Journal of Organometallic Chemistry, 52 (1973) 261–268

MONOISOTOPIC MASS SPECTRA

E. MCLAUGHLIN and R. W. ROZETT*

Chemistry Department, Fordham University, New York, N.Y. 10458 (*U.S.A.*)

(Received September 19th, 1972)

SUMMARY

A computer technique which calculates a least-squares fit monoisotopic mass spectrum, restricted to positive solutions, from polyisotopic measurements, is described. Any combination of elements, sets of isotopes or isotopic abundances can be handled. Partial or complete spectra may be analyzed. The method is also useful to determine the fractional abundance of isotopes, calculate the cluster of isotope peaks due to any formula, handle the data from ion-molecule studies, and detect impurities in mass spectral data. As examples the mass spectra of $(CH_3)_2Hg$, $(CH_3)_4Pb$, $SiCl_4$ and $SOCl_2$ are resolved.

INTRODUCTION

When one uses mass spectroscopy to analyze compounds containing highly polyisotopic elements, the resulting spectra are complex and it may be difficult to discover the formulas of the ions present and the intensity of each of these formulas. This is the case for bromine, chlorine and sulfur derivatives, as well as for boron and silicon compounds. But the most widespread area of application is found among the organometallics and other metal derivatives. The 10 naturally-occurring isotopes of tin, for example, are spread over 13 mass units. Cadmium isotopes occupy 11 masses. Mercury, platinum and molybdenum atoms range over 9 mass units. Zinc, iron, chromium, nickel, tungsten and lead isotopes are distributed over five or more mass units.

Monoisotopic mass spectra may be calculated manually, of course. A typical method begins with the intensity of the peak which occurs at the lowest mass. For the sake of simplicity we assume that this peak contains only the lightest isotope of each element. From statistical considerations which are discussed later, one can calculate the abundance of every ion with the same elemental formula, but heavier because of the presence of one or more heavier isotopes. These calculated intensities are then subtracted from the experimental measurements at the appropriate masses. In this way one removes from the experimental data the effect of the presence of ions with the formula in question, but containing heavier isotopes. The residual intensity at the second lowest experimental mass is now due to an ion containing only the lightest

---

* To whom correspondence should be addressed.

isotopes and one more hydrogen than the peak at the lowest mass. The procedure can be repeated until the effect of the presence of heavier isotopes in every formula is removed.

These computations are complicated and tedious. In fact, whenever the lowest or highest peak in the experimental spectrum is not monoisotopic, this straightforward method is not possible. But even if this conventional method is automatically performed by a computer, it is not satisfactory. Errors or inaccuracies which occur early in the procedure are cumulative. Because of the consecutive nature of the calculations, special weight is put on the measurements at lower masses. A third drawback is that one often cannot use all of the information available to generate the monoisotopic intensities. There are usually more polyisotopic abundances than monoisotopic intensities to be determined, and the above procedure does not take advantage of this fact. Finally, the results of the conventional technique are often unsatisfactory because of the presence of negative intensites among the calculated monoisotopic abundances. These negative abundances almost necessarily arise during the subtractions described above. They may be due to mismeasurements, the presence of impurities, the use of incorrect formulas or incorrect fractional abundances for the isotopes, as well as other reasons[1,2]. The conventional method for calculating monoisotopic mass spectra provides no objective procedure to be followed when these negative intensities arise.

Another procedure for calculating monoisotopic mass spectra avoids these pitfalls. It is a matrix method which solves for all the monoisotopic intensities simultaneously. This avoids error accumulations and the need for iterative calculations. All the available measurements are used to establish least-squares fit monoisotopic abundances. The monoisotopic intensities may be restricted to non-negative values easily. This automatic computer matrix method will be described.

METHOD

In the mass spectrum of a polyisotopic sample, the measured ion intensity at any mass, P_i, is the sum of the contributions of the different ionic formulas which have one or more isotopic variant at that mass, P_{ij}. These contributions may be expressed in terms of the intensity of the monoisotopic ion with the same elemental formula, m_j, as in eqn. (1).

$$P_i = \sum_j P_{ij} = \sum_j (r_{ij} \cdot m_j) \tag{1}$$

The matrix of coefficients, r_{ij}, can be evaluated from statistical considerations. The probability or fractional abundance, a, of the occurrence of any one configuration of isotopes with n atoms of an element which has m different isotopes, can be calculated from the multinomial distribution given in eqn. (2)[3]. The gross fraction of each isotope in the sample is defined as f_i, and n_i atoms of each isotope occur in the configuration.

$$a = n! \cdot f_1^{n_1} \cdot f_2^{n_2} \cdot \ldots f_m^{n_m} / (n_1! \cdot n_2! \cdot \ldots n_m!) \tag{2}$$

If there are only two isotopes, this reduces to the more familiar binomial distribution, and the ratios of the factorials become binomial coefficients. Since one is not directly concerned with the probability of a configuration, but with its probability relative to the occurrence of the monoisotopic form, r, we must divide the abundance of the configuration by the abundance of the monoisotopic variant with the same elemental formula, a_0, as in eqn. (3).

$$r = a/a_0 \qquad (3)$$

When several elements are present in the ion, a further complication arises. The pertinent relative abundance for the multielement ion is the compound relative abundance, r_c. This is the product of the relative abundances for each element, as in eqn. (4).

$$r_c = r_a \cdot r_b \cdot \cdots \qquad (4)$$

The relative probabilities are multiplied because the variation of isotopes in one element is independent of the occurrence of isotopes in the other elements[4,3]. A second complication occurs when several isotopic variants of the multielement ion are found at the same mass. If the first variant has a relative abundance r_c, the second r'_c and so forth, then the total relative abundance of formula j at mass i, r_{ij}, is the sum of all of the appropriate compound abundances, eqn. (5).

$$r_{ij} = r_c + r'_c + \cdots \qquad (5)$$

In this case the respective probabilities are interdependent or mutually exclusive, as statisticians put it[5].

An example of the use of these formulas is presented in Table 1. The twelve isotopic variants of the ion C_2NCl occur at six different masses, 73 to 78. Each configuration of isotopes is described by a different set of the integers, k, 1, m, n, o, and p. They represent the number of atoms of the isotopes ^{12}C, ^{13}C, ^{14}N, ^{15}N, ^{35}Cl, and ^{37}Cl respectively. In the table the lightest isotope of each element has been chosen as the monoisotopic reference. A number calculated by the use of eqns. (2) and (3) occurs in Table 1 in the row labelled mass 78 under the column heading $r(C_2)$. The abundance of $^{13}C_2$ is reported there as 1.2599×10^{-4}. This was calculated by using eqn. (2) for $^{13}C_2$ ($n=2, n_1=0, n_2=2$), and also for $^{12}C_2$ ($n=2, n_1=2, n_2=0$) using the natural abundance of ^{12}C and ^{13}C (1.11%). Then the former number was divided by the latter according to eqn. (3). Use of eqn. (4) is illustrated in the same row of the table, mass 78. The relative abundance of $^{13}C_2{}^{15}N^{37}Cl$ is listed under the column heading $r(C_2NCl)$ as 1.5159×10^{-7}. It is the product of the relative abundance of $^{13}C_2$, 1.2599×10^{-4}, of ^{15}N, 3.7137×10^{-3}, and of ^{37}Cl, 3.2398×10^{-1}. Each of these three numbers were in turn calculated from eqns. (2) and (3) in the manner described above. Finally, eqn. (5) is exemplified by the two rows labelled mass 77 in Table 1. In the first row marked 77, the relative abundance of $^{12}C^{13}C^{15}N^{37}Cl$ is calculated as 2.7010×10^{-5}. In the second row 77, the relative abundance of $^{13}C_2{}^{14}N^{37}Cl$ is reported as 4.04818×10^{-5}. The relative abundance of all isotopic variants of C_2NCl which occur at mass 77 is the sum of these two abundances, or 6.7828×10^{-5}.

The ultimate result of the calculations for the ion C_2NCl are the six numbers listed in the column of the table marked r_{ij}. When these numbers are multiplied by the monoisotopic intensity of C_2NCl, we obtain the polyisotopic contributions of this ion at the six different masses at which it occurs.

In the second stage of the procedure to generate monoisotopic mass spectral intensities, one must solve the simultaneous equations (1), whose coefficients we have just calculated. There is one equation for each experimental intensity measured, and one unknown for each monoisotopic formula. A least-squares matrix technique which avoids the usual difficulties has been described in a recent complication of the mono-

TABLE I

THE TWELVE ISOTOPIC VARIANTS OF C_2NCl

Mass	$^{12}C_k$ k	$^{13}C_l$ l	$^{14}N_m$ m	$^{15}N_n$ n	$^{35}Cl_o$ o	$^{37}Cl_p$ p	$r(C_2)$	$r(N)$	$r(Cl)$	$r(C_2NCl)$	r_{ij}
73	2	0	1	0	1	0	1.0000	1.0000	1.0000	1.0000	1.0000
74	2	0	0	1	1	0	1.0000	3.7137×10^{-3}	1.0000	3.7137×10^{-3}	
74	1	1	1	0	1	0	2.2449×10^{-2}	1.0000	1.0000	2.2449×10^{-2}	2.6163×10^{-2}
75	2	0	1	0	0	1	1.0000	1.0000	3.2398×10^{-1}	3.2398×10^{-1}	
75	1	1	0	1	1	0	2.2449×10^{-2}	3.7137×10^{-3}	1.0000	8.3369×10^{-5}	
75	0	2	1	0	1	0	1.2599×10^{-4}	1.0000	1.0000	1.2599×10^{-4}	3.2419×10^{-1}
76	2	0	0	1	0	1	1.0000	3.7137×10^{-3}	3.2398×10^{-1}	1.2032×10^{-3}	
76	1	1	1	0	0	1	2.2449×10^{-2}	1.0000	3.2398×10^{-1}	7.2730×10^{-3}	
76	0	2	0	1	1	0	1.2599×10^{-4}	3.7137×10^{-3}	1.0000	4.6789×10^{-7}	8.4766×10^{-3}
77	1	1	0	1	0	1	2.2449×10^{-2}	3.7137×10^{-3}	3.2398×10^{-1}	2.7010×10^{-5}	
77	0	2	1	0	0	1	1.2599×10^{-4}	1.0000	3.2398×10^{-1}	4.0818×10^{-5}	6.7828×10^{-5}
78	0	2	0	1	0	1	1.2599×10^{-4}	3.7137×10^{-3}	3.2398×10^{-1}	1.5159×10^{-7}	1.5159×10^{-7}

isotopic mass spectra of the boranes[1]. The basic program used, called MIMS, or Mono Isotopic Mass Spectra, is a drastic revision and generalization of the previously published programs for boranes and boron derivatives. The power and accuracy of the general technique has been amply demonstrated with the boranes. It provided a fit about twenty-five times as good as the conventional method for the boranes[1].

APPLICATIONS AND EXAMPLES

As the computer program, MIMS, is now set up, it can handle up to 10 different elements in any one mass spectrum. Each element can have isotopes which extend over

TABLE 2

THE MONOISOTOPIC MASS SPECTRUM OF TETRAMETHYLLEAD

Mass	*Formula*	*Intensity*
208	Pb	63.97
209	PbH	5.77
220	PbC	0.85
221	PbCH	2.78
222	$PbCH_2$	6.13
223	$PbCH_3$	85.28
224	$PbCH_4$	0.37
232	PbC_2	0.04
233	PbC_2H	0.13
235	PbC_2H_3	0.05
237	PbC_2H_5	0.56
238	PbC_2H_6	24.82
239	PbC_2H_7	0.24
252	PbC_3H_8	1.11
253	PbC_3H_9	100.00
254	PbC_3H_{10}	0.22
	RMD	0.030

TABLE 3

THE MONOISOTOPIC MASS SPECTRUM OF DIMETHYL MERCURY

Mass	*Formula*	*Intensity*
202	Hg	41.99
203	HgH	4.81
214	HgC	2.75
215	HgCH	6.11
216	$HgCH_2$	14.03
217	$HgCH_3$	100.00
218	$HgCH_4$	0.13
219	$HgCH_5$	0.06
229	HgC_2H_3	0.20
230	HgC_2H_4	0.59
231	HgC_2H_5	0.73
232	HgC_2H_6	40.73
	RMD	0.071

13 mass units. This is sufficient for even the most polyisotopic elements, tin and xenon. Any isotope may be designated as the monoisotopic intensity, but generally the most abundant isotope is chosen. A quantitative measure of the fit of the least-squares solution, the root-mean-square deviation, is automatically calculated. The polyisotopic spectrum may be handled completely, or truncated in any fashion. The solution is completely accurate in the range designated if all of the appropriate formulas can be included. This feature is especially helpful if a computer with small storage is available. The largest demand on core results from the matrix of coefficients of eqn. (1) which has the dimensions Number of polyisotopic measurements × Number of monoisotopic formulas. Truncation allows one to factor the spectrum into several small problems. The program for $Hg(CH_3)_2$ in Table 3 occupied less than 40 K bytes on an IBM 360 Model 40 (*i.e.* 10 K variables).

Some examples of the results of MIMS are listed in the accompanying tables and figures. Tetramethyllead and dimethylmercury were chosen to illustrate the use of the program with elements which have many isotopes[6]. The compounds $SiCl_4$ and $SOCl_2$ exemplify its application to sulfur, the halogens and silicon[7].

The input needed for the program consists of the fractional abundance of the isotopes of the elements present, the experimental spectrum and a set of formulas. If there is any uncertainty about the formulas of the ions present, all may be tried and the best-fitting set chosen by comparison of the root-mean-square deviation, RMD, for each set. Naturally, if the abundance of a second isotope is small, as for the elements hydrogen, carbon and oxygen, this procedure may be unreliable. A parent-mass-plus-

TABLE 4

THE EXPERIMENTAL POLYISOTOPIC MASS SPECTRUM, MONOISOTOPIC MASS SPECTRUM AND RECALCULATED POLYISOTOPIC MASS SPECTRUM OF THIONYL CHLORIDE

Mass	*Monoisotopic formula*	*Monoisotopic intensity*	*Recalculated polyisotopic intensity*	*Experimental polyisotopic intensity*
83	SOCl	97.74	97.70	97.7
84			0.82	0.9
85			36.19	36.3
86			0.27	0.3
87			1.49	1.5
88			0.0	0.0
89			0.0	0.0
118	$SOCl_2$	9.80	9.79	9.8
119			0.08	0.3[a]
120			6.80	6.8
121			0.05	0.0
122			1.32	1.3
123			0.01	0.0
124			0.05	0.0
125			0.0	0.0
126			0.0	0.3[a]
RMD		0.06		

[a] Impurities suspected.

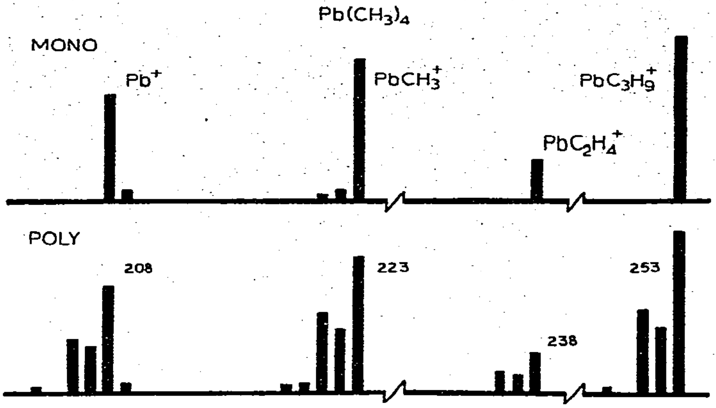

Fig. 1. The polyisotopic and monoisotopic mass spectrum of tetramethyllead.

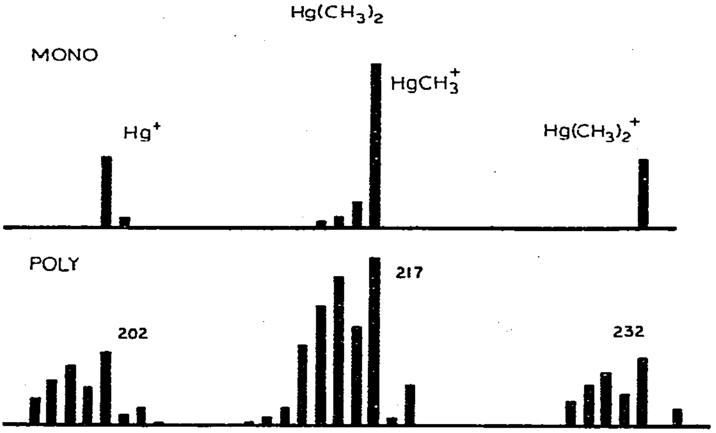

Fig. 2. The polyisotopic mass spectrum and monoisotopic mass spectrum of dimethylmercury.

one formula may be included if needed. When negative monoisotopic intensities occur, they may be eliminated by removing the formulas of the peaks in question.

Some sources of error and limitations have been described previously[1,2]. It is well to recall that some polyisotopic spectra cannot be resolved into monoisotopic spectra by any method. This occurs whenever there are more ions present than there are measurements to establish their abundance. This is found most frequently in the mass analysis of a mixture, although it may also occur with a single complex compound. In the latter case, however, it is always possible to resolve the spectrum in the region near the parent peak.

A listing of the program in Fortran IV can be obtained from the author (RWR). Arrangements can be made through the author to obtain copies of the source deck from the Fordham University Computing Center. The program will also be filed with the American Society for Mass Spectrometry.

TABLE 5

THE EXPERIMENTAL POLYISOTOPIC MASS SPECTRUM, MONOISOTOPIC MASS SPECTRUM AND RECALCULATED POLYISOTOPIC MASS SPECTRUM OF SILICON TETRACHLORIDE

Mass	*Monoisotopic formula*	*Monoisotopic intensity*	*Recalculated polyisotopic intensity*	*Experimental polyisotopic intensity*
98	$SiCl_2$	6.94	6.90	6.8
99			0.35	0.4
100			4.71	4.8
101			0.23	0.4
102			0.87	0.9
103			0.04	0.1
104			0.02	0.0
133	$SiCl_3$	100.00	99.46	100.0
134			5.07	5.2
135			100.00	99.2
136			4.93	5.0
137			34.56	33.6
138			1.60	1.6
139			4.43	4.3
140			0.17	0.2
141			0.11	0.1
168	$SiCl_4$	41.09	40.87	41.3
169			2.08	2.1
170			54.33	54.0
171			2.70	2.7
172			27.51	27.0
173			1.31	1.3
174			6.42	6.3
175			0.28	0.3
176			0.64	0.6
177			0.02	0.1
178			0.02	0.1
	RMD	0.17		

ACKNOWLEDGEMENTS

This research was supported in part by the Petroleum Research Fund administered by the American Chemical Society (PRF 1233–G2). The instrumentation was supplied in part by the National Science Foundation (GP 8220) and by the New York State Science and Technology Foundation (NYSSF (6)-13).

REFERENCES

1 E. McLaughlin, T. E. Ong and R. W. Rozett, *J. Phys. Chem.*, 75 (1971) 3106.
2 E. McLaughlin and R. W. Rozett, *J. Phys. Chem.*, 76 (1972) 1860.
3 W. Feller, *Probability Theory and Its Applications*, Wiley, New York, 1966, p. 157.
4 Ref. 3, p. 114.
5 Ref. 3, p. 22.
6 *Mass Spectral Data*, American Petroleum Institute Research Project 44, serial nos. 698 and 700.
7 E. Stenhagen, S. Abrahamsson and F. W. McLafferty, *Atlas of Mass Spectral Data*, Interscience, New York, 1969, pp. 445, 1093.