QUALITATIVE ANALYSIS OF GAS MIXTURES
BY MICROWAVE SPECTROSCOPY

By William F. White

NASA - Langley Research Center
Langley Station, Hampton, Virginia

Presented at the 155th National Meeting
of the American Chemical Society

San Francisco, California
March 31 - April 5, 1968

QUALITATIVE ANALYSIS OF GAS MIXTURES BY MICROWAVE SPECTROSCOPY

By William F. White
Langley Research Center

INTRODUCTION

The Langley Research Center of NASA has been engaged for several years in the development of microwave spectroscopy as a means of detecting atmospheric contaminants in space cabin simulators. The need for better analytical methods is illustrated in such reports as reference 1. In an experiment involving manned and unmanned simulator operation for 30 days, four cryogenically trapped contaminant samples were taken per day and sent to different laboratories for analysis. The conclusion reached was that the art of contaminant analysis had not reached the point of valid repeatability.

The goal of the work at Langley is to develop an unambiguous method of identifying the components of gas and vapor mixtures, and of measuring the concentrations of each. Only the qualitative aspects will be covered in this paper. Research in this area is primarily concerned with the techniques and reference data needed for the detection of contaminants in air, but most of the work is directly applicable to other analytical problems.

The purpose of this paper is to provide an introduction to the characteristics and capabilities of microwave spectroscopy as an analytical tool, through a review of some of the work at Langley. First, the choice of frequency range will be discussed with respect to the number of molecules which can be detected and the effect of frequency on sensitivity to trace gases. The versatility of the method will be demonstrated with a series of spectra of the same molecule run under different conditions, and the application of the various types of data to analytical work will be discussed. One of the possible approaches to the analysis of an unknown sample will be presented. The paper will conclude with a review of the work now being done to completely automate the operation and data reduction of the spectrometer, and of the work in double resonance spectroscopy.

MOLECULAR COVERAGE AND SENSITIVITY

The frequency range needed is determined by the distribution of absorption lines throughout the microwave range, which is, in turn, a function of molecular structure. Except for a few very simple gases, most polar gases and vapors have absorptions throughout the range accessible to typical microwave spectrometers. The solid curve of figure 1 is a plot of the percentage of molecules with microwave spectra which have one or more absorption lines in a given frequency interval centered at 25 GHz (ref. 2). As the center frequency of the interval is increased, the curve tends to change in the manner indicated by the dashed line, since line intensities generally increase with frequency and a higher proportion of the allowed transitions for a molecule become strong enough to detect. The rectangular area is the interval covered by an R-band (26.5-40 GHz) spectrometer and indicates that approximately 95 percent of the molecules with microwave spectra can be detected in this band. If K-band (18-26.5 GHz) coverage is also provided, there will be few molecules with microwave spectra which cannot be detected.

It is not necessary to use the entire coverage of the spectrometer to detect most molecules, since, as can be seen from the curve, a high percentage of molecules will have lines in a frequency interval much smaller than the total spectrometer range. The data catalogued in R-band so far includes molecules with anywhere from 0 to 1500 lines, with the average number being somewhat less than 400 per molecule. There is, therefore, a good chance of finding one or more absorptions due to the majority of molecules by scanning only a portion of the range available, and any of the lines may be used in identification of a molecule. Normally, the stronger lines are preferred since they provide better sensitivity to trace gases, but any observed absorption provides valid information.

When sensitivity to trace gases is important, it is desirable to operate at the highest frequencies available, due to the increase in line strength with frequency. There is an optimum frequency beyond which sensitivity declines, but this is for most molecules considerably higher than the frequencies available with commercial spectrometers. Figure 2 is a cumulative curve of percentage of molecules which can be detected at various concentrations in a mixture. This curve is based on published peak absorption coefficients below 30 GHz and is obtained by assuming that the coefficient decreases in proportion to the dilution of the sample. In general, the coefficient does not vary linearly with concentration but is dependent on sample composition. The experimental points reflect improved sensitivity over the predicted values, partly due to operation at higher frequencies, but also due to the fact that they were taken using air as a diluent gas. When the diluent is a light non-polar gas, the line strengths do not decrease as rapidly as the sample concentration. The reason is the smaller effective cross-section for collisions in such a gas, which results in less pressure broadening of the absorption lines. Since the integrated intensity over the entire line does vary in proportion to the number of molecules, the narrower lines must be compensated by greater peak heights.

Figure 3 shows the variation of the absorption coefficient and the signal as a function of microwave power in the sample cell. The dashed curve is representative of the behavior to be expected if the line shape were independent of sample composition, and the solid curve shows the effect of using a gas like air as a diluent. The partial pressure of the absorbing molecule in the sample is assumed to be the same. Less power is required to reach the maximum signal in this case; but if sufficient power is not available, there is still more signal than expected on a one-to-one dilution basis. Further, since the linewidths are less at a given pressure in air than for the pure sample, it is possible to operate at higher pressures and increase the signal even more without a loss in resolution as compared to the pure sample.

Figure 4 is an illustration of this effect for methyl chloride. On the right is the spectrum at approximately .03 torr pressure, and the spectrum on the left includes absorptions (numbered 1 and 2) corresponding to those in the pure sample. The latter spectrum is from a calibration sample containing 75 ppm methyl chloride and 76 ppm Freon-22 in air, run at .3 torr. There is a Freon doublet in this spectrum, but it is evident that the resolution is comparable in the two cases in spite of the large pressure difference. Also, note that the signal-to-noise ratio is sufficient to detect somewhat smaller concentrations than 75 ppm, although even this sensitivity should not be expected on the basis of the absorption coefficient for the pure methyl chloride. This improvement in sensitivity over the expected value is peculiar to dilution with a non-polar gas. If the diluent is polar, there may be little effect on linewidths, or the effect may be the opposite of that described, resulting in less sensitivity than predicted.

EFFECTS OF SAMPLE PRESSURE AND SCAN RATE

Both the total scan time and the frequency range scanned are continuously variable over a wide range with modern microwave spectrometers, so that there is literally an infinite number of scan rates. The maximum rate which may be used in a given case depends on the strength and width of the lines to be observed, since these factors determine the necessary detector time constant. These absorption parameters are, in turn, dependent on the sample composition and pressure. High pressures (a few tenths of a torr or greater) naturally allow the fastest scans, but also result in the poorest resolution and frequency measurement accuracy. There is also the problem of glow discharges if too high pressures and Stark modulation fields are used at the same time. When operating at pressures of the order of one torr or more, it is usually necessary to use relatively low modulating fields. For analytical purposes, this does not appear to be a disadvantage, since many of the lines are severely undermodulated and are not observed, resulting in a simpler spectrum. Figures 5-9 illustrate the effects of pressure on resolution. Progressively smaller portions of the range are covered to keep the scan times comparable in each case.

Figure 5 is the spectrum of methylene chloride at about two torr. There are around 1200 discrete lines catalogued in this region, which comprises all of R-band. These lines vary in intensity about three orders of magnitude. The wave numbers are shown for comparison with infrared spectra. The only distinctive feature of this spectrum is the strong line at 40 GHz, which will be seen later in a mixture. In general, one finds more of a discrete line spectrum, even at this high pressure. The lack of resolution and accuracy is obvious here, but this type of scan covers the entire range of the spectrometer and can be done in as little as 1 minute.

As the pressure is decreased, the broad bands of figure 5 split up into resolvable sharper lines, as shown in figure 6. This is the portion of the spectrum indicated by brackets in figure 5, run at .15 torr pressure. Resolution and frequency measurement accuracy are greatly improved here, at the expense of speed. A scan of R-band with this resolution can be accomplished in an hour or more, depending on the strength of the lines. Note the larger quantities of data available here. The amount of information available from a low resolution spectrum can be duplicated by arbitrarily choosing a small portion of R band for observation of the methylene chloride spectrum with the improved resolution of figure 6. However, there is more information available than is evident in this spectrum. The portion shown in brackets is repeated in figure 7 at .03 torr pressure. The single strong line is now seen to be a triplet. Frequency measurements can readily be made to 5 or 6 digit accuracy at pressures of this order, but the scan rate is such as to take several hours to scan the entire band, and longer if the lines are weak.

Further reducing the pressure to .01 torr and slowing the scan rate correspondingly, shows each component of the triplet in figure 7 to be a triplet itself. Frequencies can be measured to seven digit accuracy under these conditions; and, in this particular case, there are nine lines where only one is observed at higher pressures. There is additional information available from the relative intensity pattern of the lines. However, attaining this resolution and accuracy over the entire range involves one or more days scan time. This is the type of operation necessary to observe all of the 1200 catalogued lines in R-band. Fortunately, it is not necessary for most purposes, other than initial data cataloguing, to operate over wide frequency ranges at maximum sensitivity and resolution.

Figure 9 represents the maximum resolution which can be achieved with ordinary spectrometers. The mean free path of molecules in the gas is such at these low pressures that collisions with the cell walls predominate and further reduction of pressure reduces the signal strength without affecting linewidth. It is also difficult to avoid power saturation broadening at very low pressures. Operation at maximum resolution should be necessary only if interference is experienced from closely adjacent lines or if maximum accuracy in frequency measurement is desired.

USE OF MICROWAVE SPECTRA TO IDENTIFY MOLECULES

One of the areas of research at Langley is how best to use the various types of data illustrated. The first approach which comes to mind is to utilize the high resolution and accuracy of the spectrometer by choosing a few of the strong lines for identification purposes and ignoring the rest of the spectrum. The presence of absorptions at several of these frequencies is taken as proof of the molecule's identity. This is quite satisfactory for molecules with few absorption lines, as is the case for simple molecules or when dilution has made all but the strongest lines undetectable. For the more realistic case where some molecules may be present in high concentrations, and where some of the molecules may have more than 1000 absorption lines in R-band alone, this approach can lead to undetected errors in analysis. In particular, it cannot be used if there is a chance of more than two components in the gas, as the possibility of error increases very rapidly with the number of components in a mixture.

Figure 10 illustrates the type of ambiguity which may occur in a mixture. This sample contains methanol, ethanol, isopropanol, methylene chloride, and chlorobenzene in air. On the left, there are two lines due to methylene chloride which are coincident with two lines of isopropanol. At higher frequencies, the situation is repeated with interference between methylene chloride and methanol lines. Thus, the information that absorption occurs at these four frequencies is not sufficient to know whether it is due to methylene chloride, to methanol and isopropanol combined, or to all three - not to mention the possibility of one or more of the lines being due to some substance which is not in the catalog. This points up the importance of having complete spectral data available for reference, even though it is not all necessary for a given identification. This problem does not exist for one or two components in a sample but can be serious for complex mixtures.

The illustration is given not to imply that the microwave method is inherently ambiguous, but to emphasize that attention must be given to the entire spectrum of a molecule, rather than just a few lines. Another advantage of knowing the entire spectrum is that it allows identification of all lines observed for a particular unknown sample so that absorptions due to molecules which may not be in the catalog can be recognized. It is then known that there are one or more components in the sample which have not been identified; and, if desired, the data may be retained and will be valid whenever a more comprehensive catalog is obtained.

The situation may be summed up as follows: It is necessary to have available for reference the entire spectrum of every molecule, in order that ambiguities such as the one illustrated may be recognized and eliminated. Since data catalogs will be incomplete for at least the near future, this does not allow determination of all possible errors. However, if a suitable amount of data is is available, statistical analyses can be done to determine the probability of interference. Such a study is in progress at Mississippi State University where about 10,000 line frequencies have been put on punched cards and are being analyzed by computer. Information gained in this manner can be used to determine how much data of a given accuracy are needed to provide the desired confidence level in an analysis. It has been seen that, in general, more than two or three measurements will be needed. The entire spectrum of a molecule provides a unique fingerprint, but this is much more information than needed to identify the molecule. The answer is somewhere in between and depends strongly on how many components are in the sample.

A general approach to the solution of the data problem is to consider the case of N randomly distributed lines with M resolvable "slots" in which they may be placed. This approach can be equally applied to results from any of the types of spectra illustrated. In fact, there is a great similarity in appearance if the frequency interval is reduced enough with each increase in resolution. Compare figures 7 and 11. The latter is a scan over all of R-band of a sample containing n-propyl chloride, which is responsible for the six strong lines seen, three for each of the two chlorine isotopic species. There are hundreds of lines in this region at maximum resolution. Except for the frequency scale, there is no way to tell that these spectra were not run under the same conditions. Thus, the statistics for one case are valid for the other. There is approximately the same amount of information available in each case, and it can be obtained in about the same length of time, even though the absolute accuracy and resolution are different in the two cases by almost two orders of magnitude. The choice of which mode to use must be based on the following considerations: (1) the low resolution scan covers the entire range of the spectrometer, whereas the high resolution scan covers only a small portion, and a single such scan will not be enough to detect all molecules; (2) in the case of a mixture, the low resolution scan may not provide enough information to complete the analysis, whereas as much more range as needed may be scanned in the high resolution case; (3) not all molecules exibit useful spectra at both extremes of pressure. Examples are the methylene chloride spectrum in figure 5, which does not show discrete lines at high pressures, and molecules like nitrobenzene which have strong lines at high pressure which split up into dense clusters of weak lines at lower pressure.

One way in which the desirable characteristics of both types of spectra may be used is to begin with a high pressure sample (of the order of 1 torr) and do a full band scan. A fast scan (2 minutes) will show if the sample has one or more molecules in high concentrations and with strong lines. This may be repeated at a slower rate and higher gain to detect the minority constituents, or if nothing is seen the first time. There will be many cases where no further data are needed; but, in general, it will not be possible at this stage to positively identify every line observed, due to the lack of resolution and accuracy. What will be achieved is narrowing down the possible constituents of the sample to a reasonably short list. Figure 12 shows a scan of the same mixture as that in figure 10, run in about 15 minutes. All of the strong lines are due to methanol, which was easily identified on a 2-minute scan. The 40 GHz methylene chloride line can be seen, and the lines due to the other components are weak. Methanol can be positively identified at this stage, but more data is needed to positively establish the identity of the others.

The final identification is done by lowering the pressure and searching for some of the strong lines for each molecule at high resolution and accuracy. Figure 13 illustrates several such lines picked from a mixture of methylene chloride, methyl ethyl ketone, and five alcohols in air. In this case, some of the chosen lines were obscured by interference from the other molecules in the sample, but enough are available without interference to complete the analysis. These narrow-band scans are done automatically at Langley by means of a punched tape frequency programmer which allows scans of any number of small pre-selected intervals. A library of tapes may be kept on hand to suit various situations.

In the event that the results of the low-resolution scan are completely negative, it is an indication that any absorbing molecules which may be in the sample are only present in trace amounts. In this case, maximum sensitivity will be needed, but since only a few lines will be detectable for each molecule, the chances of interference are almost eliminated and a minimum of data is needed. The time saved in not scanning a number of frequency intervals at high sensitivity may be much more than that used in making the low-resolution scan. In any event, some idea of the complexity of the spectrum is obtained and this is important in the determination of how many lines per molecule should be used for identification.

AUTOMATED QUALITATIVE ANALYSIS

Whether 3 or 10 or 50 lines per molecule are used, it is evident that a great amount of data will be needed to confirm the presence (or absence) of very many molecules. As data catalogs become more complete, there will be less advantage in scanning numerous small intervals over making a wider scan and using all of the lines detected. However, in order to make use of all of the information available, it is almost necessary to use a computer for data reduction, and an automated spectrometer is highly desirable to reduce the amount of operator time spent in making the measurements. An automated spectrometer is expected to be in operation at Langley next year, and the data reduction program is presently being developed. The ultimate goal is real-time analysis, so that the operation of the spectrometer may be continuously modified as appropriate to provide only the amount of data necessary. The development and testing of this program is also providing valuable insight into analytical techniques and data handling in general.

Figure 14 is a simplified flow diagram of the computer program. The catalog, for test purposes, contains an average of 20 line frequencies and intensities for each of 15 molecules. The intensities are defined on a logarithmic scale related to the absorption coefficient. The first version of the program uses only frequencies to identify the molecules, and the second version also uses intensity information. The sample data input consists of anywhere from 30 to 75 line frequencies and intensities taken from a full band scan of the sample. Real-time operation is presently simulated by keeping the sample data in the same order in which it comes from the spectrometer. As each sample frequency is read, it is compared with the tables of data in the catalog. The program looks for matching frequencies within a preselected error limit. A running tabulation is kept on the possible identifications, including all information which might be useful such as frequency and intensity errors and the number of lines tentatively assigned to each molecule compared to the number expected. A printout of the status is made after each sample line is checked. This part of the program will ultimately be the one where decisions are made and instructions issued to the spectrometer rather than printing out the results.

Figure 15 is a typical printout after about two-thirds of the data had been analyzed for a mixture of ethanol, isopropanol, and methylene chloride in air. The top portion is from version one of the program, and the bottom one is the corresponding report from version two. The score for version one is simply the percentage of catalog lines which have been matched for each molecule in the range covered. Version two of the program adjusts the score for lines which may have been missed due to dilution by comparing the observed intensities with those in the catalog and determining the weakest line which could be seen in the sample. This has improved the scores somewhat but is especially important when a more complete catalog is included which has a wider range of intensities than the test catalog.

In addition to the scores, the calculated errors aid in determining correct assignments of lines to molecules. Note the small positive frequency error for the three molecules actually present. This is a systematic error due to the combination of scan rate and detector time constant used, and a similar error should hold for any line which has been correctly assigned. The variance is useful here to insure that the error is actually systematic and as an indication of the measurement precision. The average intensity error may be large as a result of dilution of the sample; but if the variance is small, it is an indication that not only the frequencies match the catalog but the relative intensities as well.

It is difficult to predict the final form of the program, but it will probably operate in a manner similar to that described earlier for analysis of an unknown. That is, after some initial data has been obtained to allow a tentative analysis, the computer will choose specified frequencies for the verification of the preliminary analysis rather than continuing to scan through the entire band.

DOUBLE RESONANCE SPECTROSCOPY

All of the preceding discussion has concerned means of handling the large quantities of data obtained from conventional spectrometers. Double resonance modulation is being studied as a means of simplifying the observed spectrum. Figure 16 shows some energy level systems which exibit double resonance spectra. Pairs of transitions are indicated, and each individual transition represents one which can be observed by a conventional spectrometer tuned to the proper frequency. As has been shown, in a mixture there are likely to be other adjacent lines, and the object is to prevent the observation of these simultaneously with the desired line. If one of the lines of a pair is being observed and a second microwave field is applied at the frequency of the other line, an intensity change in the observed line will be seen due to coupling through the common energy level. The second (pumping) microwave field may be frequency or amplitude modulated, and the intensity of the first line will vary at the modulation frequency of the pumping field. This effect may be used instead of the Stark effect or other modulation method. The advantage of this method is that it selectively modulates only the line which is connected to the pumping line frequency by a common energy level, and none of the adjacent lines are observed. Thus, the presence of a signal output requires not only that the molecule observed have an absorption at a given frequency, but that this line be connected with a specific second line at a different frequency through a common energy level. This reduces the number of absorptions which must be observed since the probability of more than one molecule having a pair of lines which coincide with those of another molecule is not negligible, but the chance of these pairs being connected by a common energy level in each case is slight.

The simplification achieved is evident from figure 17, which is a portion of the conventional R-band spectrum of a mixture of n-propyl chloride, dimethyl sulfide, and propionic acid in air. There are many lines present, mostly due to n-propyl chloride, making the detection of the other two components somewhat difficult. The three transitions indicated can be observed by the double resonance technique, using a pumping frequency in the P-band (12.4-18 GHz) region. Figure 18 shows the double resonance lines obtained in a coiled waveguide cell by Research Systems, Inc. (ref. 3). There is no interference since each line must be observed individually by tuning the pumping source to the proper frequency as well as tuning the observing frequency properly. Dimethyl sulfide has two lines in this region pumped at the same frequency.

The long waveguide cell has undesirable microwave and vacuum properties, and a new type of cell is being tested at Langley to improve these. Figure 19 is a block diagram of the spectrometer with the new cell, which consists of a pair of semi-confocal Fabrey-Perot interferometers mounted at right angles in a suitable vacuum enclosure. One acts as a resonant cavity at the pumping frequency, and the second acts as a resonant cavity at the observing frequency. Double resonance transitions occur in the space common to both. A long effective path and efficient pumping are achieved in a more compact cell than the waveguide, and the vacuum properties are greatly improved. The observed signal-to-noise ratios are comparable with conventional spectrometers. The double resonance technique does not appear suited to accurate quantitative work and is not expected to replace Stark modulation as the most widely used modulation technique. However, it is highly specific and has simple data output and can be made into a small package. The version with the cavity cell is particularly suited to fixed frequency operation and might be better than a Stark system for jobs involving continuous monitoring of a certain gas in a complex or varying mixture where there could be interference from other molecules when using conventional methods.

CONCLUSION

The value of microwave spectroscopy in qualitative analysis of gas and vapor mixtures has been established. A substantial amount of work is still necessary to establish guidelines for the amount and nature of data necessary to avoid ambiguities and to increase the amount of reference data available in the catalogs. Some of the factors which make microwave spectroscopy so attractive, such as its high resolution and accuracy, sensitivity, and versatility, contribute to the problem by making possible so many alternate approaches to the solution of a problem. To maintain perspective, however, it is only necessary to recall that none of the other techniques available can be considered to have been perfected yet, but this has not prevented their utilization. Modern microwave spectrometers are capable of providing data with an accuracy, specificity, and repeatability unequalled by any other single instrument used for gas analysis.

REFERENCES

1. Conkle, J. P., et al, "A Detailed Study of Contaminants Produced by Man in a Space Cabin Simulator at 760 mm Hg," USAF SAM-TR-67-16, 1967.
2. Townes, Charles H., and Schawlow, Arthur L.; "Microwave Spectroscopy," McGraw-Hill Book Co., Inc., 1955.
3. Volpicelli, Richard J.; Stiefvater, Otto L.; and Flynn, George W.: "A Study of Trace Contaminant Identification by Microwave Double Resonance Spectroscopy;" NASA CR-967, 1967.

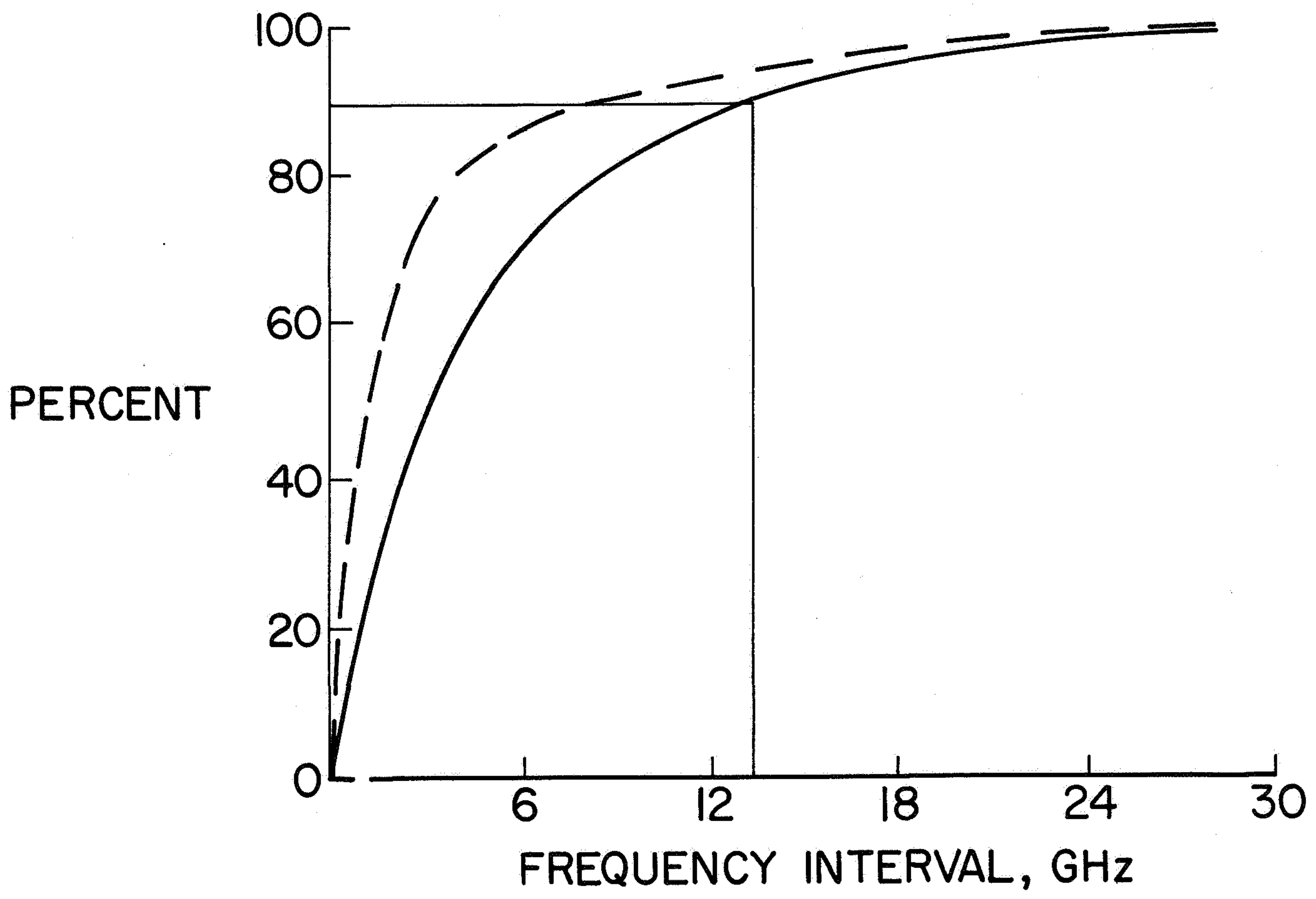


Figure 1.- Percentage of molecules detected versus frequency interval.

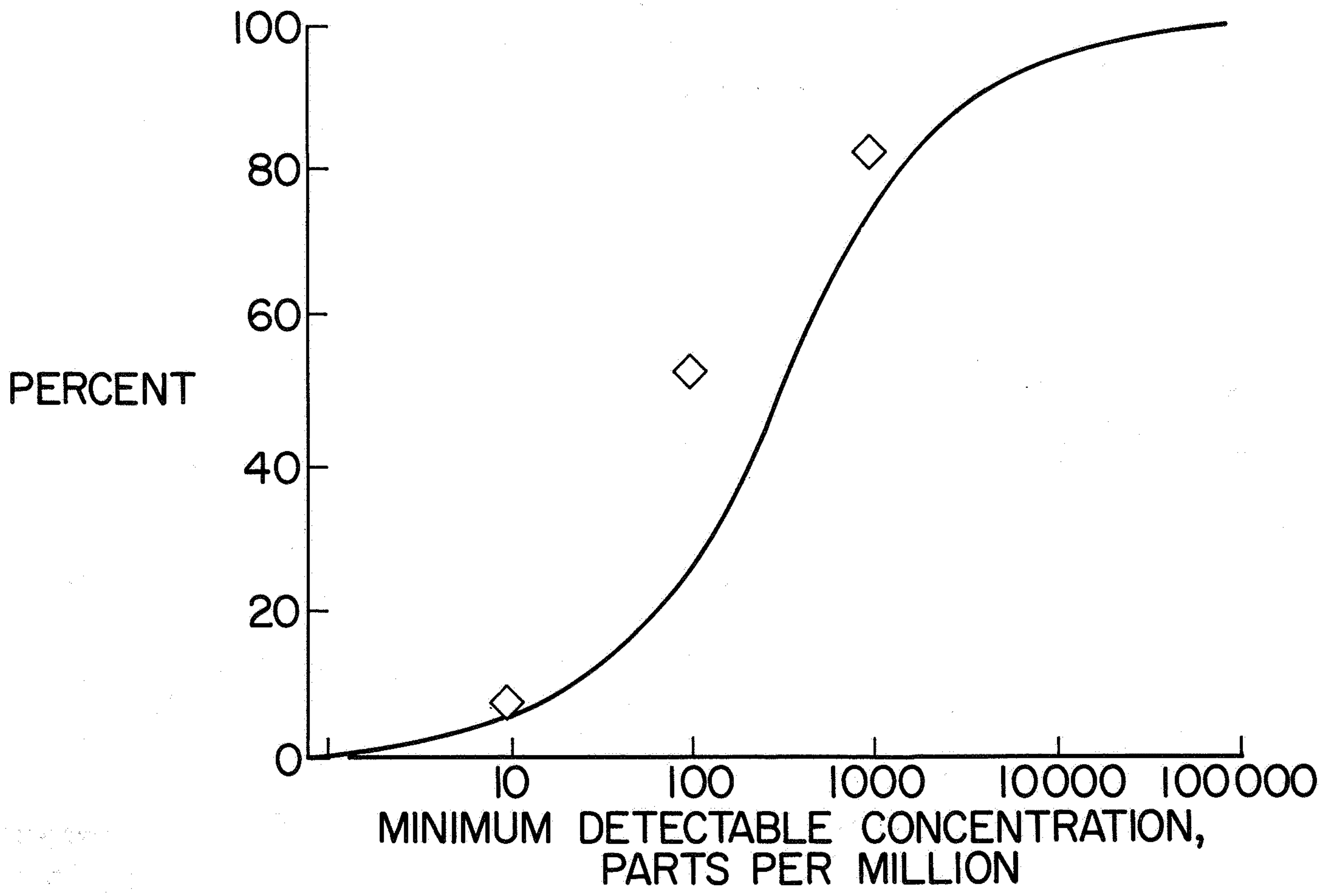


Figure 2.- Percentage of molecules detected versus concentration in air.

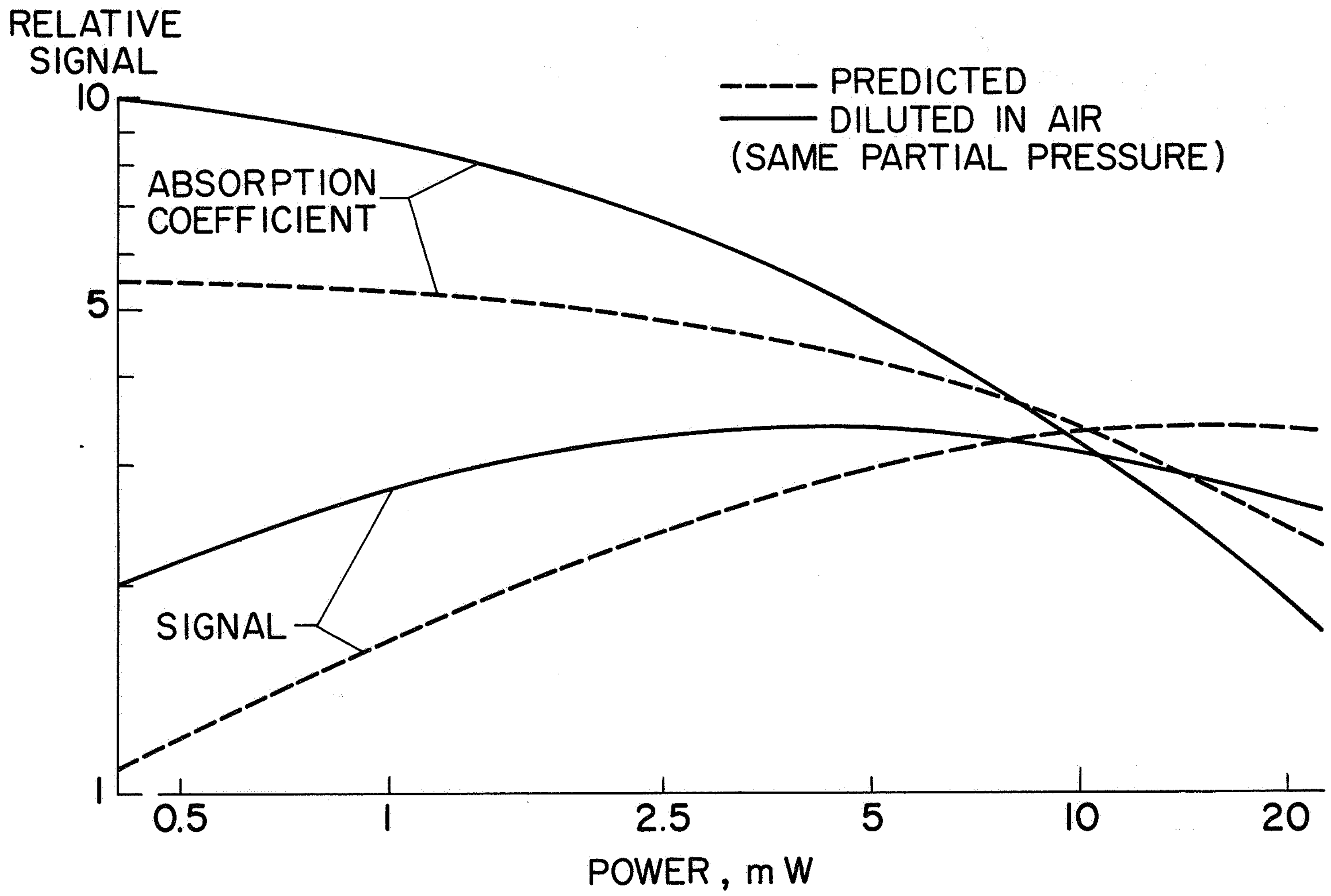


Figure 3.- Absorption coefficient and signal versus power.

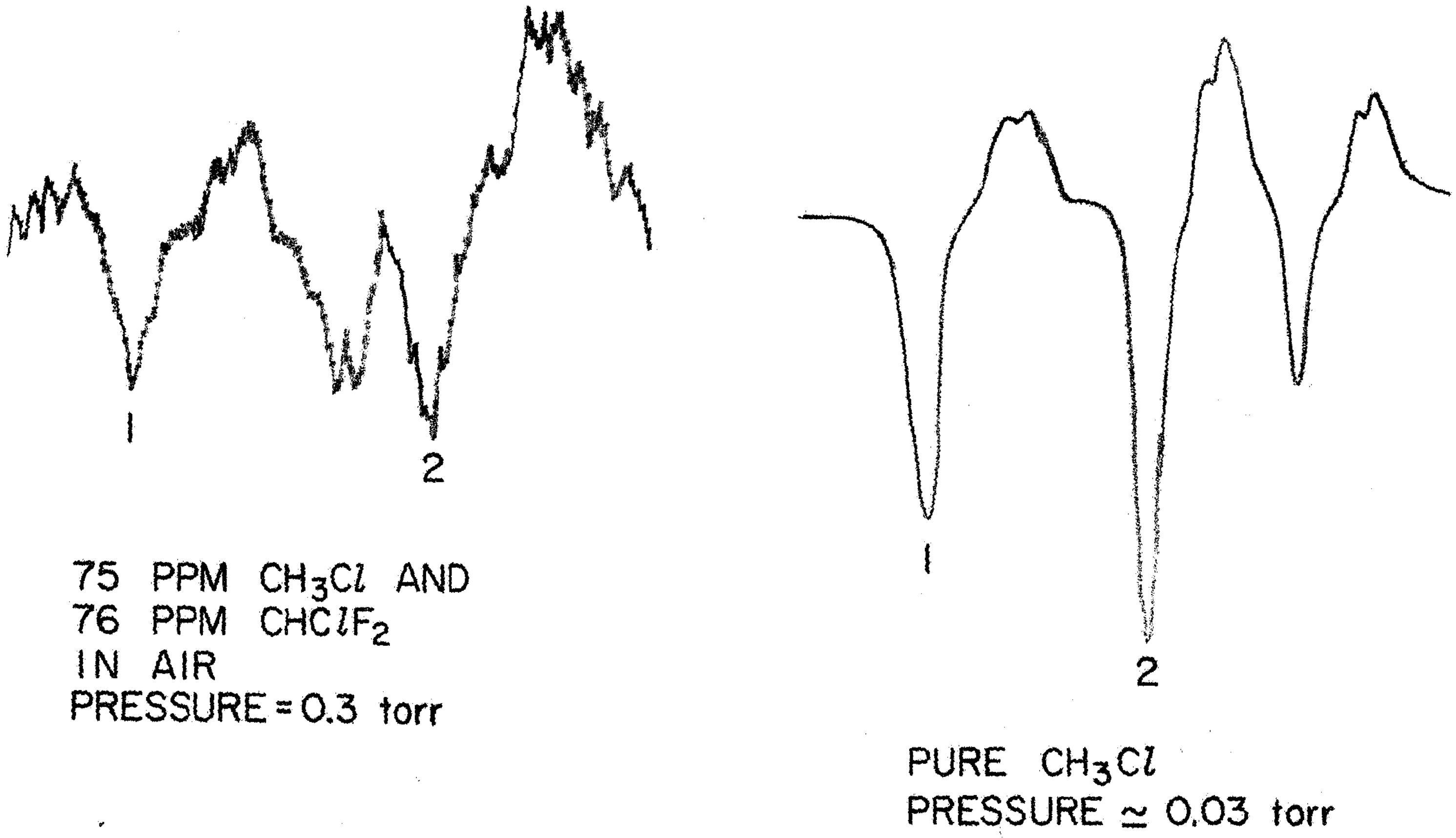


Figure 4.- Effect of dilution in air.

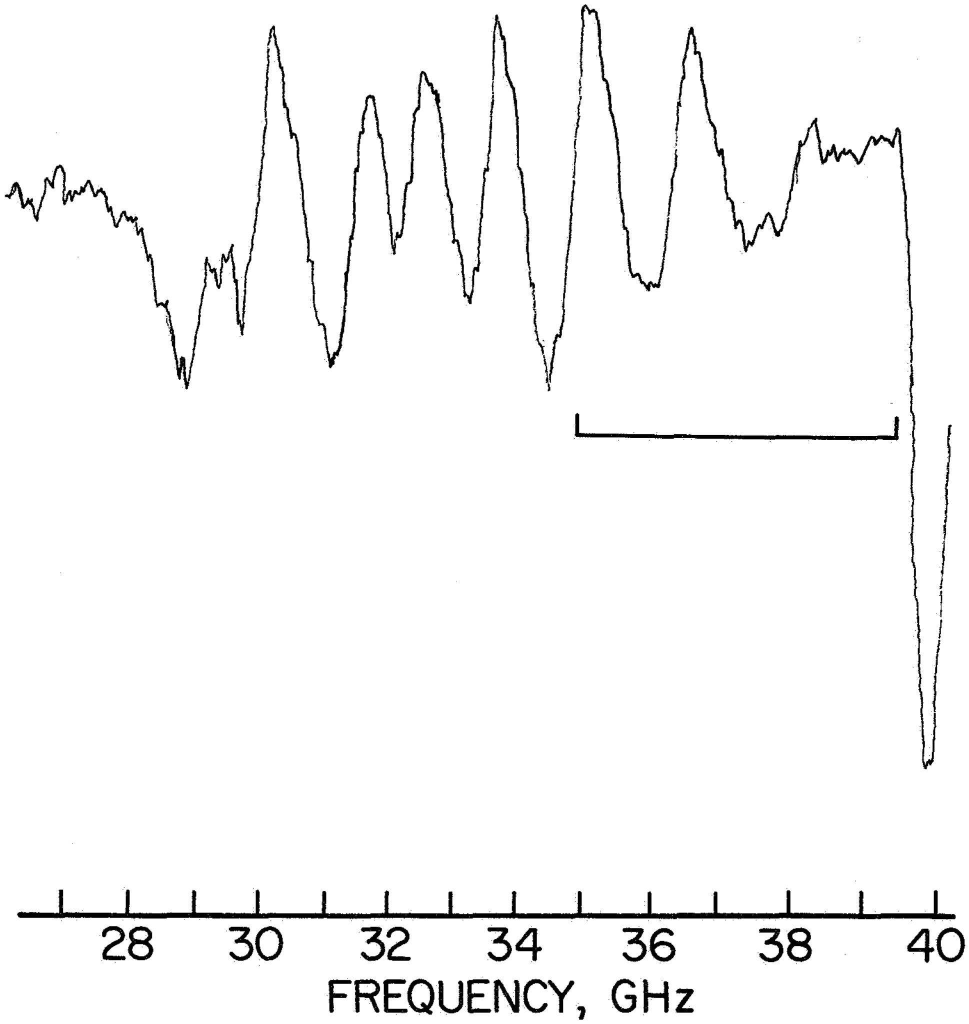


Figure 5.- Spectrum of methylene chloride at 2 torr.

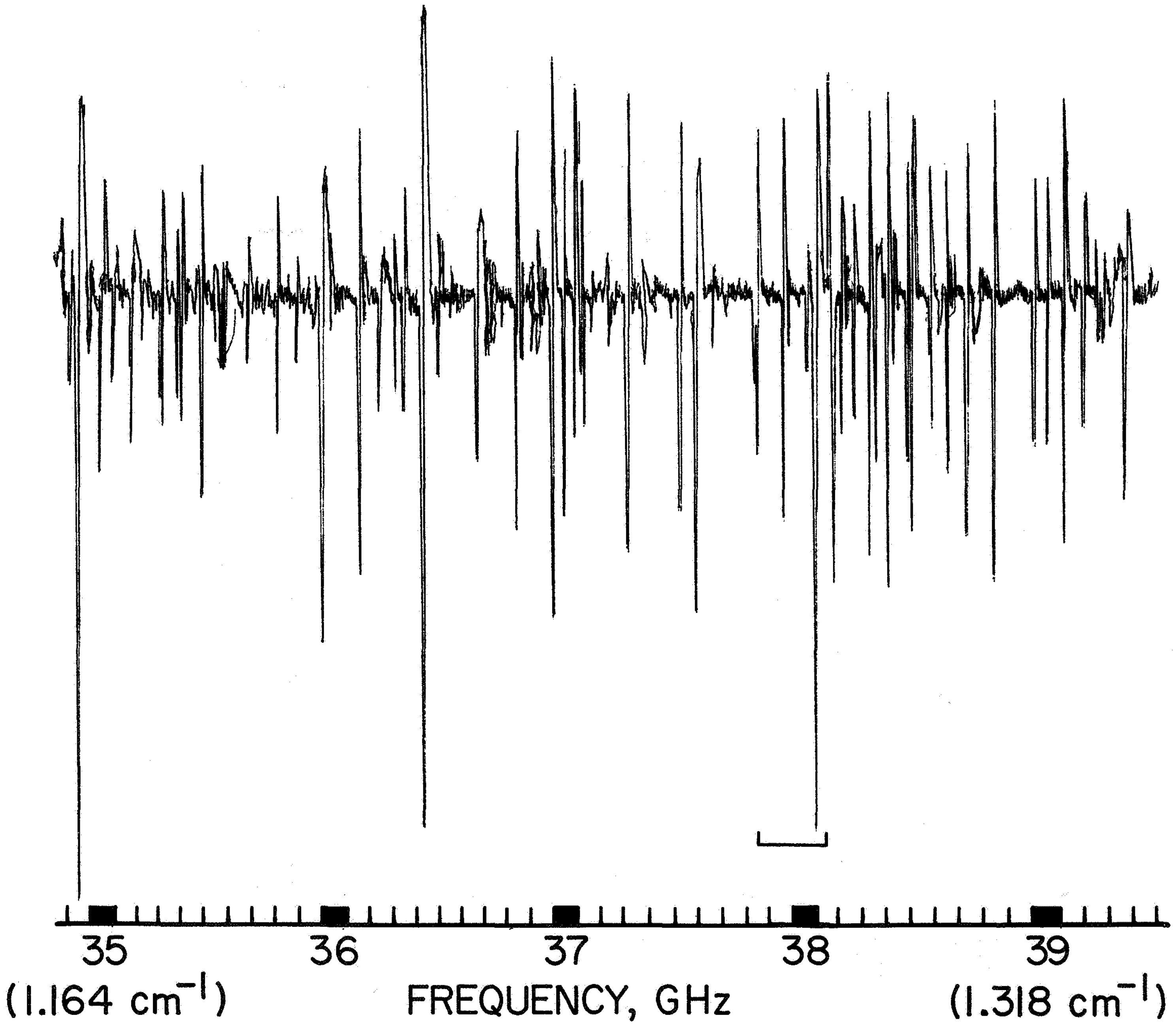


Figure 6.- Portion of methylene chloride spectrum at 150 microns pressure.

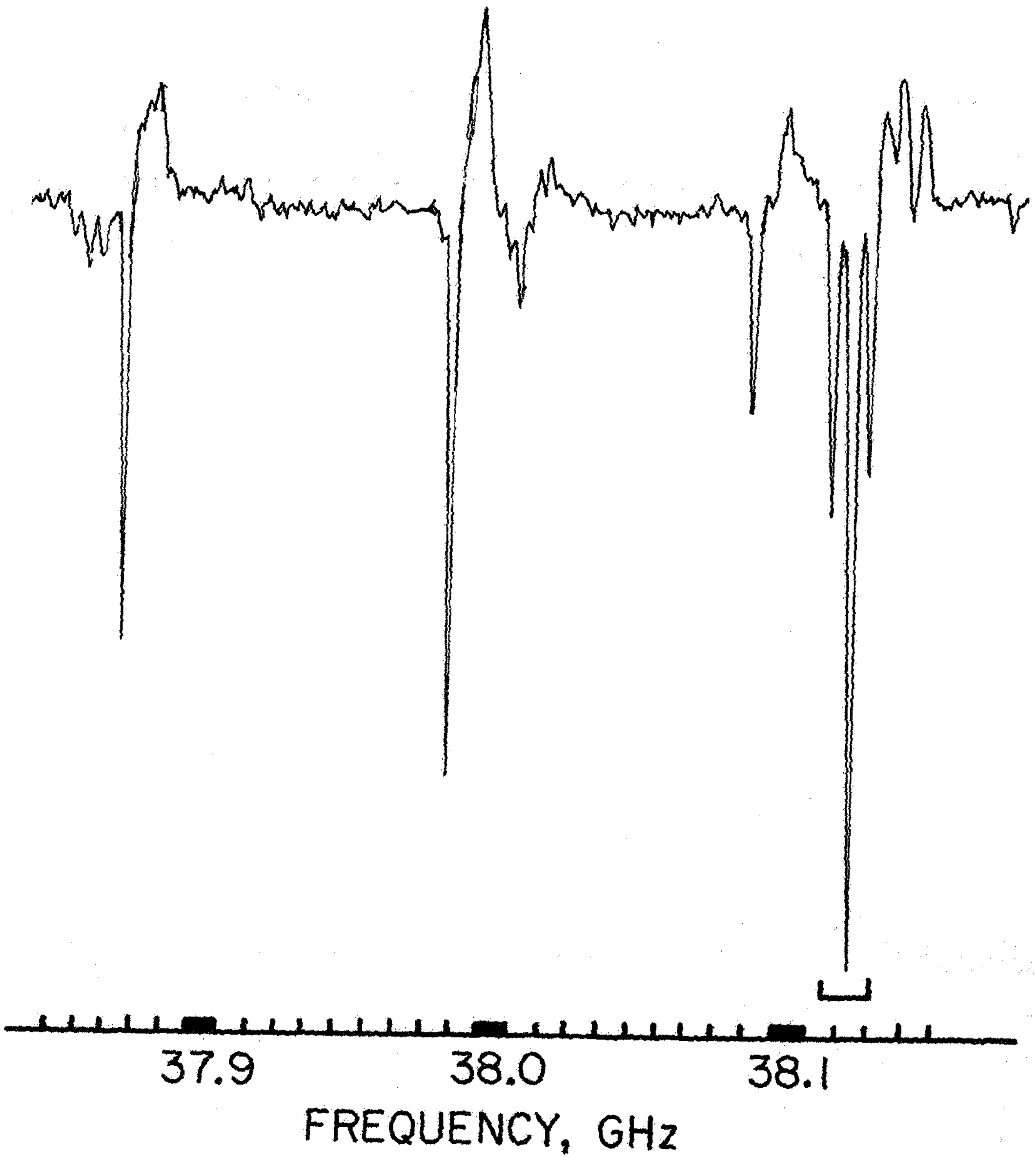


Figure 7.- Portion of methylene chloride spectrum at 30 microns pressure.

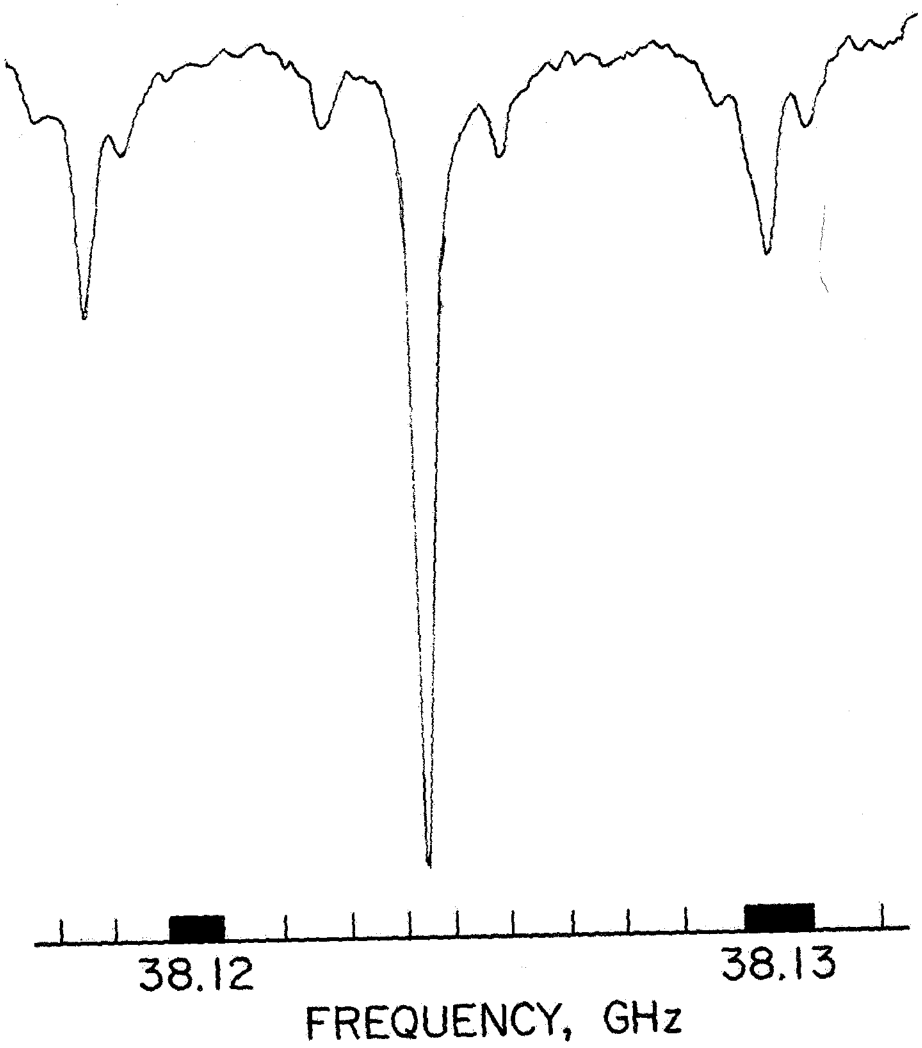


Figure 8.- $11_{0,11} \rightarrow 11_{1,10}$ transition of $CH_2{}^{35}Cl_2$ at 10 microns pressure.

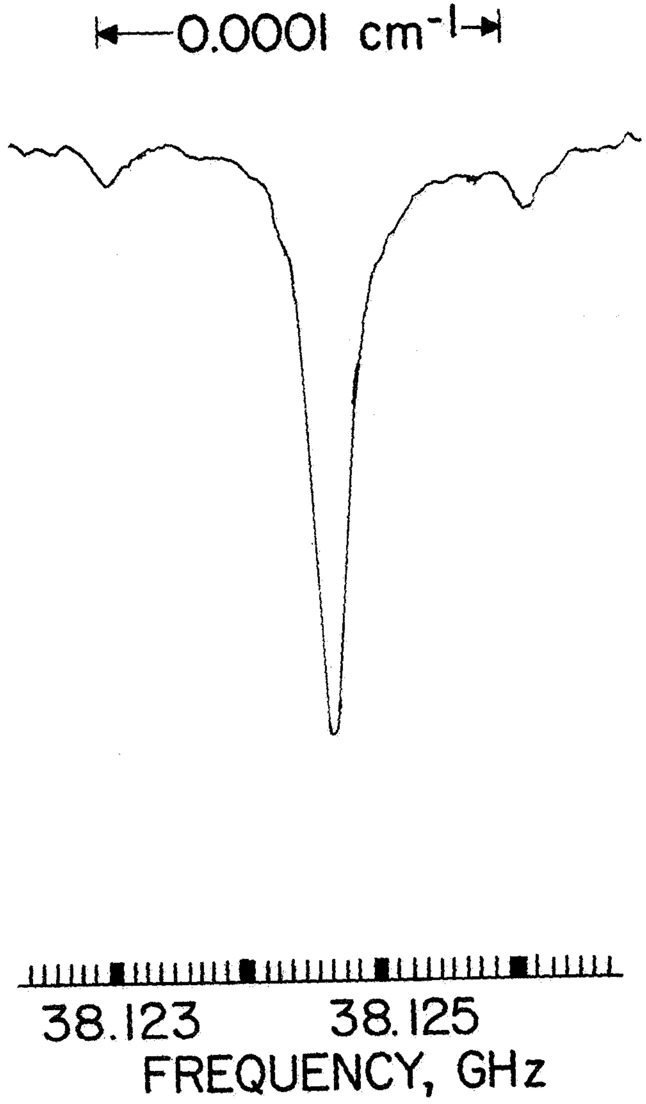


Figure 9.- $11_{0,11}$ - $11_{1,10}$ transition of methylene chloride (Cl^{35}) at 5 microns pressure.

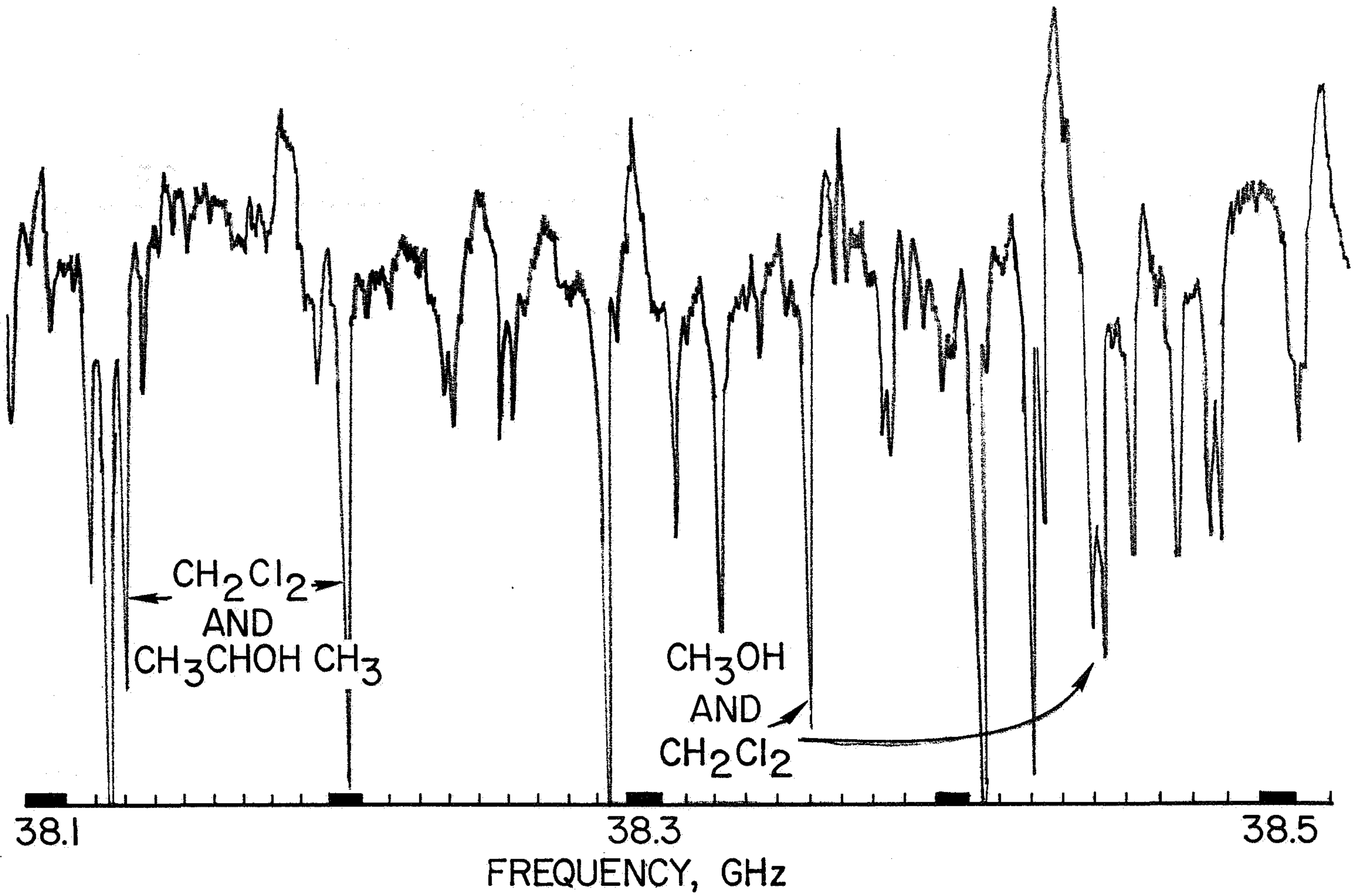


Figure 10.- Five organic vapors in air.

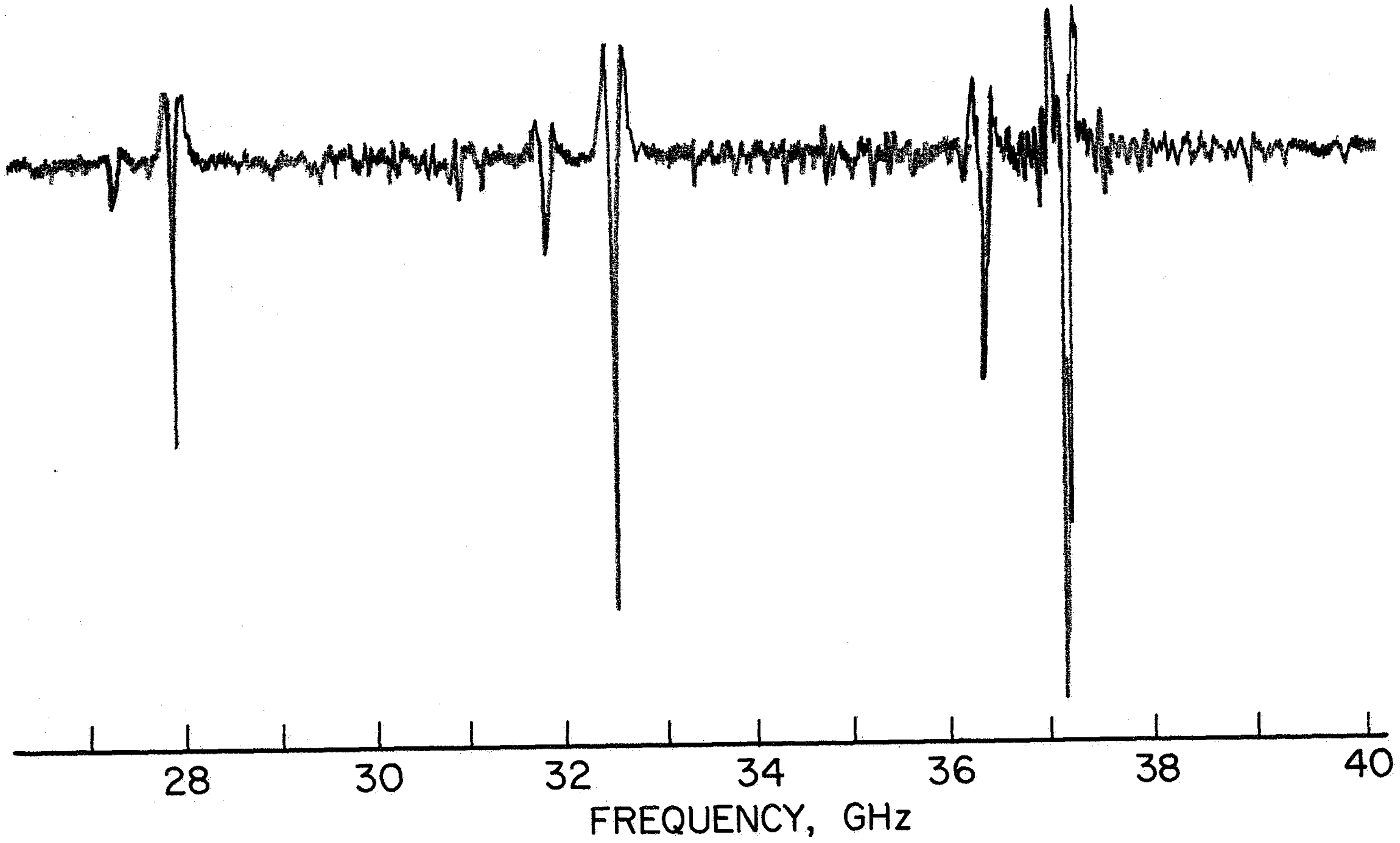


Figure 11.- $CH_3(CH_2)_2Cl$, $(CH_3)_2S$, CH_3CH_2COOH in air.

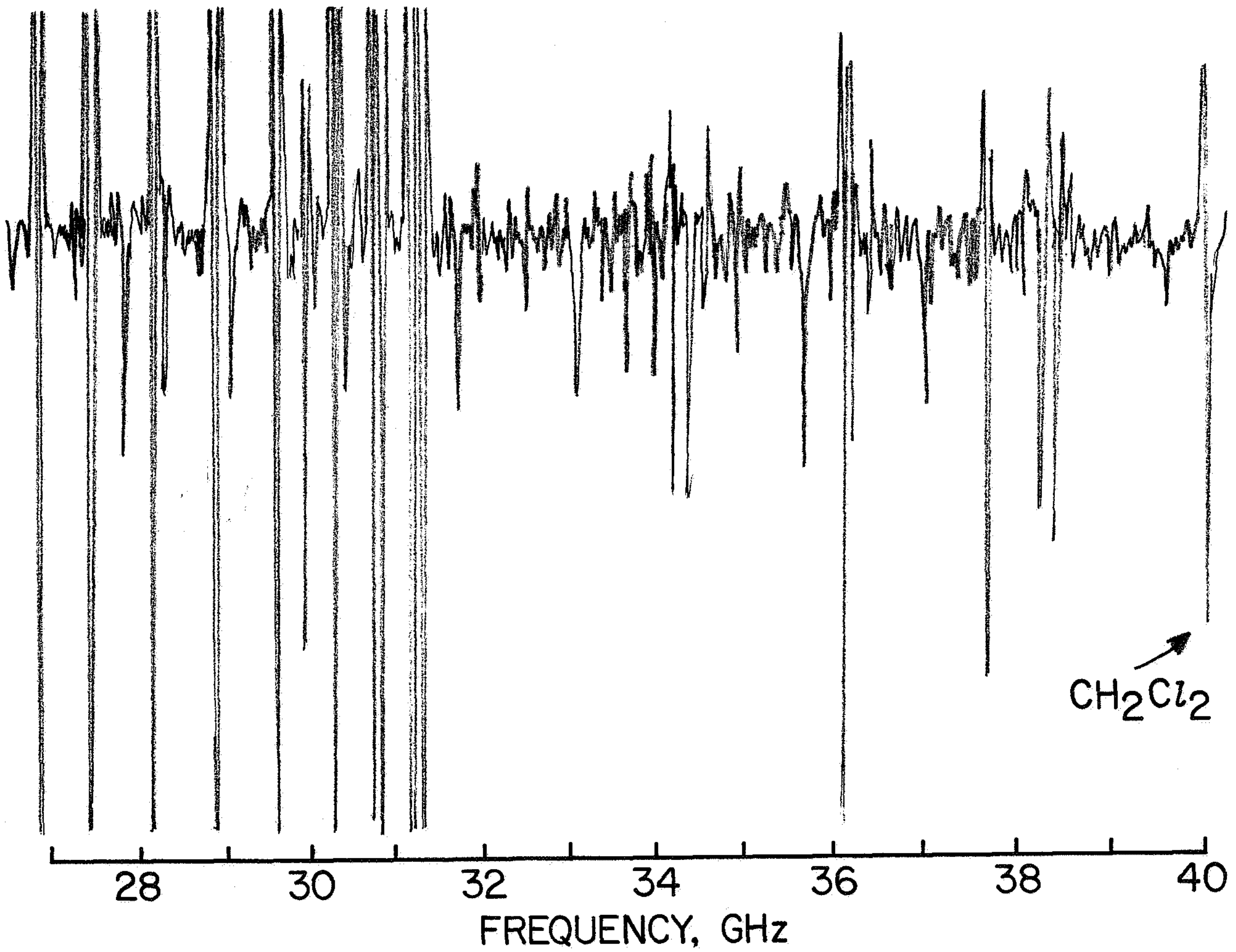


Figure 12.- Mixture of 2 percent each CH_2Cl_2, CH_3OH, CH_3CH_2OH, $CH_3CHOHCH_3$, C_6H_5Cl in air.

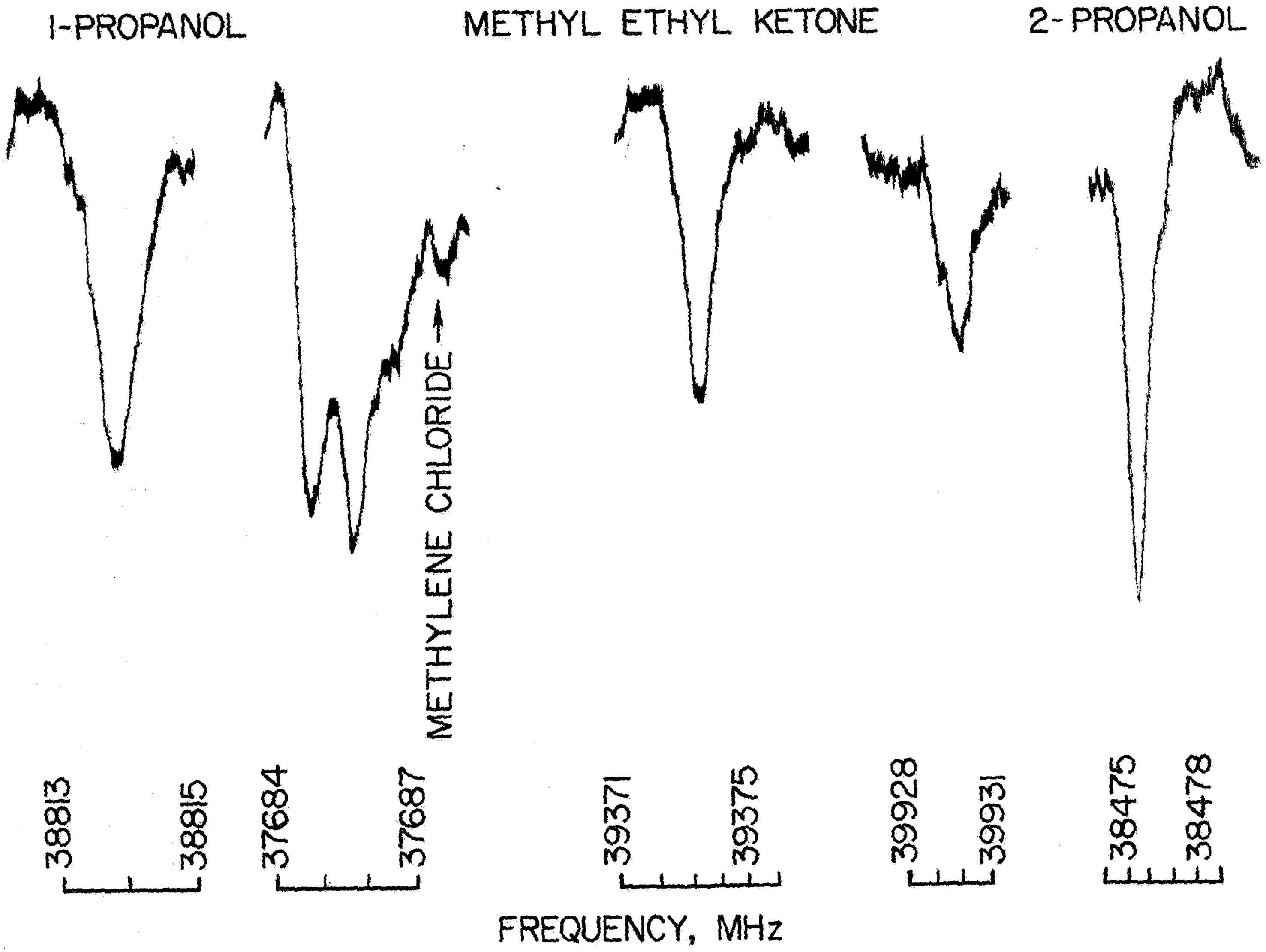


Figure 13.- High resolution spectra of gas mixture.

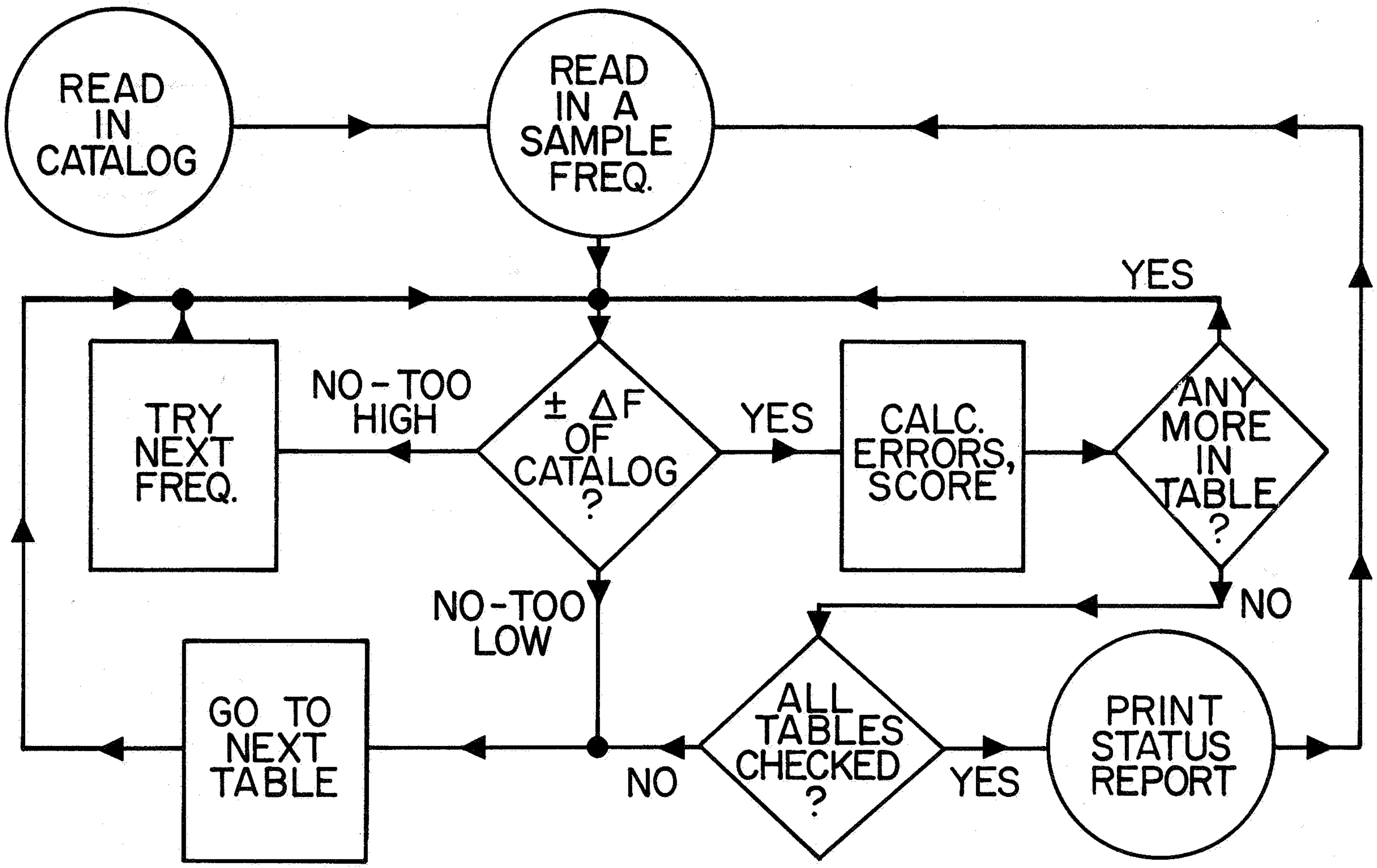
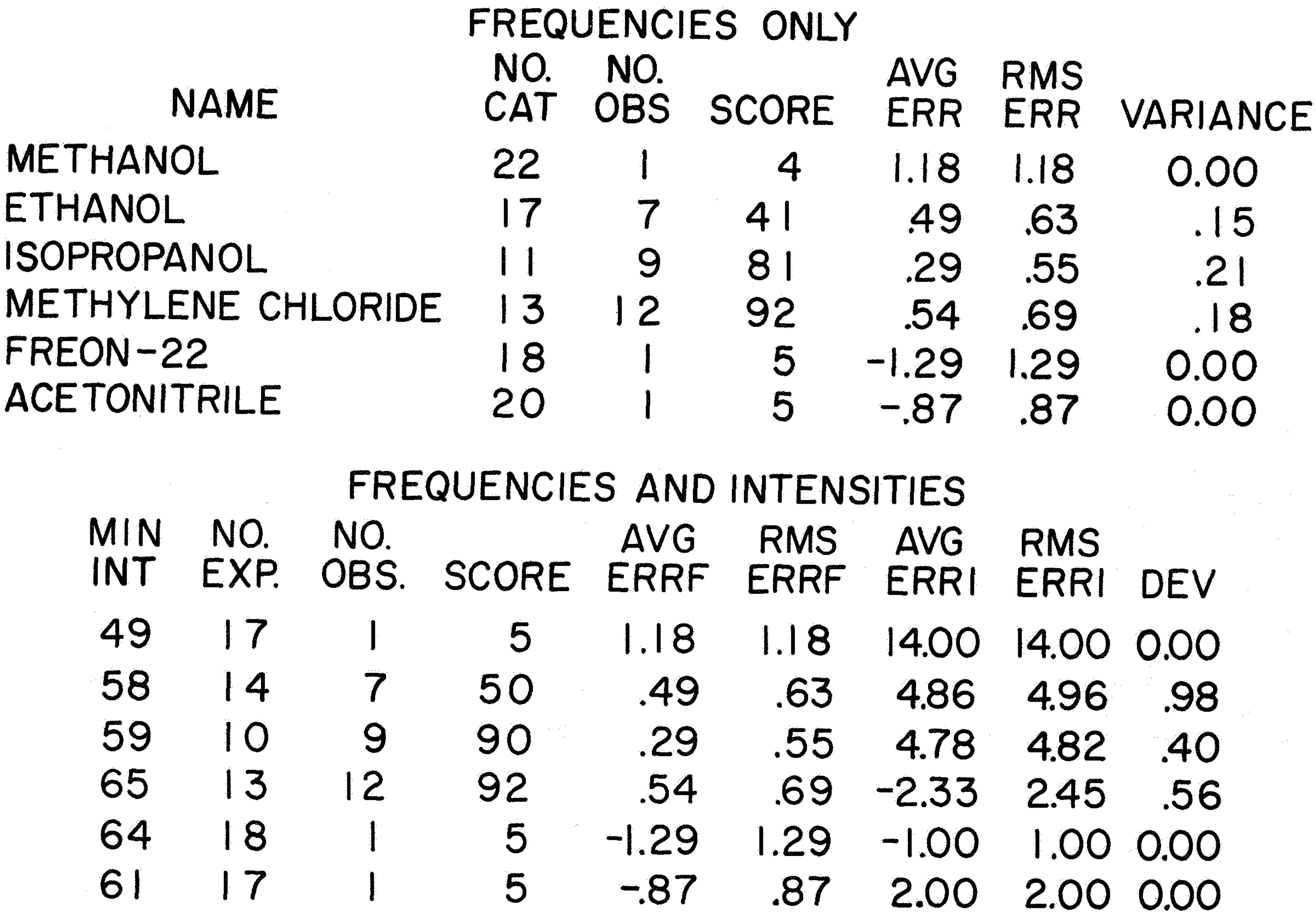


Figure 14.- Basic functions of data analysis program.



```
                      FREQUENCIES ONLY
                      NO.   NO.           AVG    RMS
     NAME             CAT   OBS   SCORE   ERR    ERR    VARIANCE
METHANOL              22    1       4     1.18   1.18     0.00
ETHANOL               17    7     41       .49    .63      .15
ISOPROPANOL           11    9     81       .29    .55      .21
METHYLENE CHLORIDE    13   12     92       .54    .69      .18
FREON-22              18    1       5    -1.29   1.29     0.00
ACETONITRILE          20    1       5     -.87    .87     0.00

                 FREQUENCIES AND INTENSITIES
MIN   NO.   NO.            AVG    RMS    AVG    RMS
INT   EXP.  OBS.  SCORE   ERRF   ERRF   ERRI   ERRI   DEV
49    17     1      5     1.18   1.18  14.00  14.00  0.00
58    14     7     50      .49    .63   4.86   4.96   .98
59    10     9     90      .29    .55   4.78   4.82   .40
65    13    12     92      .54    .69  -2.33   2.45   .56
64    18     1      5    -1.29   1.29  -1.00   1.00  0.00
61    17     1      5     -.87    .87   2.00   2.00  0.00
```

Figure 15.- Typical computer output for mixture of CH_3CH_2OH, CH_2Cl_2, and $CH_3CHOHCH_3$ in air.

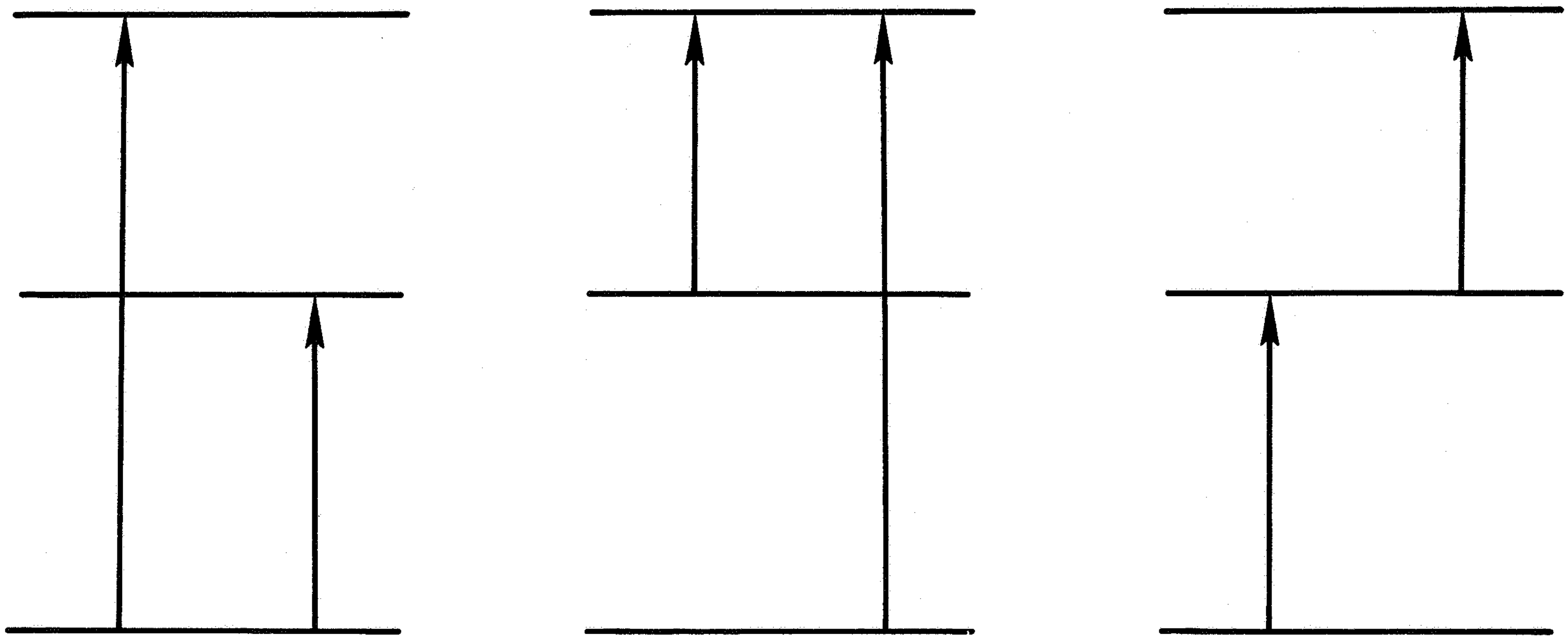


Figure 16.- Some three-level systems.

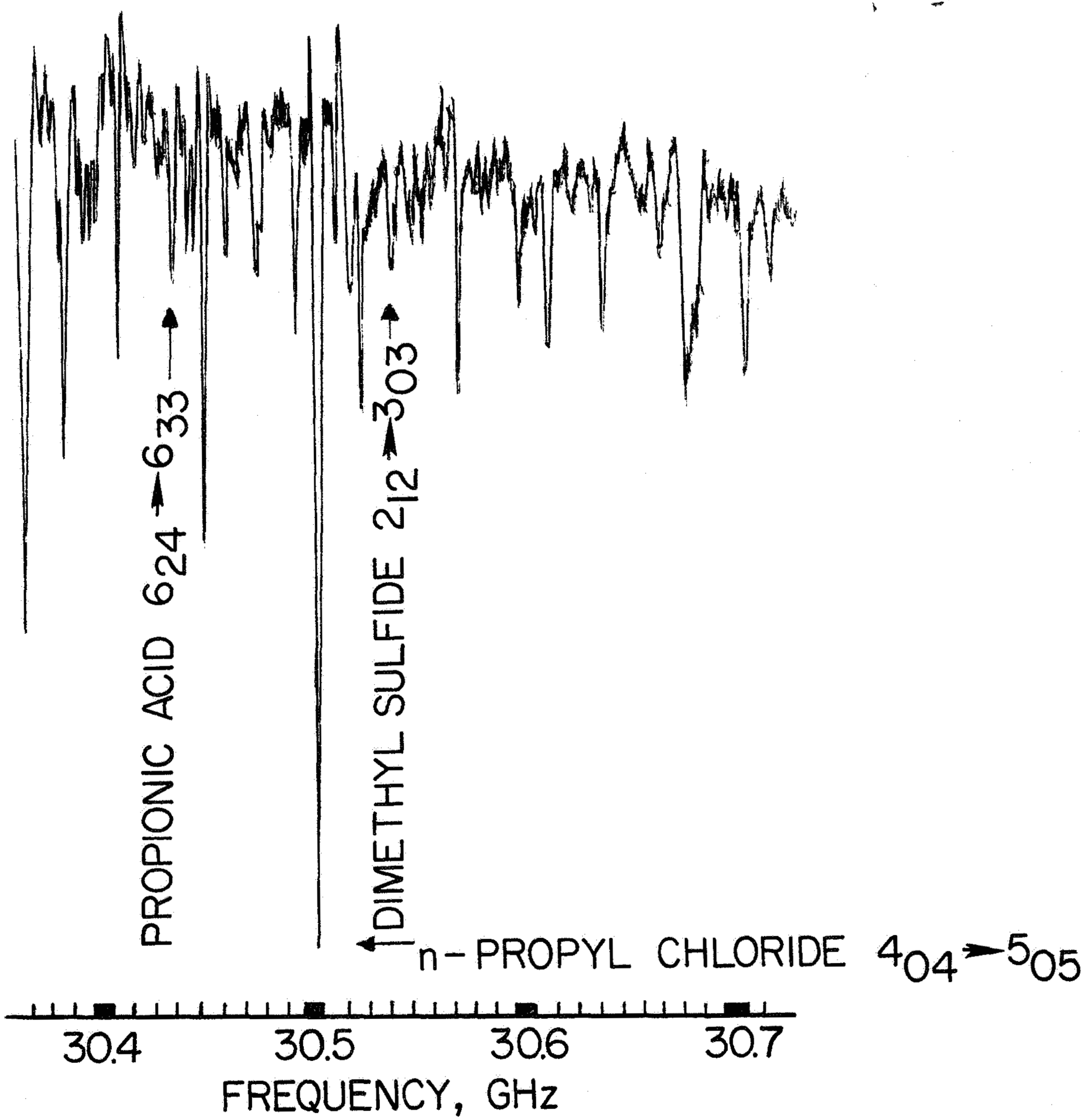


Figure 17.- Mixture of propionic acid, n-proply chloride, dimethyl sulfide, and air.

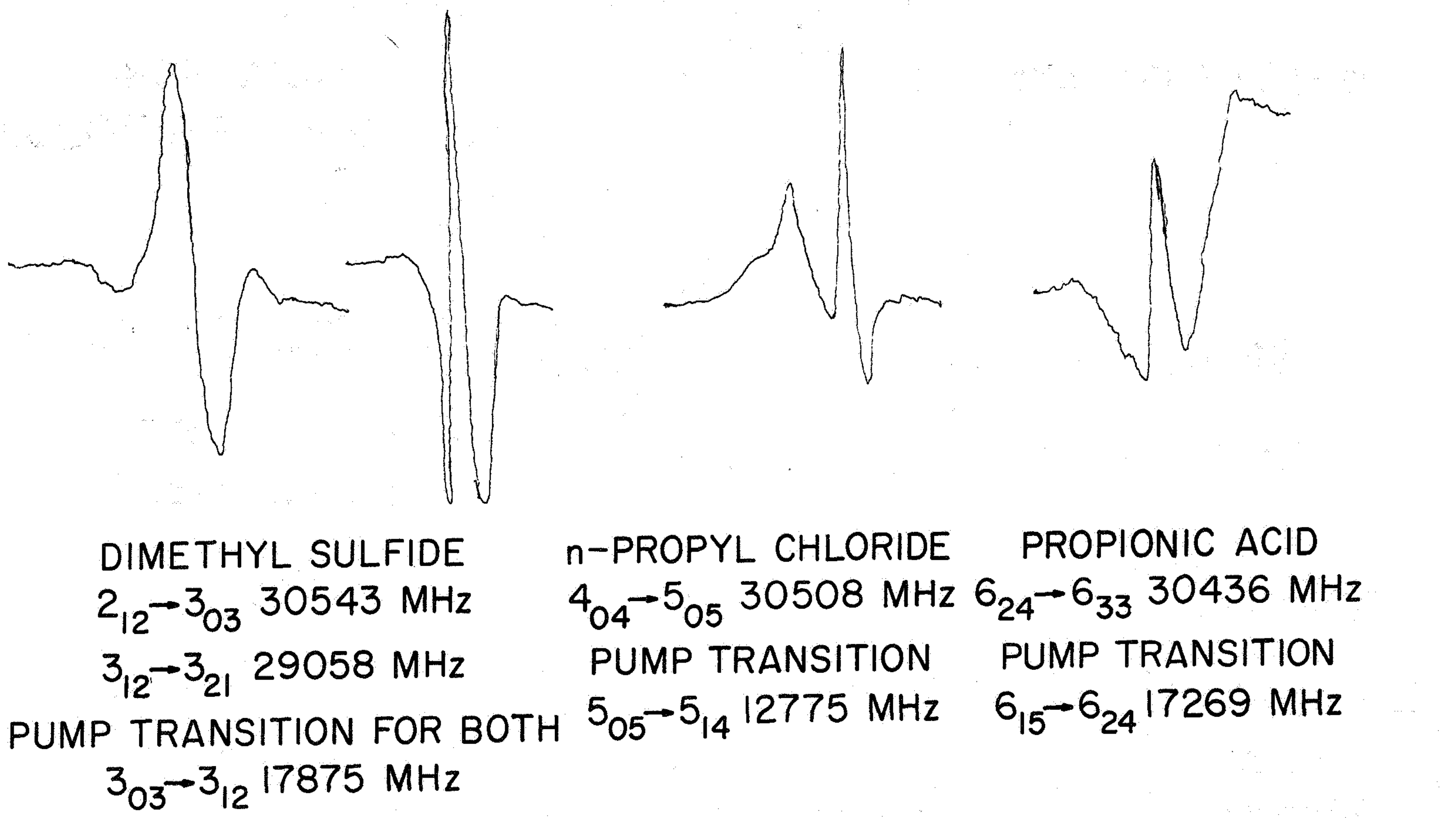


Figure 18.- Double resonance spectra of gas mixture in 100-foot cell.

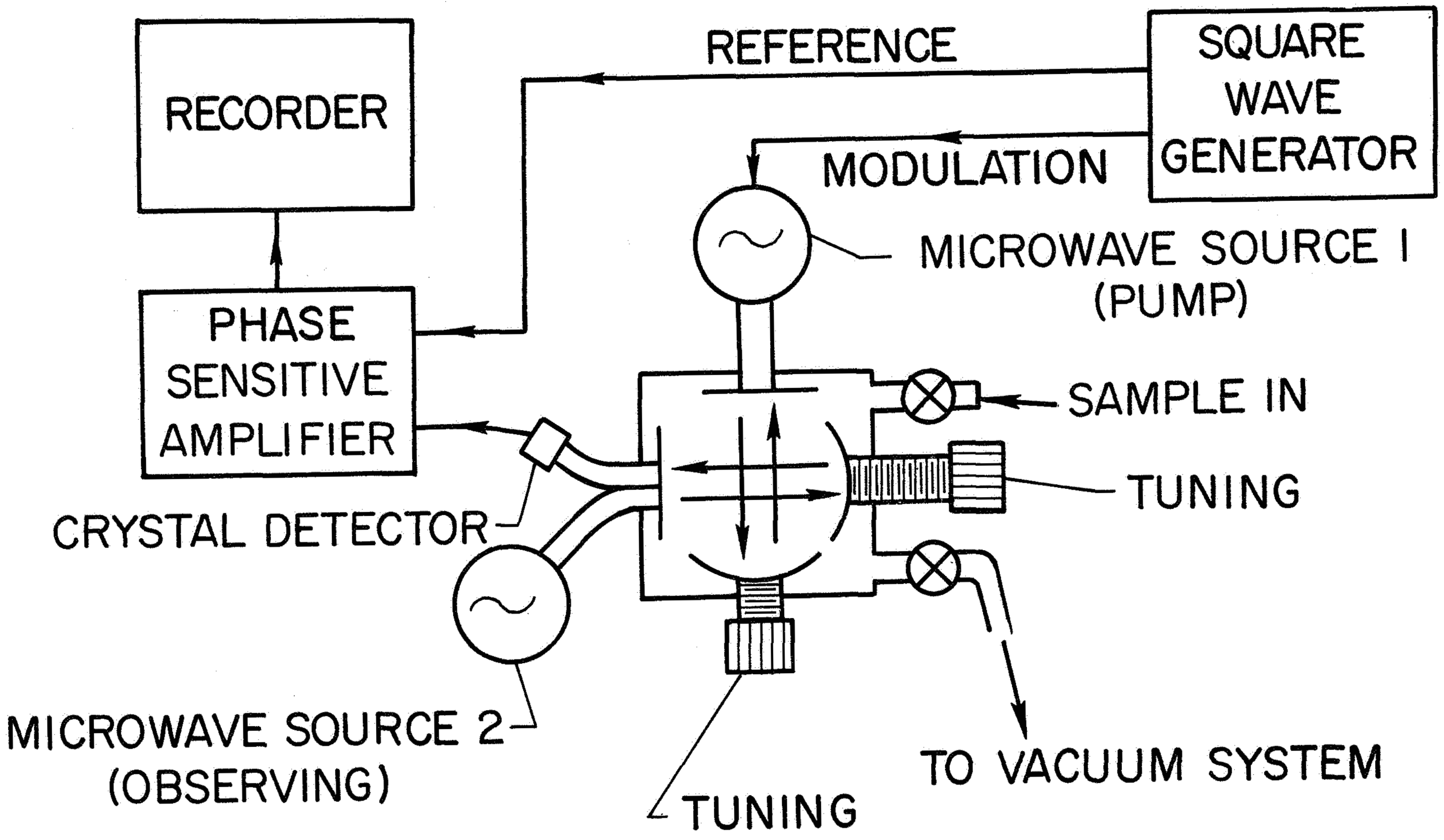


Figure 19.- Double resonance spectrometer.

NASA-Langley, 1968