
660 LET Z5=Z5+X(I)*Y(I)
670 NEXT I
675 LET T1=SQR(N1*Z2-Z1↑2)
676 LET T2=SQR(N1*Z4-Z3↑2)
680 LET R=N1*Z5-Z1*Z3
681 IF T1<=0 THEN 710
682 IF T2<=0 THEN 710
690 LET R=R/(T1*T2)
695 LET R=INT(1000*R)/1000
700 RETURN
710 LET R=999
720 GOTO 700
1000 PRINT "ION NUMBER (TERMINATE BY 0"
1005 INPUT J
1010 IF J=0 THEN 550
1020 LET A=B=0
1030 FOR I=1 TO N1
1040 IF A(I,J)=0 THEN 1500
1050 LET A=A(I,J)
1060 IF A<B THEN 1100
1070 LET B=A
1080 LET Z=I
1100 NEXT I
1120 PRINT "ION";V(J)
1130 PRINT "####################"
1135 LET K2=K1
1140 FOR I=1 TO N1
1150 LET Q=A(I,J)/B
1160 LET Y1=100*Q
1170 LET Y1=INT(Y1)
1180 LET Q=60*Q
1190 LET Q=INT(Q)
1210 PRINT K2+I-1;
1220 FOR L=4 TO Q
1230 PRINT "-";
1240 NEXT L
1245 PRINT
1250 NEXT I
1310 GOTO 1000

*
```

PROGRAM QUANT7

This program detects peaks in limited mass reconstructed gas chromatograms and determines their areas. The program performs well on chromatograms with sharp, clearly separated peaks. It may fail on chromatograms containing poorly resolved peaks and drifting baseline.

The sign on routine is standard (lines 1-9). The array L(1000) is used in the background correction of spectra (lines 2030, 2040), since the array G(1000) contains the reconstructed gas chromatogram. The ions for quantitation are entered as 20-30 DATA (actually an additional 9 lines are available), and the input is terminated by "0."

The limited mass reconstructed gas chromatogram is obtained in line 200. M1 is the current mass; reconstruction of the chromatogram begins in scan "1," and the chromatogram is placed into the G(1000) array. The chromatogram is examined from scan 1 to scan 250 (lines 210, 220). This range should be adjusted as required.

The peak detecting algorithm is the I5 loop (lines 500-700). The first condition for a peak is that the next two scans have higher intensities (lines 510, 520). If these conditions are true, the current I5 becomes the background for the peak (P0), and the algorithm examines up to 10 following scans (loop I2, lines 540-580) to detect the maximum. The maximum (P1) is found, however, if the I2 + 1 scan is less intense than the I2 + 2 scan; the following I2 loop (lines 590-630) is shifted by one scan (J1) to avoid premature termination of the peak to be detected in this loop. End of the peak (P3) is found (line 610).

The peak width must be more than three scans (line 632). This condition tends to minimize the effect of noisy baseline.

If a peak has been detected, its area is determined by the subroutine in lines 820-890. In a somewhat unusual programming procedure, I5 is then modified so that the examination of the reconstructed gas chromatogram resumes only after the end of the currently detected peak (line 650).

Baseline correction is determined by linear interpolation (lines 820-830), intensities within the peak are summed up (lines 850-870), and the background area is subtracted (line 880). The statements 885 and 886 eliminate peaks due to a noisy baseline. Parameters of the peak: scan number of maximum (P1) and its boundaries (P0 and P3), and their respective intensities (G(P1), G(P0), and G(P3)) are printed (lines 887, 888), followed by the area B4, background area B2, and baseline slope B1 (line 890). These values help to judge the quality of the peak and the accuracy of area determination.

To confirm the identification of the quantitated compound, the background-corrected mass spectrum is retrieved (lines 2000-2040), normalized (lines 2050-2110), and printed (lines 2120-2180), if the intensity of the base ion is at least 300 absolute units. The mass range for the printout is 70-450 (line 2060) and only peaks with intensity ≥10% are printed. These parameters should be adjusted as required.

PROGRAM ELANAL

This program, developed by Kavanagh (1980), derives all empirical formulae for a given ion, calculates isotope abundancies, and compares these to the observed values. Listing of the program, obtained from the author, was translated into Tektronix Plot 50 BASIC. The formula computing algorithm was slightly modified and the isotope abundancies algorithm has not been changed. Program output has been directed to magnetic tape and several minor modifications have been implemented. A detailed documentation of the new version is given below.

N$(60) in line 10 is the name of the file on tape for recording the output. The file is found and the title is recorded (lines 16-20).

The arrays dimensioned in lines 50-60 are used as follows: N(10) contains maximum possible number of atoms of elements in a formula of given mass. L(10) contains the list of elements under consideration (1-iodine, 2-bromine, 3-chlorine, 4-sulfur, 5-phosphorus, 6-silicon, 7-fluorine, 8-oxygen, 9-nitrogen, and 10-carbon). Input of "0" or "1" into L(I) means that the I-th element is or is not considered, respectively. Carbon and hydrogen are considered automatically. The array Z(51) is expanded compared to the original version and contains factorials from 0! to 50!. T(20) contains atomic weights of the lightest stable isotopes of the 10 elements potentially under consideration, and "apparent" atomic weights used in the formula algorithm. O(10) contains numbers of atoms of individual elements in a given formula. A(18) contains calculated isotope abundancies, B(18) final isotope abundancies, and F(18) is an array of normalized intensities. C(10,3) describes a given isotopic formula, rows represent elements, columns represent isotopes. S(10,3) contains natural stable isotope abundancies of the 10 elements. Masses of ions investigated are stored in W(10). The array X(50) contains "procedural instructions" stored as follows: positions 1-10 contain either "0" or "1" for analysis (calculations of formulae, isotopic abundancies, and comparison with observed abundancies) or formulae calculation only, respectively; positions 11-20 contain "0" or "1" for molecular or even electron ion, respectively; positions 21-30 contain numbers of ions; positions 31-40 abundancies of "A-1" ions; and positions 41-50 are used to store the required "goodness of fit." Y(10,15) is used to store observed isotope abundancies for 10 ions (rows) and up to 15 masses. G(10) contains numbers of formulae found.

The arrays are initialized (lines 61-76) and elements are chosen (lines 115-210). The choice is recorded on tape (line 212). Arrays T, Z, and S are filled in lines 230-275.

A parameter input loop K (lines 280-390) is used to enter nominal masses of up to 10 ions, procedural instructions (molecular or even electron ion, comparison of observed and calculated isotope abundancies or calculation of formula only, number of nominal mass peaks, abundancies of A-1 ions, and "goodness of fit"), and observed isotope abundancies.

PROGRAM QUANT7

```
1 DIM M(1000),G(1000)
2 DIM L(1000)
3 PRINT "PROGRAM QUANT7"
4 PRINT "INPUT MZ'S AS 20-30 DATA;TERMINATE BY 0"
8 PRINT "10 FILES/GOTO 10"
9 STOP
10 FILES LSWA1
15 MAT G=ZER
20 DATA 149,85,105,0
40 READ M1
41 PRINT "XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX"
42 IF M1=0 THEN 750
43 PRINT TAB(30),"RGCLM";M1
45 PRINT TAB(30),"#########"
200 RGCLM #1,M1,1,G
210 LET S1=1
220 LET S2=250
500 FOR I5=S1 TO S2
510 IF G(I5)>=G(I5+1) THEN 700
520 IF G(I5)>=G(I5+2) THEN 700
530 LET P0=I5
535 LET J1=0
540 FOR I2=I5+1 TO I5 +10
550 IF G(I2)<G(I2+1) THEN 580
560 LET P1=I2
570 IF G(I2+1)>G(I2+2) THEN 590
575 LET J1=1
577 GOTO 590
580 NEXT I2
590 FOR I2=P1+J1+1 TO P1+J1+20
600 IF G(I2)>G(I2+1) THEN 630
610 LET P3=I2
620 GOTO 632
630 NEXT I2
632 IF P3-P0<=3 THEN 700
640 GOSUB 820
650 LET I5=I5+P3-P0-1
700 NEXT I5
710 GOTO 40
750 END
820 LET B1=(G(P3)-G(P0))/(P3-P0)
830 LET B2=(G(P0)+G(P3))*(P3-P0-1)/2
840 LET B3=0
850 FOR I1=P0+1 TO P3-1
860 LET B3=B3+G(I1)
870 NEXT I1
880 LET B4=B3-B2
885 IF B4<.1*B2 THEN 2190
886 IF B4<200 THEN 2190
887 PRINT "MAX";P1;"(";G(P1);")";"BACK";P0;"(";G(P0);")";
888 PRINT "END";P3;"(";G(P3);")"
889 PRINT "-------------"
890 PRINT "AREA=";B4,"BACKGR.=";B2,"BACKGR.SLOPE=";B1
2000 MAT M=ZER
2010 MAT L=ZER
2015 LET X=P1
2016LET Y=P0
2020 DREAD #1,X,M
2030 DREAD #1,Y,L
2040 MAT M=M-L
2050 LET C=D=J=0
2060 FOR K=70 TO 450
2070 LET C=M(K)
2080 IF C<D THEN 2110
2090 LET D=C
2095 LET J=K
2110 NEXT K
2120 PRINT TAB (10),"BASE M/Z";J,"(";D;")"
2125 IF D<300 THEN 2190
2126 LET V=INT(100*M(M1)/D)
2127 IF V<20 THEN 2190
2130 FOR K=70 TO 450
2140 LET V=INT(100*M(K)/D)
2150 IF V<10 THEN 2170
2160 PRINT K;V,
2170 NEXT K
2180 PRINT
2190 RETURN
```

*

The entered values are recorded on tape (lines 391-393).

The Finnigan BASIC function MOD(X,Y), X modulo Y is translated as user-defined function FNM(X), line 395.

The main loop (loop I1, lines 405-1800) carries out the complete calculation for up to 10 ions.

Although all results are written on tape, some are displayed on the screen to give an indication of the status of the calculations. The first such message is in line 425.

Line 426 causes the system to PAGE (clear screen) when full and continue (otherwise the program execution would be halted).

Lines 430-445 establish addresses for procedural instructions for current ion: M2 for molecular or even electron ion; M3 for number of isotope formulae; M4 for A-1 abundance; and M5 for goodness of fit.

Lines 450-460 obtain the values stored at these positions in the array X(50); A2 analysis or formula; A1 molecular or even electron ion; and A4 number of isotope formulae; P3 (line 465) counts the formulae found; and P9 is the current goodness of fit.

If formula only is required, according to line 480, the abundance normalization loop (lines 485-495) and normalization of the A-1 ion in line 500 are bypassed. It should be noted that the abundancies are normalized to the A ion (not to the most abundant ion). The other normalization is more convenient for display purposes and is used in the program ISOTOPE. The values of the goodness of fit are different for these two modes of normalization.

W9 (line 515) becomes the mass of the current ion. Maximum numbers of atoms of elements considered in a formula are obtained by dividing given mass by atomic weight (line 535). This operation is executed for all elements by the loop in lines 525-550.

Maximum number of rings and double bonds (R or D2 in the program) is calculated in line 650. If a formula of weight M contains only C and H, the maximum value of R is INT(M/12) + 1 (C_n can form a maximum of $\underline{n}$ double bonds and for this the atoms must form a ring, consequently R = n + 1).

The second major loop in the program is the "Rings and Double Bonds" loop D1 (lines 680-1720). The loop starts with paraffins and derivatives (D1 = 0) and ends with the maximum possible number of rings and double bonds, D2.

The number of rings and double bonds is the "hyrdogen deficit" relative to the corresponding paraffin (or its derivative). The formula of paraffins is C_nH_{2n+2}. Consequently, for a formula $C_{nc}H_{nh}$ the hydrogen deficit or rings and double bonds is given by (1).

$$R = nc - 0.5nh + 1 \qquad (1)$$

For a paraffin, R = 0 (substitute nh = 2nc + 2 in (1)). Values R < 0 are not possible (the hydrocarbon would be "oversaturated"). In the mode of bonding, halogens are equivalent to hydrogen (one bond) and (1) can be expanded to (2).

$$R = nc - 0.5nh - 0.5\,ncl - 0.5nbr - 0.5ni - 0.5nf + 1 \qquad (2)$$

Oxygen bound to carbon by a single bond does not affect R. Double-bonded oxygen replaces two hydrogens which is equivalent to one double bond in (1) and (2). Sulfur behaves in the same way. On the other hand, silicon forms four bonds and is analogous to carbon.

In contrast, nitrogen normally forms three bonds. Thus, for the formula $C_{nc}H_{nh}N_{nn}$, the term 0.5nn in (3) accounts for the "extra" hydrogen due to the presence of nitrogen.

$$R = nc - 0.5nh + 0.5\,nn + 1 \qquad (3)$$

Phosphorus forms five bonds and, consequently, provides three "extra" hydrogens. As a result, the formula including all the discussed elements is:

$$R = nc - 0.5(nh + ncl + nbr + ni + nf) + nsi + 0.5nn + 1.5np + 1 \qquad (4)$$

Mass spectrometry deals with ions. Positive ions may be formed by the loss of one electron (molecular or odd electron ions). R calculated from (4) will be an integer for these ions. Ions formed by the loss of one bond (even electron ions), which is equivalent to a "hydrogen deficit" of 1, will not have integer R values (the fraction of R will always be 0.5). To accommodate even-electron ions, a term "j" is subtracted from the right-hand side of (4). Its value is "0" for odd- and "0.5" for even-electron ions.

ELANAL calculates formulae from equation (4) and from equation (5) for the mass of the ion, W9.

$$W9 = nh + 35ncl + 79nbr + 127ni + 19nf + 12nc + 28nsi + 14nn + 31np + 16no + 32ns \qquad (5)$$

By rearranging (4)

$$nh + ncl + nbr + ni + nf - 2nc - 2nsi - nn - 3np = 2 - 2R - j \qquad (6)$$

After subtracting (6) from (5)

$$W9 - 2 + 2R + j = 34ncl + 78nbr + 126ni + 18nf + 14nc + 30nsi + 15nn + 34np + 16no + 32ns \qquad (7)$$

Value of the left-hand side of (7) is calculated in line 700 (W1). The variable A1 is "J" and is determined by the procedural instructions. The "apparent" masses of the right-hand side of (7) are stored in T(11)-T(20) in decreasing order of magnitude.

Equation (7) is solved first for carbon and nitrogen. In lines 710-905, elements not considered are bypassed, numbers of those considered are set to "0" in the first pass through the loop (U1,U2,...U8 = 0) and are stored in the array O(10).

At this stage, the "unallocated" mass K5 equals W1 (line 910), which is the left-hand side of (7). If all mass had been allocated (K5 = 0), the formula is not possible since it does not contain carbon, and the program proceeds to calculate another

formula (line 920). The same situation arises when K5 ≤ 14 (line 925).

The innermost loop in the empirical formulae calculation is the U9 loop. The loop operates from U9 = 0 to the maximum numbers of nitrogen atoms, currently possible (unallocated mass/15, line 940). The loop is solving the equation

$$W9 = 14nc + 15nn \qquad (8)$$

for interger values of nc and nn. If (W9-15nn)/14 is an integer (950 FNM = 0), then nn = U9 = 0(9) and nc = 0(10). If 950 is not true, another value of U9 is tried.

If nitrogen is not to be considered (L(9) = 1, line 935), then the unallocated mass must be divisible by 14 (line 970). If it is not, another formula is sought.

Line 995 calculates the number of hydrogens (W2 = nh) from equation (6); lines 1000 and 1005 calculate the formula mass. W3 in line 1010 is the difference between given and calculated formula mass and is 0 for the correct formula.

At this stage, additional constraints may be introduced in lines 1021-1024. For example, in current applications of ELANAL, the number of oxygens or nitrogens is not to exceed the number of carbons (line 1022), and the sum of oxygens and nitrogens is not to exceed the number of carbons (line 1023).

If all constraints are satisfied, the formula is counted (line 1025), displayed (line 1026) and, if required (A2 = 0, line 1030), the isotope abundancies are calculated. The abundancies calculations (lines 1070-1590) are discussed in connection with the program ISOTOPE.

Depending on the specified goodness of fit (line 1535), the formula may be displayed (lines 1550-1557) and recorded on tape (line 1563). If formulae only were calculated, the results are written on the tape in a different format (line 1622).

Lines 1625-1700 close the formulae loops. Line 1705 initializes the 0(10) array and the "Ring and double bond" loop is closed in line 1720. The number of formulae is stored in the array G(10) and the ion loop is closed in line 1800. Lines 1810, 1815 conclude the output operations.

The subroutine in lines 1825-1870 revises down the maximum number of atoms of elements, depending on allocated mass. J1 is the current element, K1 is the next element, and K2 is the allocated mass.

Examples

For a fragment ion m/z 235 and isotopic abundance 234 0, 235 100, 236 16, 237 68, 238 10, and 239 12, ELANAL examines 17 formulae (Br and Cl considered). Two formulae have goodness of fit (SUSQ) < 20: $C_{12}H_{21}Cl_2$ (13) and $C_{13}H_9Cl_2$ (7). The second formula is correct (isotope abundancies of the m/z 235 ion of DDT from the data base MSSS were used).

For a nonachloroterphenyl molecular ion $C_{18}H_5Cl_9$, m/z 536 (Wright et al. 1978), and isotope abundancies read off Fig. 8 in the original paper, ELANAL finds the following best fitting (SUSQ < 10000) formulae out of a total of 135 examined:

Formula C	H	Br	Cl	Rings and double bonds	SUSQ
18	31	1	6	0	4420
19	28	0	8	2	8900
17	17	0	9	5	5340
19	19	1	6	7	4700
20	16	0	8	9	8880
18	5	0	9	12	6050
20	7	1	6	14	5100
21	4	0	8	16	8900

For a nonachlor ion $C_{10}H_5Cl_8$, m/z 405 from a sample of tuna liver (Zitko 1980), ELANAL examines 64 formulae. The 3 best fitting (SUSQ) are $C_{10}H_5Cl_8$ (15000), $C_9H_8BrCl_6$ (17000), and $C_{11}H_{10}Br_2Cl_3$ (18200).

Ions in the spectrum of a chlordane component, postulated to be a monochlorinated adduct of tetrachlorocyclopentadiene and cyclopentadiene, $C_{10}H_7Cl_5$ (Sovocool et al. 1977) were analyzed by ELANAL. The isotope abundancies were obtained from Fig. 3 of the original paper. Br, Cl, 0 were considered.

For the low intensity parent ion, 27 out of 144 formulae had SUSQ < 10000. The "correct" formula had SUSQ 7900. Best fitting formulae were $C_6H_4OCl_6$ (1330) and $C_7H_8Cl_6$ (1450). For the (M-Cl) ion, m/z 267, presumably $C_{10}H_7Cl_4$, 49 out of 111 formulae had SUSQ < 10000. The correct formula had SUSQ 1100. Eight other formulae had a better fit, $C_4H_6O_4BrCl_2$ giving the best fit at SUSQ 680. The corresponding parent ion $C_4H_6O_4BrCl_3$ had SUSQ of 3000. On the other hand, the (M-Cl) ion corresponding to the best fit for the M ion, $C_7H_8Cl_5$ had SUSQ 1600. The base cluster at m/z 202, assumed to be formed by retro Diels-Alder fragmentation, $C_5H_2Cl_4$ had SUSQ 41. This was the best fit out of 49 formulae examined. An additional 4 formulae had SUSQ of about 1100.

PROGRAM TAPE

This program reads from tape and prints files written by ELANAL. The arrays (lines 100, 101) are the same as in ELANAL. The file is selected (lines 105-115), the system is set for output on a teletype (line 120), file name is read from tape (line 125) and displayed. Goodness of fit is selected (line 128). For formulae exceeding this level, formulae only will be printed.

In lines 130-455, the "header" information of the file, including elements considered, ion masses, and operational instructions, is read and printed. Record reading starts in line 462. I1 = 100000 indicates that the next record is a summary table at the end of the file (lines 655-690).

Depending on the type of calculation specified in ELANAL (formula only or isotope abundance calculation), X(I1) is either "0" or "1" and the respective records (line 465 or 468) are read. I1 is ion number, D1 is the number of rings and double bonds, 0 contains number of atoms of elements in the formula, W2 is the number of hydrogen atoms, D3 is the intensity of the (A-1) peak, B is the calculated

PROGRAM ELANAL

```
10 DIM N$(60)
15 PRINT "DATA FILE NUMBER"
16 INPUT G9
17 FIND G9
18 PRINT "FILE NAME"
19 INPUT N$
20 WRITE @33:N$
50 DIM N(10),L(10),Z(51),T(20),O(10),A(18),B(18),F(18)
55 DIM C(10,3),S(10,3),W(10)
60 DIM X(50),Y(10,15),G(10)
61 N=0
62 L=0
63 Z=0
64 T=0
65 V=0
66 O=0
67 A=0
68 B=0
69 F=0
71 C=0
72 S=0
73 W=0
74 X=0
75 Y=0
76 G=0
105 PRINT "WHEN THE ELEMENTAL SYMBOL APPEARS INPUT 0 IF YOU WAIT"
110 PRINT "THAT ELEMENT CONSIDERED, OTHERWISE INPUT 1"
115 PRINT "I"
120 INPUT L(1)
125 PRINT "BR"
130 INPUT L(2)
135 PRINT "CL"
140 INPUT L(3)
145 PRINT "S"
150 INPUT L(4)
155 PRINT "P"
160 INPUT L(5)
165 PRINT "SI"
170 INPUT L(6)
175 PRINT "F"
180 INPUT L(7)
185 PRINT "O"
190 INPUT L(8)
195 PRINT "N"
200 INPUT L(9)
210 L(10)=0
212 WRITE @33:L
215 PRINT "H&C ARE CONSIDERED AUTOMATICALLY"
220 PRINT "INPUT THE NUMBER OF IONS TO BE ANALYZED(MAY 10)"
225 INPUT M1
230 DATA 127,79,35,32,31,28,19,16,14,12,126,78,34,32,34,33,18,16,15,14
235 READ T
240 DATA 1,1,2,6,24,120,720,5040,40320,362880,3628800,3.99168E+7
250 DATA 4.79E+8,6.227E+9,8.718E+10,1.3077E+12,2.092E+13
251 DATA 3.557E+14,6.402E+15,1.21645E+17,2.4329E+18,5.1091E+19
252 DATA 1.124E+21,2.5852E+22,6.2045E+23,1.55112E+25,4.03292E+26
253 DATA 1.08889E+28,3.04888E+29,8.84176E+30,2.65253E+32
254 DATA 8.22283E+33,2.63131E+35,8.68332E+36,2.95233E+38
255 DATA 1.0333IE+40,3.71993E+41,1.37637E+43,5.23023E+44
256 DATA 2.03988E+46,8.15915E+47,3.345253E+49,1.40501E+51
257 DATA 6.04513E+52,2.65825E+54,1.1962E+56,5.50262E+57
258 DATA 2.58623E+59,1.24139E+61,6.08282E+62,3.04141E+64
259 READ Z
260 DATA 1,0,0,0.5054,0.4946,0,0.7547,0.2453,0,0.95,0.0076,0.0422
265 DATA 1,0,0,0.9221,0.047,0.0309,1,0,0,0.9976,3.994E-4,0.00199
270 DATA 0.9963,0.0037,0,0.9889,0.0111,0
275 READ S
280 FOR K=1 TO M1
285 M2=K+10
290 M3=K+20
295 PRINT "FOR ION NUMBER";K;"*******"
300 PRINT "INPUT MOLECULAR WEIGHT"
305 INPUT W(K)
```

PROGRAM ELANAL (cont'd)

```
310 PRINT "INPUT 0 OR 1 FOR MOLECULAR OR EVEN ELECTRON ION"
315 INPUT X(M2)
320 PRINT "INPUT 0 OR 1 FOR ANALYSIS OR FORMULA"
325 INPUT X(K)
330 IF X(K)=1 THEN 390
335 PRINT "INPUT NUMBER OF PEAKS"
340 INPUT X(M3)
345 PRINT "INPUT RELATIVE INTENSITIES STARTING WITH A-1"
350 M4=K+30
355 M5=K+40
360 INPUT X(M4)
365 FOR I=1 TO X(M3)-1
370 INPUT Y(K,I)
375 NEXT I
380 PRINT "INPUT SUSQ"
385 INPUT X(M5)
390 NEXT K
391 WRITE @33:W
392 WRITE @33:X
393 WRITE @33:Y
395 DEF FNM(X)=X-14*INT(X/14)
405 FOR I1=1 TO M1
425 PRINT "ION NUMBER ";I1,"MASS ",W(I1)
426 PRINT @32,26:2
430 M2=I1+10
435 M3=I1+20
440 M4=I1+30
445 M5=I1+40
450 A2=X(I1)
455 A1=X(M2)
460 A4=Y(M3)
465 P3=0
470 P9=Y(M5)
480 IF A2=1 THEN 515
485 FOR I=1 TO A4-1
490 F(I)=Y(I1,I)*100/Y(I1,1)
495 NEXT I
500 D3=X(M4)*100/Y(I1,1)
515 W9=W(I1)
520 REM ESTABLISH MAXIMUM LIMITS FOR ELEMENT NUMBERS
525 FOR I=1 TO 10
530 IF L(I)=1 THEN 545
535 N(I)=INT(W9/T(I))
540 GO TO 550
545 N(I)=0
550 NEXT I
555 REM MAXIMUM LIMITS FOR RINGS & DOUBLE BONDS
650 D2=1+N(10)
680 REM RINGS&DOUBLE BONDS LOOP
685 FOR D1=0 TO D2
690 PRINT "R&DB= ";D1
700 W1=W9-2+2*D1+A1
705 REM SOLUTION OF EQUATIONS
710 IF L(1)=1 THEN 735
715 FOR U1=0 TO N(1)
720 J1=1
725 O(1)=U1
730 GOSUB 1825
735 IF L(2)=1 THEN 760
740 FOR U2=0 TO N(2)
745 J1=2
750 O(2)=U2
755 GOSUB 1825
760 IF L(3)=1 THEN 785
765 FOR U3=0 TO N(3)
770 J1=3
775 O(3)=U3
780 GOSUB 1825
785 IF L(4)=1 THEN 810
790 FOR U4=0 TO N(4)
795 J1=4
```

PROGRAM ELANAL (cont'd)

```
800 O(4)=U4
805 GOSUB 1825
810 IF L(5)=1 THEN 835
815 FOR U5=0 TO N(5)
820 J1=5
825 O(5)=U5
830 GOSUB 1825
835 IF L(6)=1 THEN 860
840 FOR U6=0 TO N(6)
845 J1=6
850 O(6)=U6
855 GOSUB 1825
860 IF L(7)=1 THEN 885
865 FOR U7=0 TO N(7)
870 J1=7
875 O(7)=U7
880 GOSUB 1825
885 IF L(8)=1 THEN 910
890 FOR U8=0 TO N(8)
895 J1=8
900 O(8)=U8
905 GOSUB 1825
910 K5=W1-K2
915 O(9)=0
916 O(10)=0
920 IF K5=0 THEN 1625
925 IF K5<14 THEN 1625
935 IF L(9)=1 THEN 970
940 FOR U9=0 TO INT(K5/15)
950 IF FNM(K5-15*U9)<>0 THEN 1620
955 O(9)=U9
960 O(10)=(K5-15*U9)/14
965 GO TO 995
970 IF FNM(K5)<>0 THEN 1625
975 O(10)=K5/14
995 W2=2*O(10)+O(9)+2*O(6)+3*O(5)+2-2*D1-O(1)-O(2)-O(3)-O(7)-A1
1000 K8=O(1)*T(1)+O(2)*T(2)+O(3)*T(3)+O(4)*T(4)+O(5)*T(5)+O(6)*T(6)
1005 K9=K8+O(7)*T(7)+O(8)*T(8)+O(9)*T(9)+O(10)*T(10)+W2
1010 W3=W9-K9
1015 IF W3<>0 THEN 1625
1020 IF W2<0 THEN 1625
1022 IF O(10)<O(9) OR O(10)<O(8) THEN 1625
1023 IF O(10)<O(8)+O(9) THEN 1625
1025 P3=P3+1
1026 PRINT "ION ";I1;" R&DB ";D1;" MAX ";D2;"FORMULA ";P3,O;
1030 IF A2=1 THEN 1550
1040 A=1
1045 B=0
1055 REM CALCULATIONS OF INTENSITIES
1070 C=0
1085 IF O(2)=0 THEN 1105
1090 FOR N1=0 TO O(2)
1095 C(2,2)=N1
1100 C(2,1)=O(2)-N1
1105 IF O(3)=0 THEN 1125
1110 FOR N2=0 TO O(3)
1115 C(3,2)=N2
1120 C(3,1)=O(3)-N2
1125 IF O(4)=0 THEN 1170
1130 FOR N3=0 TO O(4)
1135 C(4,2)=N3
1140 C(4,1)=O(4)-N3
1145 FOR N4=0 TO O(4)
1150 C(4,3)=N4
1155 C(4,2)=N3
1160 C(4,1)=O(4)-N3-N4
1165 IF C(4,1)<0 THEN 1425
1170 IF O(6)=0 THEN 1215
1175 FOR N5=0 TO O(6)
1180 C(6,2)=N5
1185 C(6,1)=O(6)-N5
1190 FOR N6=0 TO O(6)
1195 C(6,3)=N6
```

PROGRAM ELANAL (cont'd)

```
1200 C(6,2)=N5
1205 C(6,1)=O(6)-N6-N5
1210 IF C(6,1)<0 THEN 1410
1215 IF O(8)=0 THEN 1260
1220 FOR N7=0 TO O(8)
1225 C(8,2)=N7
1230 C(8,1)=O(8)-N7
1235 FOR N8=0 TO O(8)
1240 C(8,3)=N8
1245 C(8,2)=N7
1250 C(8,1)=O(8)-N7-N8
1255 IF C(8,1)<0 THEN 1395
1260 IF O(9)=0 THEN 1280
1265 FOR N9=0 TO O(9)
1270 C(9,2)=N9
1275 C(9,1)=O(9)-N9
1280 IF O(10)=0 THEN 1300
1285 FOR P1=0 TO O(10)
1290 C(10,2)=P1
1295 C(10,1)=O(10)-P1
1300 REM CALCULATION OF PEAK NUMBER
1302 A5=C(10,2)+C(9,2)+C(8,2)+C(6,2)+C(4,2)+2*C(3,2)+2*C(2,2)
1305 A5=A5+2*C(8,3)+2*C(6,3)+2*C(4,3)+1
1314 IF A5>A4 THEN 1365
1315 FOR I=1 TO 10
1320 IF O(I)=0 THEN 1350
1325 FOR J=1 TO 3
1330 IF S(I,J)=0 THEN 1340
1335 A(A5)=A(A5)*S(I,J)↑C(I,J)/Z(C(I,J)+1)
1340 NEXT J
1345 A(A5)=A(A5)*Z(O(I)+1)
1350 NEXT I
1355 B(A5)=B(A5)+A(A5)
1360 A(A5)=1
1365 IF O(10)=0 THEN 1375
1370 NEXT P1
1375 IF O(9)=0 THEN 1385
1380 NEXT N9
1385 IF O(8)=0 THEN 1400
1390 NEXT N8
1395 NEXT N7
1400 IF O(6)=0 THEN 1415
1405 NEXT N6
1410 NEXT N5
1415 IF O(4)=0 THEN 1430
1420 NEXT N4
1425 NEXT N3
1430 IF O(3)=0 THEN 1440
1435 NEXT N2
1440 IF O(2)=0 THEN 1450
1445 NEXT N1
1450 REM SUM PEAKS
1455 H=0
1460 A7=B(1)
1465 P5=B(2)
1470 FOR I=1 TO A4-1
1475 J=I+1
1480 B(I)=B(I)+B(J)*D3*A7/(F(1)*A7-P5*D3)
1485 NEXT I
1490 P6=B(1)
1495 FOR I=1 TO A4-1
1500 B(I)=B(I)*100/P6
1505 H=H+(F(I)-B(I))↑2
1510 NEXT I
1535 IF H>P9 THEN 1580
1550 PRINT "FORMULA ";P3
1555 PRINT "CJ←";O(10);"K←HJ←";W2;"K←OJ←";O(8);"K←BRJ←";O(2);
1556 PRINT "K←CLJ←";O(3);"K←FJ←";O(7);"K←IJ←";O(1);"K←NJ←";O(9);
1557 PRINT "K←PJ←";O(5);"K←SJ←";O(4);"K←SIJ←";O(6)
1558 IF A2=1 THEN 1570
1563 WRITE @33:I1,D1,O,W2,D3,B,H,P3
1565 GO TO 1580
1570 WRITE @33:I1,D1,O,W2,P3
1580 IF L(9)=1 THEN 1625
```

PROGRAM ELANAL (cont'd)

```
1620 NEXT U9
1625 IF L(8)=1 THEN 1635
1630 NEXT U8
1635 IF L(7)=1 THEN 1645
1640 NEXT U7
1645 IF L(6)=1 THEN 1655
1650 NEXT U6
1655 IF L(5)=1 THEN 1665
1660 NEXT U5
1665 IF L(4)=1 THEN 1675
1670 NEXT U4
1675 IF L(3)=1 THEN 1685
1680 NEXT U3
1685 IF L(2)=1 THEN 1695
1690 NEXT U2
1695 IF L(1)=1 THEN 1705
1700 NEXT U1
1705 O=0
1720 NEXT D1
1780 G(I1)=P3
1800 NEXT I1
1810 WRITE @33:100000,G
1815 PRINT @32,26:0
1820 END
1825 K1=J1+1
1830 IF J1=10 THEN 1870
1835 K2=0
1840 FOR K3=1 TO J1
1845 K4=K3+10
1850 K2=K2+O(K3)*T(K4)
1855 NEXT K3
1860 J2=J1+1
1865 N(J2)=INT((W1-K2)/T(K1))
1870 RETURN
```

PROGRAM TAPE

```
100 DIM N$(60),L(10),W(10),F(18),Y(10,15),X(50),O(10),B(18)
101 DIM G(10)
105 PRINT "FILE NUMBER "
110 INPUT N1
115 FIND N1
120 PRINT @37,26:1
125 READ @33:N$
126 PRINT N$
127 PRINT "SUSQ UPPER LIMIT"
128 INPUT H1
130 PRINT @40:"                    FILE NUMBER ";N1
135 PRINT @40:"*********************************"
140 PRINT @40:N$
141 PRINT @40:"SUSQ<";H1
145 PRINT @40:"ELEMENTS"
150 READ @33:L
155 IF L(1)=1 THEN 165
160 PRINT @40:"I ";
165 IF L(2)=1 THEN 175
170 PRINT @40:"BR ";
175 IF L(3)=1 THEN 185
180 PRINT @40:"CL ";
185 IF L(4)=1 THEN 195
190 PRINT @40:"S ";
195 IF L(5)=1 THEN 205
200 PRINT @40:"P ";
205 IF L(6)=1 THEN 215
210 PRINT @40:"SI ";
215 IF L(7)=1 THEN 225
220 PRINT @40:"F ";
225 IF L(8)=1 THEN 235
230 PRINT @40:"O ";
235 IF L(9)=1 THEN 245
240 PRINT @40:"N ";
245 PRINT @40:"C "
250 PRINT @40:"IONS"
255 READ @33:W
260 N9=0
265 FOR I=1 TO 10
270 IF W(I)=0 THEN 290
275 N9=N9+1
280 PRINT @40:W(I);
285 NEXT I
290 PRINT @40:" "
295 READ @33:X
300 PRINT @40:"ANALYSIS (0) OR FORMULA ONLY (1) "
305 FOR I=1 TO N9
310 PRINT @40:X(I);
315 NEXT I
316 PRINT @40:" "
320 PRINT @40:"MOLECULAR (0) OR EVEN ELECTRON (1) ION"
325 FOR I=1 TO N9
330 PRINT @40:X(I+10);
335 NEXT I
336 PRINT @40:" "
340 PRINT @40:"NUMBER OF PEAKS "
345 FOR I=1 TO N9
350 PRINT @40:X(I+20);
355 NEXT I
356 PRINT @40:" "
360 PRINT @40:"(A-1) INTENSITIES"
365 FOR I=1 TO N9
370 PRINT @40:X(I+30);
375 NEXT I
376 PRINT @40:" "
380 PRINT @40:"SUSQ"
385 FOR I=1 TO N9
390 PRINT @40:X(I+40);
395 NEXT I
396 PRINT @40:" "
```

PROGRAM TAPE (cont'd)

```
400 PRINT @40:"INTENSITY MATRIX"
405 READ @33:Y
410 FOR I=1 TO N9
412 IF X(I)=1 THEN 445
415 PRINT @40:W(I)
416 PRINT @40:"----"
420 FOR J=1 TO 15
430 PRINT @40:Y(I,J);
435 NEXT J
440 PRINT @40:" "
445 NEXT I
450 PRINT @40:" "
455 PRINT @40:"======================================="
460 ON EOF (0) THEN 710
462 READ @33:I1
463 IF I1=100000 THEN 655
464 IF X(I1)=1 THEN 468
465 READ @33:D1,O,W2,D3,B,H,P3
466 GO TO 470
468 READ @33:D1,O,W2,P3
469 GO TO 472
470 PRINT @40:"# ";I1;" R ";D1;" F ";P3
471 GO TO 475
472 PRINT @40:"# ";I1;" R ";D1;" F ";P3
475 PRINT @40:"--------------"
480 PRINT @40:"C";O(10);
485 IF W2=0 THEN 495
490 PRINT @40:"H";W2;
495 IF O(8)=0 THEN 505
500 PRINT @40:"O";O(8);
505 IF O(2)=0 THEN 515
510 PRINT @40:"BR";O(2);
515 IF O(3)=0 THEN 525
520 PRINT @40:"CL";O(3);
525 IF O(7)=0 THEN 535
530 PRINT @40:"F";O(7);
535 IF O(1)=0 THEN 545
540 PRINT @40:"I";O(1);
545 IF O(9)=0 THEN 555
550 PRINT @40:"N";O(9);
555 IF O(5)=0 THEN 565
560 PRINT @40:"P";O(5);
565 IF O(4)=0 THEN 575
570 PRINT @40:"S";O(4);
575 IF O(6)=0 THEN 579
576 PRINT @40:"SI";O(6);
579 PRINT @40:" "
580 IF X(I1)=1 THEN 460
585 H=INT(100*H)/100
586 IF H>H1 THEN 645
590 PRINT @40:"SUSQ ";H
597 D3=INT(100*D3)/100
600 PRINT @40:"A-1 ";D3
605 FOR I=1 TO 18
610 B(I)=INT(10*B(I))/10
615 NEXT I
620 FOR I=1 TO 18
625 IF B(I)=0 THEN 635
630 PRINT @40:B(I);
635 NEXT I
640 PRINT @40:" "
645 PRINT @40:"=============="
650 GO TO 460
655 READ @33:G
660 PRINT @40:"ION       NUMBER OF FORMULAE EXAMINED"
665 PRINT @40:"-----------------------------------"
670 FOR I=1 TO N9
675 PRINT @40:I,G(I)
680 NEXT I
685 PRINT @40:"-------------------------------"
690 PRINT @40:"END OF FILE"
700 PRINT @37,26:0
710 END
```

isotope abundance array, H is goodness of fit, and P3 the number of the formula.

The information is printed by lines 470-579 (heading and formula), 585-590 (goodness of fit), and 597-635 (calculated isotope abundance).

PROGRAM ISOTOPE

Program ISOTOPE calculates isotope abundancies of specified ions containing C, H, O, Br, Cl, F, I, N, P, S, and Si. The algorithm developed by Kavanagh (1980) in ELANAL is used. The program ISOTOPE is applicable when the isotope abundance of a particular ion is to be calculated and, optionally, compared with the observed abundance. The program avoids the formula calculation procedure of ELANAL since the formula is specified by the user. Consequently, the result is obtained much faster than by ELANAL. On the other hand, no alternative and possibly better fitting formulae for the observed isotope abundance are presented.

The arrays dimensioned in lines 50-60 are similar or the same as in ELANAL. Z(31) is array of factorials, T(10) contains atomic weights of the lightest isotopes of the 10 elements, O(10) contains numbers of atoms of the elements in the specified formula, A(18) contains the calculated, and B(18) the final isotope abundance, C(10,3) stores isotopic formulae, S(10,3) contains the stable isotope abundancies of the 10 elements, and Y(2,15) is used to store entered isotope abundancies. The arrays are initialized and filled as applicable in lines 63-492.

The formula is entered in lines 500-566 and the formula weight is calculated (lines 710-726). Rings and double bonds are calculated in line 740. The information is displayed (lines 750-800).

At this stage, the observed isotope abundance may be entered (lines 840-880) or the program may proceed directly to the calculation of the isotope abundance. If the observed abundance is available, its maximum is found (lines 890-920), the abundance distribution is normalized and stored in the second row of the array Y (lines 930-950).

Lines 1040-1445 contain the Kavanagh algorithm calculating the isotopic abundance distribution. All isotope distributions for the given formula are calculated consecutively in lines 1055-1295 and stored in the array C(10,3). Since iodine is monoisotopic, the calculation begins with Br (line 1085) or the first polyisotopic element present in the specified formula.

In the first pass through the loops the array C contains the lightest isotopes of all elements (the A ion). For example, for the formula $C_3H_{10}N_2S_2$, the rows 4, 9 and 10 of the array C are 2,0,0; 2,0,0; and 3,0,0, respectively. All other rows contain "0"'s. This represents the formula

$$^{12}C_3H_{10}{}^{14}N_2{}^{32}S_2$$

The probability that all three carbon atoms are ^{12}C isotopes is $0.9889^3 = 0.96707$ (0.9889 is the abundance of ^{12}C). Similarly, the probabilities of $^{14}N_2$ and $^{32}S_2$ are 0.9926 and 0.9025, respectively. The probability of simultaneous occurrence of $^{12}C_3$, $^{14}N_2$, and $^{32}S_2$ is the product of these probabilities (0.86633). This calculation is performed in lines 1315-1355.

After the second pass (in the innermost loop P1, lines 1285-1370), the array C is

```
2   0   0    (row 4)
2   0   0    (row 9)
2   1   0    (row 10)
```

and represents the formula $^{12}C_2{}^{13}C_1H_{10}{}^{14}N_2{}^{32}S_2$. The probability of $^{12}C_2{}^{13}C_1$ is $3 \times 0.9889^2 \times 0.0111 = 0.03256$, and the overall probability for this formula is 0.02917.

On the next pass the array C becomes

```
2   0   0    (row 4)
2   0   0    (row 9)
1   2   0    (row 10)
```

and represents the formula $^{12}C_1{}^{13}C_2H_{10}{}^{14}N_2{}^{32}S_2$. The calculations continue as indicated until all isotopic formulae are constructed and their probabilities calculated.

The index A5 (lines 1302-1305) determines the mass of the formula relative to the A ion (A5 = 1 for the A ion, A5 = 2 for the A +1 ion, etc.).

The summed calculated abundancies, stored in array B, are normalized (lines 1455-1495) and displayed (lines 1511-1518). If observed abundancies were entered, goodness of fit is calculated in lines 1550-1570, and the observed abundancies are also displayed (lines 1587-1590).

The results may also be printed on the teletype (lines 1620-1794 and 1900-1950). Lines 1900-1950 contain a "graphic" subroutine for a more illustrative "graphic" presentation on the teletype.

An example of ISOTOPE output for a hexachlorobiphenyl, $C_{12}H_4Cl_6$, is given below. In addition to the formula, the observed isotope distribution was entered as well. Consequently, the "goodness of fit" (SUSQ) was calculated and the observed isotope distribution was normalized and printed for comparison.

```
FORMULA  C12H4Cl6
WEIGHT 358
R&DB 8
 358-------------------- 51
 359- 6.8
 360---------------------------------------- 99.9
 361— 13.4
 362---------------------------------- 81.7
 363- 10.9
 364-------------- 35.7
 365- 4.7
 366- 8.8
 367- 1.1
 368- 1.1
 369- 0.1
SUSQ 28.15
NORMALIZED INPUT
 358---------------------- 54.6
 359- 6.6
 360---------------------------------------- 100
 361— 14.6
 362------------------------------- 81.3
 363- 8
 364-------------- 36
 365- 2.6
 366- 8
**********************
```

PROGRAM ISOTOPE

```
50 DIM Z(31),T(10),O(10),A(18),B(18)
55 DIM C(10,3),S(10,3)
60 DIM Y(2,15)
63 Z=0
64 T=0
66 O=0
67 A=0
68 B=0
71 C=0
72 S=0
75 Y=0
230 DATA 127,79,35,32,31,28,19,16,14,12
235 READ T
240 DATA 1,1,2,6,24,120,720,5040,40320,362880,3628800,3.99168E+7
250 DATA 4.79E+8,6.227E+9,8.718E+10,1.3077E+12,2.092E+13
251 DATA 3.557E+14,6.402E+15,1.21645E+17,2.4329E+18,5.1091E+19
252 DATA 1.124E+21,2.5852E+22,6.20448E+23,1.55112E+25
253 DATA 4.03292E+26,1.08889E+28,3.04888E+29,8.84176E+30
254 DATA 2.65253E+32
255 READ Z
260 DATA 1,0,0,0.5054,0.4946,0,0.7547,0.2453,0,0.95,0.0076,0.0422
265 DATA 1,0,0,0.9221,0.047,0.0309,1,0,0,0.9976,3.99E-4,0.00199
270 DATA 0.9963,0.0037,0,0.9889,0.0111,0
275 READ S
490 PAGE
491 O=0
492 Y=0
500 PRINT "INPUT FORMULA "
505 PRINT "CARBON"
506 INPUT O(10)
510 PRINT "HYDROGEN "
511 INPUT W2
515 PRINT "OXYGEN "
516 INPUT O(8)
517 PRINT "HALOGENS YES(1), NO (2) "
518 INPUT Q2
519 IF Q2=2 THEN 540
520 PRINT "BROMINE"
521 INPUT O(2)
525 PRINT "CHLORINE"
526 INPUT O(3)
530 PRINT "FLUORINE "
531 INPUT O(7)
535 PRINT "IODINE "
536 INPUT O(1)
540 PRINT "NITROGEN "
541 INPUT O(9)
550 PRINT "PHOSPHORUS "
551 INPUT O(5)
560 PRINT "SULFUR"
561 INPUT O(4)
565 PRINT "SILICON"
566 INPUT O(6)
568 PAGE
710 V=0
715 FOR I=1 TO 10
720 V=V+O(I)*T(I)
725 NEXT I
726 V=V+W2
740 R=0.5*(2*O(10)+O(9)+2*O(6)+3*O(5)+2-W2-O(3)-O(2)-O(1)-O(7))
750 PRINT "FOR FORMULA:"
760 PRINT "CJ-";O(10);"K-HJ-";W2;"K-OJ-";O(8);"K-BRJ-";O(2);
770 PRINT "K-CLJ-";O(3);"K-FJ-";O(7);"K-IJ-";O(1);"K-NJ-";O(9);
780 PRINT "K-PJ-";O(5);"K-SJ-";O(4);"K-SIJ-";O(6)
790 PRINT "FORMULA WEIGHT IS ";V
800 PRINT "R&DB= ";R
810 PRINT "INPUT OF ION INTENSITIES: YES(1),NO(2) "
820 INPUT Q2
830 IF Q2=2 THEN 1040
840 PRINT "INPUT INTENSITIES STARTING WITH (A)ION,END BY -1"
845 Q7=0
850 FOR I=1 TO 15
```

PROGRAM ISOTOPE (cont'd)

```
855 INPUT Y9
860 IF Y9=-1 THEN 890
870 Y(1,I)=Y9
875 Q7=Q7+1
880 NEXT I
890 M1=0
891 M2=0
900 FOR I=1 TO Q7
905 M1=Y(1,I)
906 IF M1<M2 THEN 920
908 M2=M1
920 NEXT I
930 FOR I=1 TO Q7
940 Y(2,I)=INT(1000*Y(1,I)/M2)/10
950 NEXT I
1040 A=1
1045 B=0
1055 REM CALCULATIONS OF INTENSITIES
1070 C=0
1085 IF O(2)=0 THEN 1105
1090 FOR N1=0 TO O(2)
1095 C(2,2)=N1
1100 C(2,1)=O(2)-N1
1105 IF O(3)=0 THEN 1125
1110 FOR N2=0 TO O(3)
1115 C(3,2)=N2
1120 C(3,1)=O(3)-N2
1125 IF O(4)=0 THEN 1170
1130 FOR N3=0 TO O(4)
1135 C(4,2)=N3
1140 C(4,1)=O(4)-N3
1145 FOR N4=0 TO O(4)
1150 C(4,3)=N4
1155 C(4,2)=N3
1160 C(4,1)=O(4)-N3-N4
1165 IF C(4,1)<0 THEN 1425
1170 IF O(6)=0 THEN 1215
1175 FOR N5=0 TO O(6)
1180 C(6,2)=N5
1185 C(6,1)=O(6)-N5
1190 FOR N6=0 TO O(6)
1195 C(6,3)=N6
1200 C(6,2)=N5
1205 C(6,1)=O(6)-N6-N5
1210 IF C(6,1)<0 THEN 1410
1215 IF O(8)=0 THEN 1260
1220 FOR N7=0 TO O(8)
1225 C(8,2)=N7
1230 C(8,1)=O(8)-N7
1235 FOR N8=0 TO O(8)
1240 C(8,3)=N8
1245 C(8,2)=N7
1250 C(8,1)=O(8)-N7-N8
1255 IF C(8,1)<0 THEN 1395
1260 IF O(9)=0 THEN 1280
1265 FOR N9=0 TO O(9)
1270 C(9,2)=N9
1275 C(9,1)=O(9)-N9
1280 IF O(10)=0 THEN 1300
1285 FOR P1=0 TO O(10)
1290 C(10,2)=P1
1295 C(10,1)=O(10)-P1
1300 REM CALCULATION OF PEAK NUMBER
1302 A5=C(10,2)+C(9,2)+C(8,2)+C(6,2)+C(4,2)+2*C(3,2)+2*C(2,2)
1305 A5=A5+2*C(8,3)+2*C(6,3)+2*C(4,3)+1
1314 IF A5>18 THEN 1365
1315 FOR I=1 TO 10
1320 IF O(I)=0 THEN 1350
1325 FOR J=1 TO 3
1330 IF S(I,J)=0 THEN 1340
1335 A(A5)=A(A5)*S(I,J)^C(I,J)/Z(C(I,J)+1)
1340 NEXT J
1345 A(A5)=A(A5)*Z(O(I)+1)
```

PROGRAM ISOTOPE (cont'd)

```
1350 NEXT I
1355 B(A5)=B(A5)+A(A5)
1360 A(A5)=1
1365 IF O(10)=0 THEN 1375
1370 NEXT P1
1375 IF O(9)=0 THEN 1385
1380 NEXT N9
1385 IF O(8)=0 THEN 1400
1390 NEXT N8
1395 NEXT N7
1400 IF O(6)=0 THEN 1415
1405 NEXT N6
1410 NEXT N5
1415 IF O(4)=0 THEN 1430
1420 NEXT N4
1425 NEXT N3
1430 IF O(3)=0 THEN 1440
1435 NEXT N2
1440 IF O(2)=0 THEN 1450
1445 NEXT N1
1450 REM SUM PEAKS
1455 H=0
1460 M1=0
1461 M2=0
1465 FOR I=1 TO 18
1470 M1=B(I)
1472 IF M1<M2 THEN 1480
1475 M2=M1
1480 NEXT I
1490 FOR I=1 TO 18
1492 B(I)=INT(1000*B(I)/M2)/10
1495 NEXT I
1511 VIEWPORT 65,130,20,80
1512 MOVE 0,0
1513 SCALE 0.25,2
1514 AXIS 0,10
1515 FOR I=1 TO 18
1516 MOVE I,0
1517 DRAW I,B(I)
1518 NEXT I
1545 IF Q2=2 THEN 1600
1550 S1=0
1555 FOR I=1 TO 15
1560 S1=S1+(B(I)-Y(2,I))↑2
1565 NEXT I
1566 S1=INT(100*S1)/100
1570 PRINT "SUSQ= ";S1
1571 PRINT @37,26:1
1576 PRINT @37,26:0
1587 FOR I=1 TO Q7
1588 MOVE I+0.2,0
1589 DRAW I+0.2,Y(2,I)
1590 NEXT I
1591 VIEWPORT 0,130,0,100
1592 MOVE 0,5
1600 PRINT "TTY OUTPUT YES (1), NO (2)"
1605 INPUT Q1
1610 IF Q1=2 THEN 1800
1615 PRINT @37,26:0
1620 PRINT @37,26:1
1630 PRINT @40:"FORMULA"
1635 PRINT @40:"C";O(10)
1636 IF W2=0 THEN 1645
1640 PRINT @40:"H";W2
1645 IF O(8)=0 THEN 1655
```

PROGRAM ISOTOPE (cont'd)

```
1650 PRINT @40:"O";O(8)
1655 IF O(2)=0 THEN 1665
1660 PRINT @40:"BR";O(2)
1665 IF O(3)=0 THEN 1675
1670 PRINT @40:"CL";O(3)
1675 IF O(7)=0 THEN 1685
1680 PRINT @40:"F";O(7)
1685 IF O(1)=0 THEN 1695
1690 PRINT @40:"I";O(1)
1695 IF O(9)=0 THEN 1705
1700 PRINT @40:"N";O(9)
1705 IF O(5)=0 THEN 1715
1710 PRINT @40:"P";O(5)
1715 IF O(4)=0 THEN 1725
1720 PRINT @40:"S";O(4)
1725 IF O(6)=0 THEN 1735
1730 PRINT @40:"SI";O(6)
1735 PRINT @40:"WEIGHT ";V
1740 PRINT @40:"R&DB ";R
1745 FOR I=1 TO 17
1750 IF B(I)<0.1 AND B(I+1)<0.1 THEN 1770
1760 Y7=INT(0.6*B(I))
1761 GOSUB 1900
1765 NEXT I
1770 IF Q2=2 THEN 1794
1771 PRINT @40:"SUSQ ";S1
1772 PRINT @40:"NORMALIZED INPUT"
1775 FOR I=1 TO Q7
1777 B(I)=Y(2,I)
1780 Y7=INT(0.6*B(I))
1782 GOSUB 1900
1785 NEXT I
1794 PRINT @40:"********************"
1795 PRINT @37,26:0
1800 PRINT "ANOTHER FORMULA (1), END (2)"
1810 INPUT Q3
1820 IF Q3=1 THEN 490
1830 END
1900 PRINT @40:V+I-1;
1910 FOR J=4 TO Y7
1920 PRINT @40:"-";
1930 NEXT J
1940 PRINT @40:B(I)
1950 RETURN
```

ACKNOWLEDGMENTS

I thank Mr. Marcel Babineau for patience and initiative in repairing the Model 6100 Data System and for interfacing a teletype to the Tektronix Plot 50 System, Mr. Steven Bellis for advice on the intricacies of the Tektronix System, Ms. Ruth Garnett for excellent editorial assistance, Mrs. Jeanine Hurley and Ms. Brenda McCullough for skillful typing and arrangement of the report, Mrs. Madelyn Irwin and Miss Marion Haley for meticulous proofreading, and Dr. J. F. Uthe and Mr. A. Sreedharan for comments.

Listing of Finnigan Basic version of ELANAL was kindly provided by Dr. P. E. Kavanagh.

REFERENCES

Kavanagh, P.E. 1980. Program for elemental analysis using low or medium resolution mass spectra. Org. Mass. Spectrom. 15: 334-335.

Sovocool, G. W., R. G. Lewis, R. L. Harless, N. K. Wilson, and R. D. Zehr. 1977. Analysis of technical chlordane by gas chromatography mass spectrometry. Anal. Chem. 49: 734-740.

Wright, L. H., R. G. Lewis, H. L. Crist, G. W. Sovocool, and J. M. Simpson. 1978. The identification of polychlorinated terphenyls at trace levels in human adipose tissue by gas chromatography mass spectrometry. J. Anal. Toxicol. 2: 76-79.

Zitko, V. 1980. PCB's and other organochlorine compounds in sharks and tunas. ICES, C.M. 1980/E:21, Marine Environmental Quality Committee.