

## Computer programs for the analysis of PDMS-spectra

J. Curdes and E. R. Hilf

Fachbereich Physik, Universität Oldenburg, Postfach 2503, W-2900 Oldenburg, Federal Republic of Germany

Received May 7, 1992; revised July 3, 1992

**Summary.** We will present computer programs to analyze PDMS-spectra and to simulate the fragmentation of molecules in PDMS (Plasma-Desorption Mass Spectrometry), even of those with high masses.

### Introduction to PDMS

For those who are not familiar with the principles of this TOF-technique, we give some short explanations here.

In PDMS a solid surface is imposed to the impact of fast, heavy ions in the range of some MeV/nucleon. The very high energy deposited in this process (compared to other desorption techniques) results in an explosive process where even very large or thermally instable molecules are desorbed from the surface before acquiring thermal energy. The yield of desorbed molecules (or fragments thereof) is in the range of  $10^4$  per incoming primary ion, but only some 10 are charged and can thus be analyzed using time-of-flight techniques.

As a source for the incoming primary ions the radioactive  $^{252}\text{Cf}$  is most commonly used. This isotope undergoes spontaneous fission resulting in two fragments of comparable (but not necessary equal) mass and energy. The geometry of the instrument, see Fig. 1, is designed in such a way that one of the fragments is hitting a start detector at the same time when the other hits the sample of the substance in question. Molecules, clusters of molecules and fragments of molecules are desorbed, some of which are charged and can thus be accelerated with a high voltage (10 kV) and then drift along a force-free path (10–100 cm) to a stop detector. Thus a time difference is gained from which the mass of the fragment or molecule can be easily calculated, only two apparatus constants need to be known.

The time differences are saved into typically 65000 "time bins" so that a channel spectrum is the result of a measurement. Such a spectrum is a sum over many events (the fission rate of  $^{252}\text{Cf}$  is about 1000/s) so that it roughly represents the probabilities for the formation of a fragment or cluster with a certain mass. A good introduction to the method of PDMS can be found in [1].

### Method

The advantages of PDMS – in comparison to other well known techniques – are lying mainly in the possibility to measure substances with high masses (substances with an atomic weight  $> 50000$  amu have been measured) and in the high yield for the molecular peak. Obviously it is an almost unique method for the analysis of heavy, unstable organic molecules. The characteristics mentioned above are based on the high and sudden energy deposition  $dE/dx$ , see Fig. 2, so that the incoming primary ion sets free some 40 electrons/Å on its path through the solid.

Fig. 3 gives an example for a spectrum gained with this method. This spectrum has been recorded by F. Hillmann and W. Tuszynski [2]. The substance is a metal complex [bromo-bis(dimethylglyoximato)-pyridino-cobalt(III)-complex] with a molecular mass of 448 amu, from which – to the best of our knowledge – other than this PDMS spectrum is not available. The molecular peak can be seen clearly as the second highest peak in the whole spectrum, thus clearly

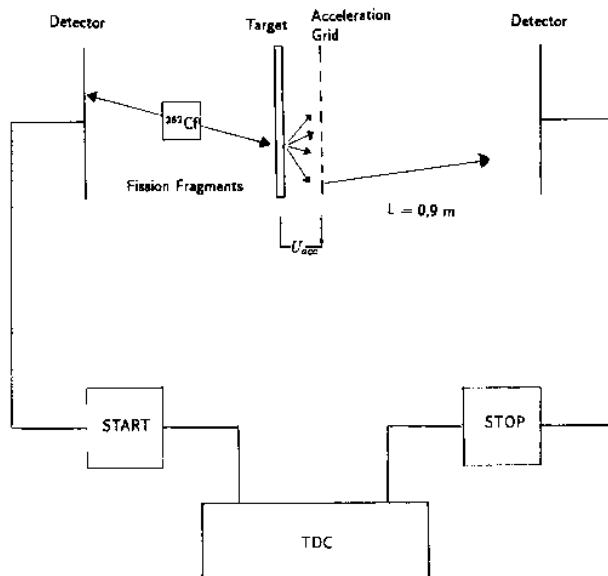


Fig. 1. Principal components of a PDMS spectrometer

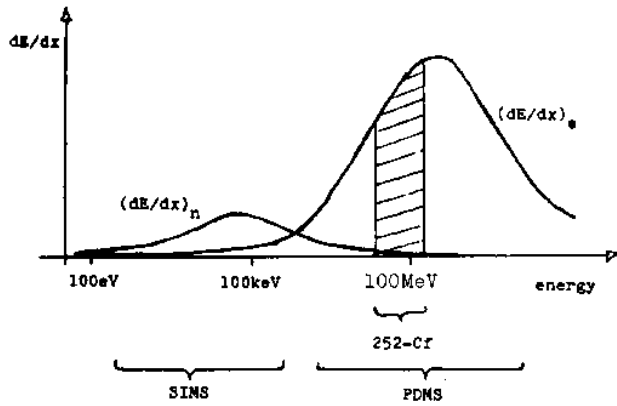


Fig. 2. Energy deposition rate of PDMS and SIMS [4]

demonstrating the characteristics mentioned above. If we take a closer view, see Fig. 4, at the region of this peak, the signals at  $m/z$  369 and 289 can be easily explained by assuming the loss of bromine resp. bromine and pyridine. The series of signals at  $m/z$  254, 335, 428, (...) are clusters formed from molecular fragments which will be not discussed further here.

**Peaks**

The huge amount of unprocessed data which forms a PDMS spectrum (typically 250 kB/spectrum) gave rise to the idea of a more or less automatic method of data reduction and analysis.

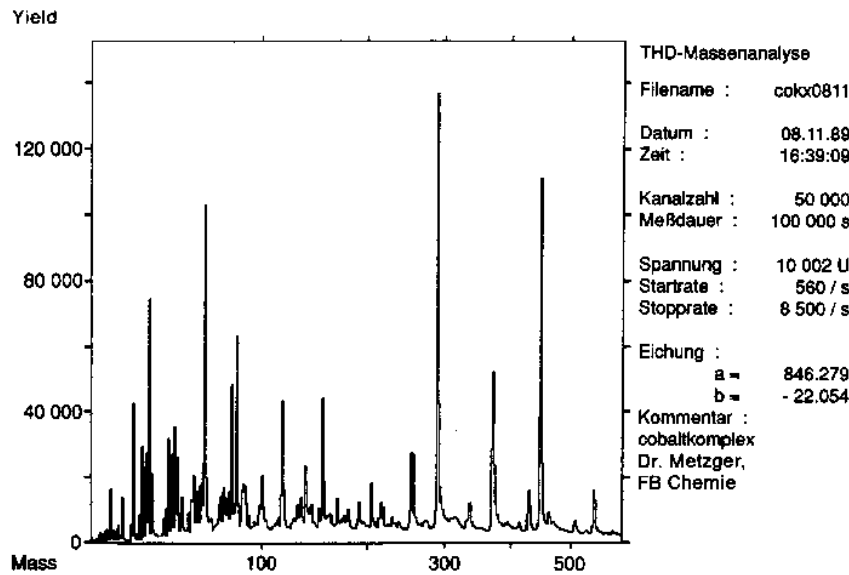


Fig. 3. PDMS-spectrum of a metal complex

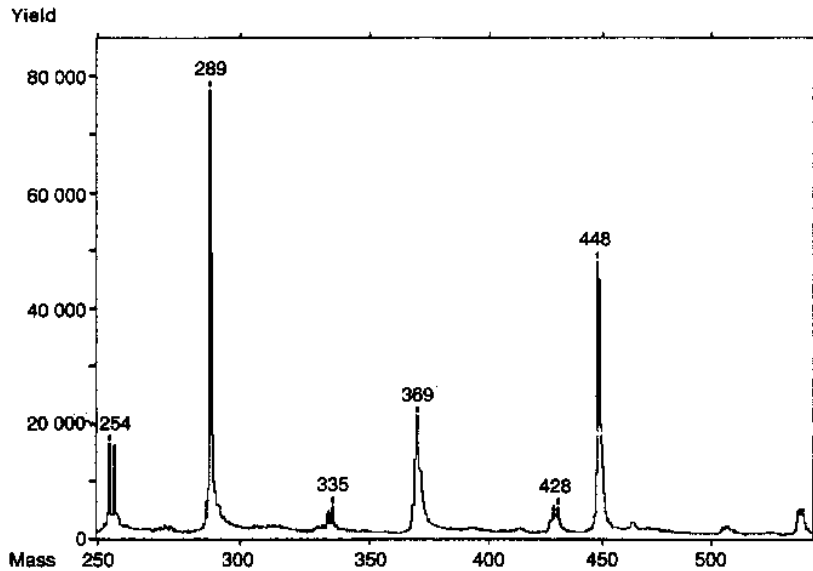


Fig. 4. Cobalt complex: region of molecular ion peak

The program PEAKS [3] developed by W. Schlez in our group extracts the relevant peaks from the noisy spectrum using a remarkably efficient algorithm. Mathematically speaking, it calculates *weighted sums*, also called *moments*

$$M_k(C_i) = \sum_{C_j \in Y} \xi^{jk} \cdot Y(j)$$

where  $M_k$  is the resulting  $k$ -th Moment calculated over the range  $C_i$  of the spectrum  $Y$ , whose channels are denominated by  $Y(i)$ .

The weight factor  $\xi$  must be chosen appropriately according to the given problem to gain the meaningful properties as moments. In our case it is chosen to be

$$\xi = (i - i_0)$$

From a reference point  $i_0$  the distances of the other channels  $i$  from this reference channel are calculated.

This set of moments is called *close - moments* because it is necessary to make an assumption for  $i_0$ , in other words,  $i_0$  must be "close" enough to the peak we are looking for. This is one of the reasons why it appears difficult to automatize this calculation process. However, if we make the contrary - assumption that our reference point  $i_0$  lies well outside the peak that we are looking for, or

$$M'_k(C_i) = \sum_{C_j \in Y} (i - i_1)^k \cdot Y(j)$$

advantageously although  $i_1$  is arbitrary, surprisingly it is still possible to calculate a mean peak position, etc. This is what we were looking for: a method for automatic scanning of a spectrum giving us the relevant mass line informations. Another advantage is that far moments give better performance in CPU-time.

These moments are calculated along the spectrum over a selected range to extract mass line information such as amplitude, position, width, asymmetry.

The interesting information can then be calculated as following skipping the dash for now,

$$A = M_0 \cdot w / \sqrt{2} \cdot \pi$$

$$t = M_1 / M_0 + i_1$$

$$w = \sqrt{M_2 / M_0 - i_1^2}$$

$$\gamma = \frac{M_3 / M_0 - i_1 \cdot M_2 \cdot (3 - i_1^2 \cdot M_0 / M_2)}{M_0 / b^3}$$

where  $A$  represents the calculated amplitude,  $t$  the center position of the peak,  $w$  the calculated line width and  $\gamma$  the asymmetry. For the foundations of this method see [4], [5].

Here we give an example of the peak position of the shown distribution as the program scans through the spectrum, see Fig. 5.  $M_0$  is the (constant) sum over the shown distribution and  $i_1$  is varied from channel # 2 to # 8. The resulting  $M_1$  changes with the starting point, but the calculated position  $t$  does not, as long as it remains well outside the peak itself.

The example shows that indeed the calculated position is almost independent from the starting point  $i_1$ .

Naturally there remain problems such as how to decide if there is something like a peak at all or how to choose the starting point. In fact, in our case the asymmetry is used as a criterium for finding a "peak": if there is a peak, however shaped, somewhere "in" this peak the calculated asymmetry will change the sign while the reference point  $i_1$  is shifted

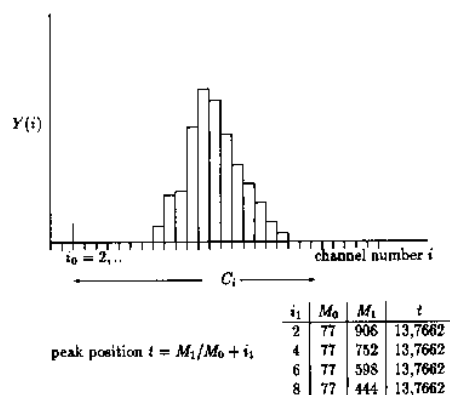


Fig. 5. Example: calculation of weighted peak position  $t$  (channels)

through the pattern. This change of sign is detected and the quantities calculated for the region are recorded as quantities of the peak.

The advantage of this method - compared to e.g. a Gauss-fit - is that there are no parameters to be fitted - the spectrum is scanned one time only with the selected "window". Only in this way the huge amounts of data can be analyzed in competitive time even on a 286-PC! In addition, it is very sensitive even for small peaks in noisy spectra because no shape is assumed [6].

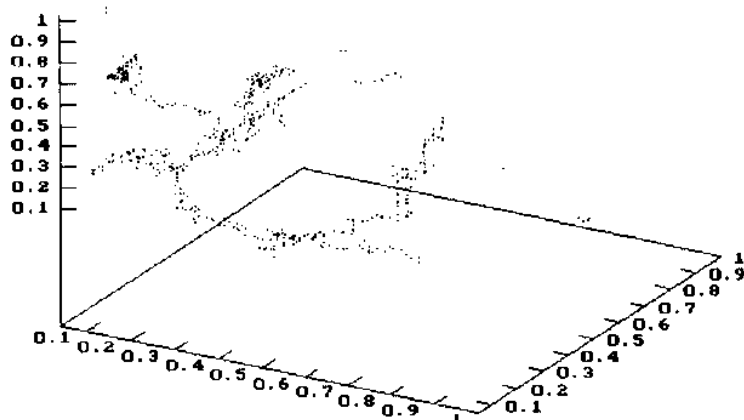
## CRUNCHER

On the way to understand the experimental spectra of pure substances in our group a powerful algorithm ("CRUNCHER") [7] was developed for the entropic breaking of the - known - chemical structure. This leads to a spectrum which is similar to the experimental spectrum for the dominant lines.

Entropic means here that the deposition of energy to the bond takes only a very short time as compared to bond oscillations. Thus a thermal equilibration of energy in the molecular structure is not possible and hence the breaking probabilities of the different bonds should be quite independent from each other.

The problem was to define an algorithm which is able to calculate fragmentation from a given bond structure of arbitrary type, with no (principal) restrictions on the shape or size of the molecule concerned. This was done by describing the atoms or components and their bonds as a list containing mainly pointers to other nodes as a memory for the structure. Using this list, an algorithm that we call "pacman" due to its winding its way along the bond-structure of the molecule is created. It includes calculations of possible fragmentation results and their probabilities. During this process of calculation it is necessary to keep in memory the parts of the molecule which have already been visited. All this has been embedded in the necessary surroundings such as bond list editors, element and principal component lists, graphic spectrum viewer etc.

At the moment we are running fits where we try to select the breaking probabilities associated with the different bonds in such a way that the main lines of the experimental



**Fig. 6.** Similarity of a real spectrum of a cobalt complex to the prediction by the code CRUNCHER versus breaking probabilities of the Co-pyridine and the Co-Br bonds

spectrum are represented well in position and amplitude. After gaining some experience with the algorithm we will now have to include the isotope distribution of elements and to develop an efficient fit which explores the multi-dimensional (typically 5..15 bond types) space of the possible values in an acceptable time.

After discussions with the field theory specialists in our group where complex parameter spaces are a well-known problem, we have chosen the following fitting method: we selected a random set of probabilities as starting point in the bond probability space followed by the search for a local maximum of the similarity (exp./theor. spectrum). The stepwidth for variation of the parameters is determined by the "history" of the local search, that is, by the slope.

During our work, we realized that to gain enough information about the parameter space one needs many calculation steps each one taking a lot of time – even with our pacman code –, depending on the size of the molecule and the accuracy desired. So we decided that PC-performance will not meet the requirements of our calculations and we are translating the code into C language for a fast implementation on a 20 MFlop workstation. Nevertheless, we have gained some results:

For the metal complex already mentioned above the results of a fit as described above are given in Fig. 6. Clearly

the 'similarity surface' is highly nonlinear with several side-minima. To pick the most suited one it is wise to integrate the expertise of the respective customer.

#### References

1. Wien K (1987) In: Hilf ER, Kammer F, Wien K (eds) PDMS and Clusters, Lect Notes Physics, vol. 269. Springer, Berlin Heidelberg New York, p 1
2. Hillmann F (1991) Aufbau und Erprobung eines PDMS-Spektrometers. Diplomarbeit, Universität Oldenburg
3. Harsdorf S, Hilf ER, Nitzschmann B, Schlez W (1990) Computer Simulations, Analysis and Transfer of PDMS-Spectra. In: Hilf ER, Tuszynski W (eds) Mass Spectrometry of Large Non-Volatile Molecules for Marine Organic Chemistry. World Scientific, p 164; and K. Kruse, A. Dullweber, and M. Hilf (unpublished 1990)
4. Häertter E (1987) Wahrscheinlichkeitsrechnung, Statistik und mathematische Grundlagen. Vandenhoeck u Ruprecht, p 57
5. Wendel MH, Hilf ER (1985) J Math Phys 26:1570
6. Hilf ER, Schlez W, Kruse K, Dullweber A, Steenen U, Harsdorf S, Koch K, Nitzschmann B, Kammer F, Tuszguski W (1991) Chemometry of plasma desorption mass spectrometry. In: Grmehling J (ed) Software development in Chemistry 5. Springer, Berlin Heidelberg New York, pp 129–138
7. Hilf ER (1990) Chemometrie von PDMS-Spektren. Diskussions-tagung Computereinsatz in der Spektroskopie, Gemen