# Analysis of experimental spectra using spreadsheets

by
M. Carlsson Göthe, L. Karlsson and S. Svensson
Uppsala University, Dept. of Physics
P.O.Box 530, S-751 21 Uppsala,
Sweden

and

J. de Sousa Pires
Apple Computer AB
P.O.Box 31, S-164 93 Kista-Stockholm
Sweden

UUIP-1237 Aug. 1990

**Abstract**

Various steps in the analysis of experimental spectra are discussed. These procedures include data acquisition, pre-analysis, simulation, modelling and presentation. A dedicated pre-analysis program, called CrunchViewer, is presented. The usefulness of spreadsheet programs, here exemplified with Microsoft Excel for the Macintosh, in simulation and modelling is demonstrated in applications related to photoelectron spectroscopy of small molecules. An example of a photoelectron spectrum prepared for publication by these programs is given.

## Introduction

Experimental physics may be divided into two activities; instrumental development and systematic measurements aiming at a determination of fundamental properties of nature. The former activity generates intruments, including mechanics, electronics and methods, to obtain high quality results. The latter uses these results in order to record data. Comparing the data with the present theories may validate the models used in current research. The interpretations often results in new questions requiring refined instrumental development. In our group, both instrumental development and systematic measurements are performed [1,2].

The handling of data, or in our case electron- and ion spectra [1], includes several phases. In the data acquisition phase, the spectrum is recorded and a preliminary evaluation is performed. In the pre-analysis phase, the spectrum is viewed, data is extracted and prepared for further processing. In the modelling and simulation phase the data is fitted to the physical models. Finally, in the presentation phase, the spectrum, and additional results, are prepared for publication. User friendly computer support is essential, both within each phase and when linking the various phases together. The data must easily flow in both directions without the need of laborious format conversions as reevaluations are usually necessary.

Modelling and simulation have become important tools, which complement the traditional experimental and theoretical research methods in basic and applied science. The high level, user friendly spreadsheet programs simplify so much the use of computational resources that modelling and simulation can be carried out on a routine like basis even by non-specialists in programming. Examples of application of such spreadsheet programs, in particular Microsoft Excel for the Macintosh, for research in atomic and molecular physics will be presented. A reference source about the use of spreadsheets in about any conceivable disciplin is illustrated in ref. [3]. This handbook also delves deeply into Computer Pedagogics and Innovative Problem Solving using a personal computer.

## Data acquisition

A control and data acquisition system based on a PC/AT-compatible computer, is presently in use in the laboratory [1]. This system has been operative since 1987. The data acquisition program is entirely written in Turbo Pascal. Identical programs and computer systems are used at four different electron spectrometers in the laboratory. The main program for the spectrometer control contains routines for defining spectrum acquisition parameters, setting and stepping voltages, collecting and displaying data and saving them to disk, displaying earlier stored spectra, scaling and moving spectra and investigating the contents of individual channels. A set of auxiliary programs is used for such purposes as generation of correction tables for the detector energy scale, redefining such parameters as the calibration constants for the voltage supplies, measuring the same constants with an online

precision voltmeter, and optimization of voltage settings with respect to resolution and intensity. All communication between the AT and the peripheral units (voltage supplies and detector interface) takes place via an IEEE-488 standard port and is in ASCII-format. The spectral data are stored as text files in a format compatible with the CRUNCH analysis package [4]. Complementary data for each spectrum are stored on separate text files. Obviously any type of PC e.g. a Macintosh could be used to host the data acquisition program. It should also be mentioned that a new generation of code generating programs, such as LabView™ by National Instruments, is new available. These programs may replace the old programming languages and results in much increased development productivity. The simplicity and modularity of these graphical object oriented development systems now enables non programming specialists to participate in the design and implementation of spectrometer control and data acquisition programs.

Since the present computer for data acquisition is of a single-task type, spectral data are transferred to other computers for further analysis, and consequently no data analysis is included in the spectrometer program.

The spectrometer computers are all connected to a VAX cluster of a μVAX II and several VAX workstations. The recorded spectra are initially transferred to the cluster for analysis and storage, using the Kermit file transfer program. Finally, the spectra are transferred to a network of Macintosh computers for the processing described in this paper. Figure 1 shows the computer systems used in the laboratory.

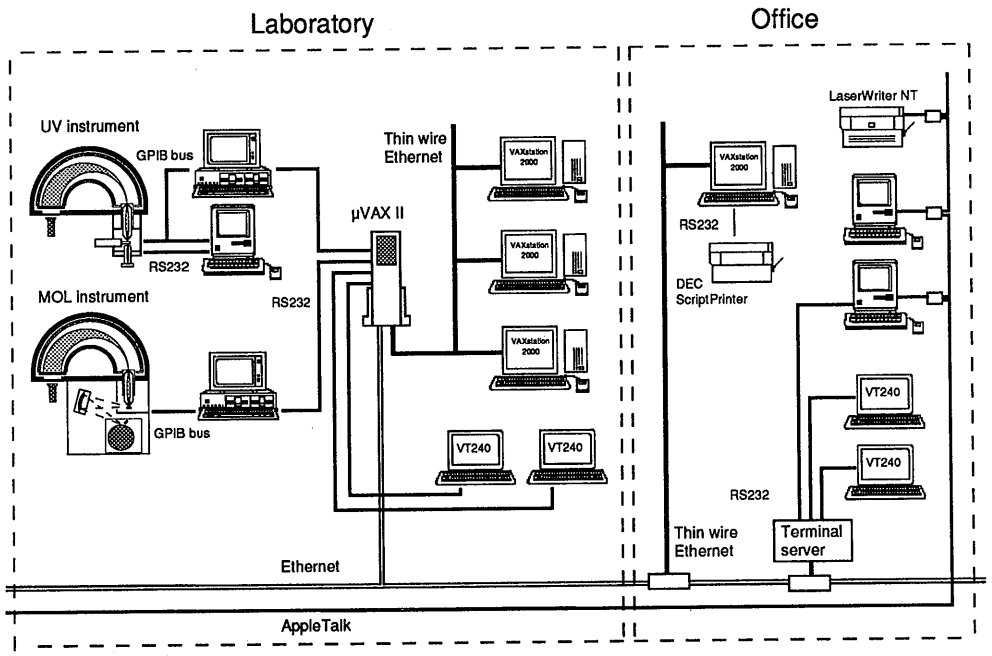


*Fig. 1. The computer systems used in the laboratory for electron- and charged particle spectroscopy of atoms and molecules.*

## Pre-Analysis

A program, CrunchViewer, specialized in pre-analysis has been developed by the group to conform to the internal Crunch file standard. CrunchViewer runs on the Macintosh and assembles a number of activities some of which were earlier performed manually. The program is written in Think Lightspeed Pascal™. Automatic code generation was performed using Prototyper™ by Smethers-Barnes.

All spreadsheet programs require that the different categories and values associated with these categories are provided in single separate columns. The Crunch format, on the other hand, stores six successive channel intensity values, separated with spaces, on each row ( cf Fig 2) and the energies are specified as a centre energy and an energy width parameters in the file header (cf Fig. 2). This format is incompatible with all spreadsheet programs tested. Thus, reading a Crunch file into a spreadsheet program results, at best, in six columns of data.

```
+FOAL4_     Run:    1  Created: 890823  14:49   By mc ESCA Ver 3.21
Formaldehyd                            300mTorr
HeI                                    Heated to 100 C
Excitation: 21.22 Experiment: 1 Repetitions: 10 Mode : SWEEP  Scale: BINDING
REAL ESCA ( 300 ),CENTER (    17.00 ),WIDTH  (     3.00 )
    1257     1204     1241     1171     1215     1311
    1298     1330     1356     1402     1371     1348
    1371     1316     1321     1251     1292     1220
    1349     1300     1302     1343     1434     1456
    1561     1428     1537     1547     1552     1494
```

***Fig. 2.*** *An example of the Crunch file format. The file contains 300 spectral channels, centred around the energy 17 eV and distributed over a 3 eV energy window. The figure shows the first 30 values of the file.*

As the Crunch data file is of text type, it may be edited with any standard text editor available. Such an editing process is, however, often time consuming. The CrunchViewer program simplifies this process considerably. It reads Crunch files and outputs, on request, text files readable by all the spreadsheet programs tested. This implies that the data is reformatted into single columns. The user may select channel number, channel energy or both to be included along with the channel intensity. The channel number, energy and intensity columns are separated by tab characters. The transfer of the file from CrunchViewer to the spreadsheet program may be carried out by copying to a clipboard or by creating a new file, which may be opened directly in the spreadsheet program. Compatibility problems are found when using a mixture of US and Swedish program versions. The US versions require a decimal point (12.34) and the Swedish version requires a decimal comma (12,34) in accordance with the national standards. The CrunchViewer program may be instructed to use either

```
EXCEL format generated from file : Formaldehyde UPS Data:Formaldehyd
UPS:FOAL4X.DAT +FOAL4_    Run:  1  Created: 890823  14:49    By mc ESCA Ver 3.21
Formaldehyd                                300mTorr
HeI                                        Heated to 100 C
Excitation:  21.22 Experiment:    1 Repetitions:  10 Mode : SWEEP  Scale: BINDING
Calibration :
Channel: 239.340 Energy: 15.8370 Step: 0.005
Comment: The 0-0 peak in the B-State.
Channel   Energy     Counts
1         17.029     1257
2         17.024     1204
3         17.019     1241
4         17.014     1171
5         17.009     1215
6         17.004     1311
7         16.999     1298
8         16.994     1330
9         16.989     1356
10        16.984     1402
```

***Fig. 3.*** *The data file converted to spreadsheet readable form.The first two columns correspond to categories (channel no and energy) and the third column contains the intensity values.*

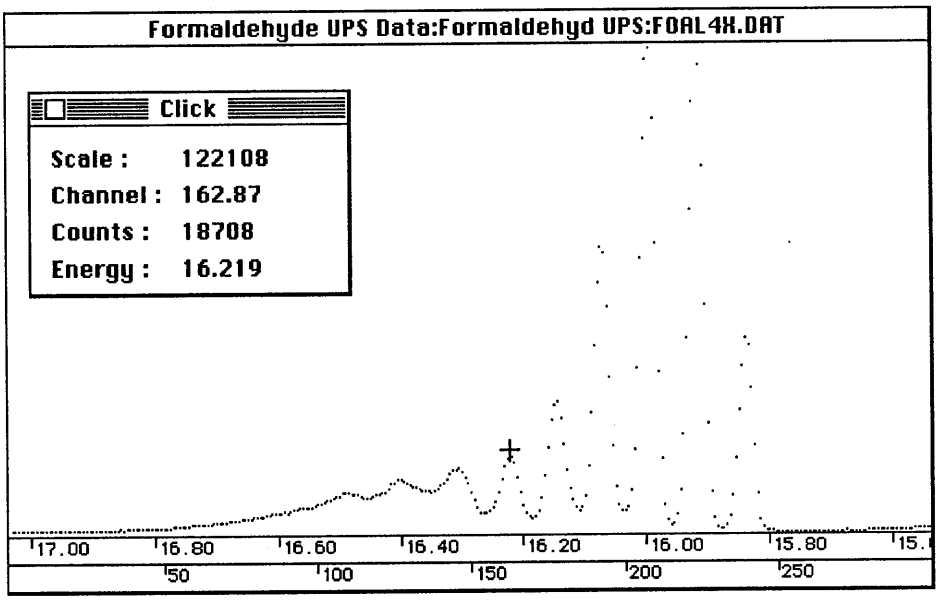


***Fig. 4.*** *A HeI excited photoelectron spectrum showing some vibrational excitations in the B and C states of the formaldehyde cation. The spectrum is displayed by the CrunchViewer program.*

a point- or a comma separator. An example of the resulting file from Figure 2 is shown in Figure 3 using the format of US KaleidaGraph™ for the Macintosh.

Figure 4 shows a part of a HeI excited photoelectron spectrum of the formaldehyde molecule displayed by the CrunchViewer program. The click information is displayed in a separate window.

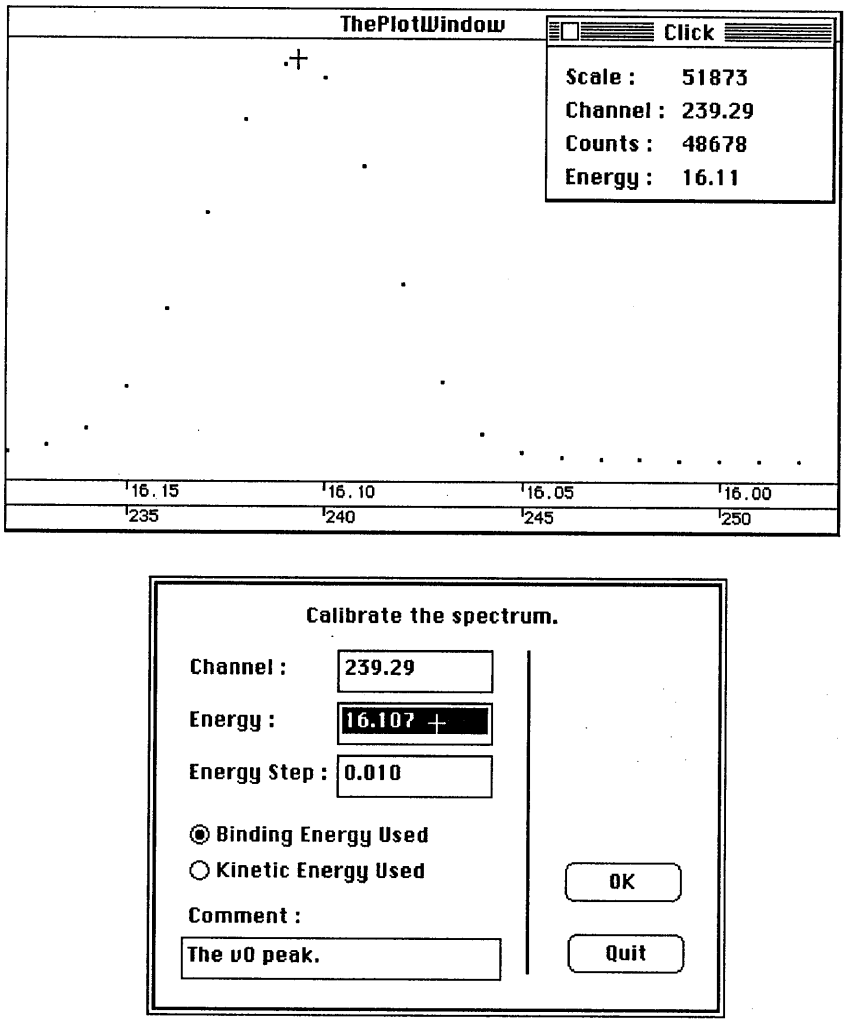


***Fig. 5.*** *A zoomed view in CrunchViewer of the $v_0$ peak (first peak) in the B state. The centre of the peak is obtained by positioning the cursor and clicking the mouse button. The lower part of the figure shows the calibration dialog window in CrunchViewer. The values from the last click are presented as default values in the dialog window.*

An exact determination of the channel energies in a spectrum requires a calibration. This is performed by referring to a spectral line of well defined energy. The CrunchViewer program allows a user to set the energy scale by assigning an energy to a channel, or a channel fraction. The channel is selected by positioning the cursor (cross) on the screen and clicking the mouse button. Figure 5 and shows a calibration of the spectrum in figure 4.

In the following extraction of data from the spectrum, channel energies, numbers and intensities can be directly obtained by means of clicking. CrunchViewer may be informed, by a menu choice, to append extracted data to the clipboard. The assembly of clicks may then be saved to file or transferred to any other program via the clipboard. This possibility, in combination with Multifinder or the Scrapbook, is valuable in the generation of tables containing line positions.

## Modelling and simulation

In photoelectron spectroscopy a spectrum can be obtained which shows a distribution of lines corresponding to transitions from an initial state of, conventionally, a neutral system to different final vibrational cationic states [5]. While the relative intensities of these lines are primarily related to the difference in geometry between the initial and final state, the spacings between the vibrational lines reflect the bonding properties of the ionic state. The shape of the individual vibrational lines is an additional feature that is strongly related to properties of the ionic state. There are thus three main sources of information in the spectrum, which can be used to draw conclusions about the system. The vibrational analysis is mainly focussed on the determination of equilibrium geometries and potential curves of molecular ions from observed experimental data. Theoretical models are used to simulate the molecular systems and provide data suitable for comparison with the experimental results. For these purposes spreadsheet programs are often ideal since the calculations are of small or medium size and has repetitive character. Excel spreadsheet templates have been developed by the group for all steps of the analysis.

## Curve fitting

The first step in the analysis involves fitting of model functions, Gaussians, Lorentzians or a combination of these, to the experimental lines. Gaussians are used to represent the spectrometer function and the Lorentzians represent life-time broadened lines. The line fitting is done by displaying both the experimental and simulated spectra directly on the computer screen and varying the curve parameters of the selected functions until the best fitting to the experimental spectrum is obtained. The experimental and theoretical intensities are stored on the spreadsheet in columns adjacent to that of the independent variable (category), the electron binding energy (or channel number for non-calibrated spectra). Curve fitting may also be performed directly by using other programs with built-in least-squares fitting routines. Hovewer, when the spectrum consists of many

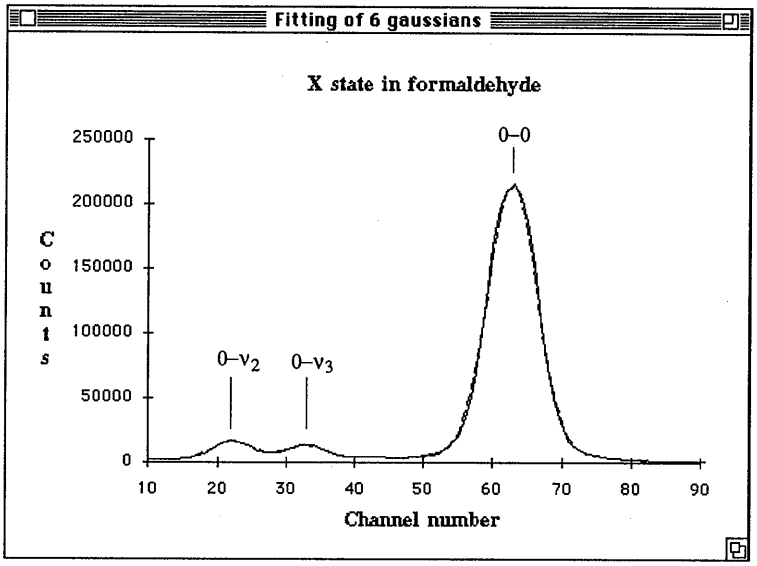


*Fig. 6a. A HeI excited photoelectron spectrum showing transitions to three vibrational levels of the X state of the ion of formaldehyde. Six gaussians are fitted using the Excel program (see text).*

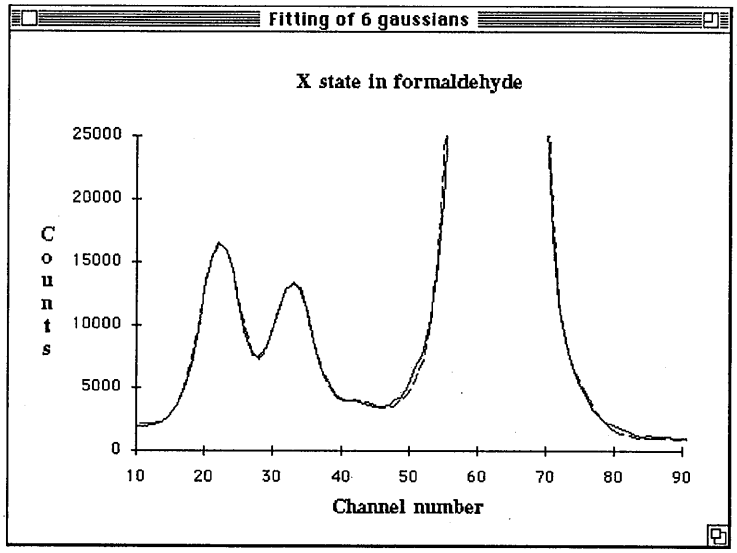


*Fig. 6b. A zoomed view of figure 6a.*

lines which may be partly overlapping, the interactive and visual fitting enabled by Excel is preferable.

Figure 6a shows an Excel fitting of six Gaussians to an experimental photoelectron spectrum from the formaldhyde molecule. In addition to the strong peak reflecting the adiabatic (0-0) transition, the experimental spectrum shows two peaks due to transitions to vibrationally excited states (excitations of single quanta of the $\nu_2$ and $\nu_3$ modes). These three peaks can be very well represented using only three Gaussians. However, due to so-called hot band excitations and the rotational degrees of freedom, low intensity structure is present, particularly in the region of the 0-0 peak. In order to account for these features three additional Gaussians have been introduced with line positions at channel numbers 42.3, 53 and 73 and with very low intensity. The result can be seen in detail in figure 6b, which shows a zoomed view of figure 6a. The structure of the program and the fitting parameters for soam of the Gaussians are shown in figure 7. As can be seen, the formula representing the Gaussian function is the samae in all cells. This implies that new Gaussians can be added whenever necessary simply by copying one of the columns, pasting it into the proper position and changing the parameters to obtain the best fitting to the experimental spectrum. The relative intensities between the spectral lines are obtained by summing all the intensity values in the columns storing the fitted Gaussians. These values are presented for each line on row 9 (Σ). On row 10 are given the same values normalised to 1 for the highest peak.

File Edit Formula Format Data Options Macro Window

E73 =Height*EXP(-((E-Line_Position)^2/(2*Width^2)))

Gaussian Fit

| | A | B | C | D | E | I |
|---|---|---|---|---|---|---|
| 1 | Gauss function fit | | | | | |
| 2 | | | | | | |
| 3 | Background | | Line number | Line_1 | Line_2 | Line_6 |
| 4 | | | | | | |
| 5 | Length | 100 | Line Position | 73 | 62,8 | 22,3 |
| 6 | Left | 2300 | Width | 3,5 | 3,5 | 3 |
| 7 | Right | 750 | Height | 4500 | 215000 | 14500 |
| 8 | | | | | | |
| 9 | | | Σ | 39479,40 | 1886237,78 | 109038,33 |
| 10 | | | \|Σ\| | 0,0209 | 1,0000 | 0,0578 |
| 11 | | | | | | |
| 12 | E | Spectrum | Line Σ | Line_1 | Line_2 | Line_6 |
| 13 | 0 | 6079 | 2300,00 | 0,00 | 0,00 | 0,00 |
| 14 | 1 | 4714 | 2284,50 | 0,00 | 0,00 | 0,00 |
| 15 | 2 | 3967 | 2269,00 | 0,00 | 0,00 | 0,00 |
| 16 | 3 | 3359 | 2253,50 | 0,00 | 0,00 | 0,00 |
| 17 | 4 | 2646 | 2238,00 | 0,00 | 0,00 | 0,00 |
| 69 | 56 | 30561 | 36791,73 | 0,03 | 32566,73 | 0,00 |
| 70 | 57 | 50592 | 58126,82 | 0,13 | 54466,01 | 0,00 |
| 71 | 58 | 83388 | 87046,16 | 0,46 | 83950,71 | 0,00 |
| 72 | 59 | 125239 | 121841,58 | 1,51 | 119253,36 | 0,00 |
| 73 | 60 | 167424 | 158296,77 | 4,54 | 156122,04 | 0,00 |
| 74 | 61 | 196709 | 190235,03 | 12,61 | 188367,18 | 0,00 |
| 75 | 62 | 211526 | 211122,00 | 32,24 | 209456,39 | 0,00 |
| 76 | 63 | 214412 | 216211,29 | 75,96 | 214649,27 | 0,00 |

*Fig. 7. The fitting parameters used in the analysis of figure 6.*

## Morse potential curves

The second step in the analysis involves a fitting of a second-order polynomial to the vibrational energies and from the result, using Excel, drawing a Morse potential curve [6], which is described by equation 1.

$$V = D_e(1 - e^{-\alpha(R-R_e)})^2 \qquad (1)$$

The constants $D_e$, the dissociation energy, and $\alpha$, are obtained directly from the polynomial fitting, performed in KaleidaGraph™, while the equilibrium bond distance $R_e$ is obtained from an analysis of the vibrational intensities, as described below. This potential function is appropriate for the resolution normally obtained in electron spectroscopy. Morse potential curves have recently been determined for several doubly ionized states of the halogen hydrides from Auger electron spectra [7,8,9].

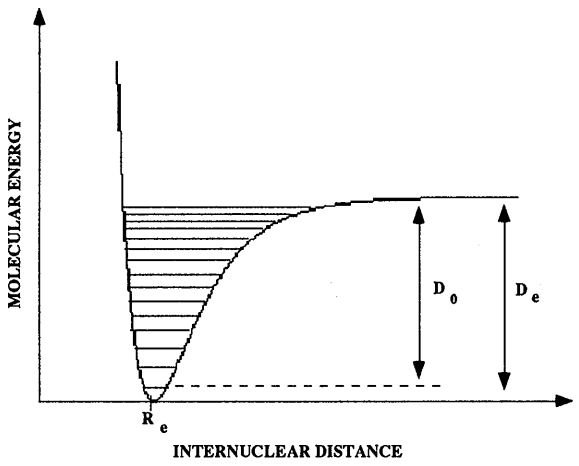


*Fig. 8. Schematic representation of a Morse potential curve. $D_e$ and $D_0$ are the dissociation energies referred to the minimum of the potential curve and the vibrationless state, respectively.*

## Franck-Condon factors

The third step in the analysis involves the determination of the equilibrium bond distance $R_e$ of the ionic state. This is done in an interactive procedure where the vibrational intensities (Franck-Condon factors) are calculated, using the bond distance as a parameter, and compared to the experimental data. In this part, we use the Ansbacher recursion formula (2).

$$R(n+1,m) = \sqrt{\frac{2}{n+1}} \cdot \frac{\beta^2\gamma}{1+\beta^2} \cdot R(n,m) + 2\sqrt{\frac{m}{n+1}} \cdot \frac{\beta}{1+\beta^2} \cdot R(n,m+1) + \sqrt{\frac{n}{n+1}} \cdot \frac{1-\beta^2}{1+\beta^2} \cdot R(n-1,m) \quad (2)$$

where n and m are the vibrational quantum numbers referring to the final and initial states in the transition, $\gamma$ is the parameter containing the change in bond distance between the two states and R(n,m) are the vibrational overlap integrals between the final and the initial states. As an example, Figure 9 shows a screen snapshot of the parameters (right) and result (left) of such a calculation of $R_e$ for $CO^+$, which gives the best agreement with the experimental spectrum [10]. The relative intensities at high vibrational quantum numbers are lower than in the experimental spectrum . This is, at least partly, due to the fact that the Ansbacher formula rests on the harmonic approximation from which the molecular potential deviates strongly for these high quantum numbers.

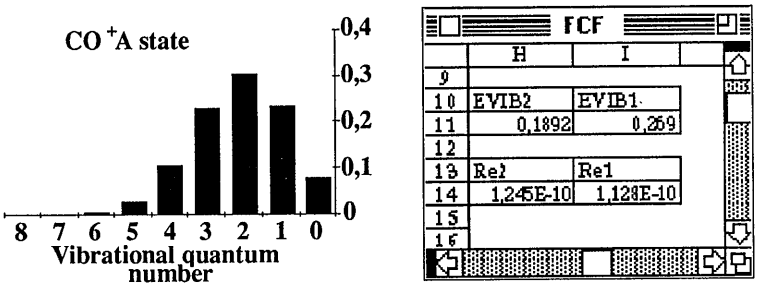


*Fig. 9. Left. Calculated intensity distribution for the photoelectron spectrum of the A state in $CO^+$. Right. The vibrational energies and internuclear equlibrium bond distances used to obtain the figure to the left. Column H represents the ionic state and column I the neutral state. The latter values were obtained from ref. [11]. The vibrational energy in cell H11 was obtained from ref. [10], while the bond distance in H14 is the required quantity.*

**Presentation**

Before submitting a publication, the figures containing the spectra must be properly prepared. The preparation involves scaling, labelling of the axes and adding text and lines to mark details of special interest. Several programs on the Macintosh are able to plot data in graphs. We find the KaleidaGraph™, by Adelbeck Software, and DeltaGraph™, by DeltaPoint Inc., to be outstanding in spectrum presentation at present. Fig. 10 shows a monochromatized x-ray excited outer and inner valence photoelectron spectrum of the HCl molecule [12]. Often a 'guide-to-the-eye' line is added to the spectrum as in this figure. This line smoothens the appearance of the spectrum and indicates the magnitude of the noise. As this line has no mathematical meaning, it is often difficult to use overall curve fitting utilities to generate this line. Polynomial spline fitting, which is available in KaleidaGraph, usually gives satisfactory results. However, other alternatives can also be used. The line shown in Figure 10 was added by transferring the spectrum, from a plot in KaleidaGraph, as a PICT file to the Canvas™ program and adding the line by means of a Bézier polynomial.

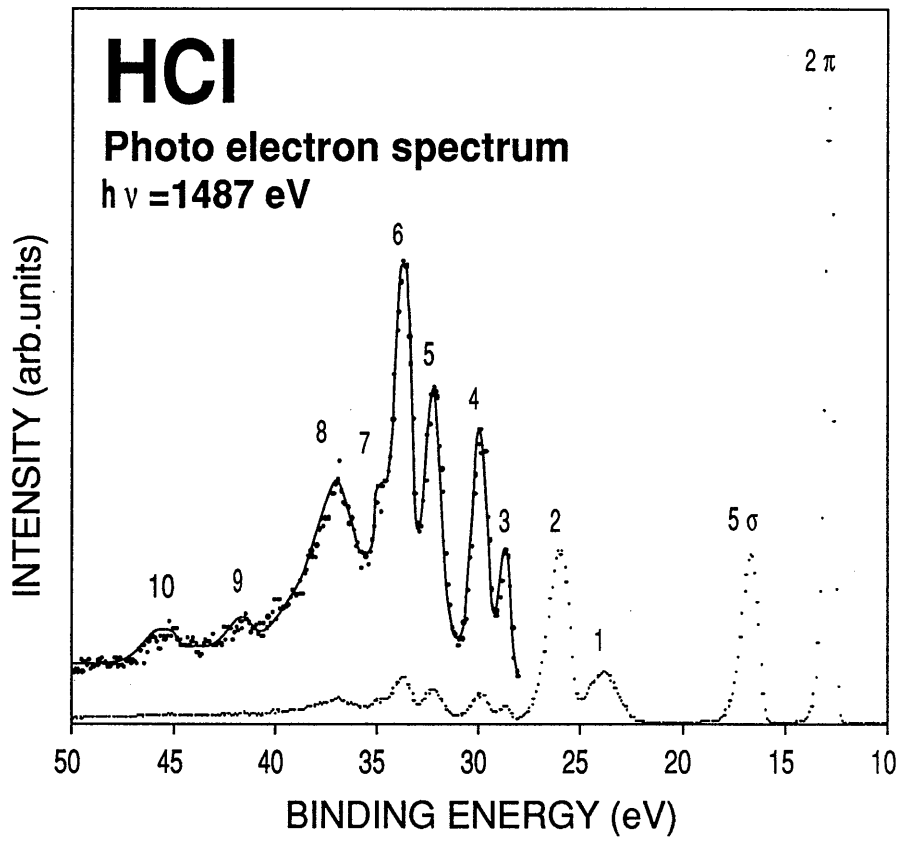


*Fig. 10. A monochromatized x-ray excited outer and inner valence photoelectron spectrum of the HCl molecule prepared for publication.*

## Conclusions

The use of a number of computers and programs is usually necessary in the creation and handling of experimental data. User friendliness in transferring the data between computers and programs is essential. The new generation of personal computers, such as the Macintosh, has been found to provide an excellent platform, along with modern spreadsheet programs, for spectrum analysis and presentation. The simplicity, uniformity and interconnectivety of the Macintosh programs gives a steep learning curve. The latter aspect is particulary important when introducing new members in the research group but also quite useful for it releases old members of the group from dealing with intricacies of interfaces in various computer systems.

## References


[1] P.Baltzer, B.Wannberg and M.Carlsson-Göthe,
Uppsala University Institute of Physics Report, UUIP-1182, 1989.
Subm. for publ. in Rev. Sci. Instrum.

[2] D.Nordfors,
Acta Universitatis Upsaliensis 248, Thesis, 1990.

[3] J. de Sousa Pires,
Electronics Handbook, ISBN 91-44-21021-3, Studentlitteratur, Lund, (1989).

[4] R.P.Vasques, J.D.Klein, J.J.Barton and F.J.Grunthaner,
J. Electron Spectrosc. **23**, 63 (1981).

[5] H.Siegbahn and L.Karlsson,
Handbuch der Physik, Vol. 31 (ed. W.Mehlhorn), Springer Verlag, 1982.

[6] G.Herzberg,
*Molecular Spectra and Molecular Structure. Spectra of Diatomic Molecules*,
Van Nostrand-Reinhold Co. New York, 1950.

[7] S. Svensson, L. Karlsson, P.Baltzer, M.P. Keane and B. Wannberg
Phys. Rev. A (In print 1989)

[8] B. Wannberg, S. Svensson, M.P. Keane, L. Karlsson and P. Baltzer
Chem. Phys. **133**, 281 (1989)

[9] L. Karlsson, S. Svensson, P. Baltzer, M. Carlsson-Göthe, M.P. Keane,
A. Naves de Brito, N. Correia and B. Wannberg
J. Phys. B **22**, 3001 (1989).

[10] B. Wannberg, D. Nordfors, K.L. Tan, L. Karlsson and L. Mattsson.
J. Electron Spectrosc. **47**, 147 (1988).

[11] K.P.Huber and G.Herzberg,
*Molecular Spectra and Molecular Structure, vol 4, Constants of Diatomic Molecules.*
Van Nostrand Co. Princeton, 1979.

[12] S.Svensson, L.Karlsson, P.Baltzer, B.Wannberg and U.Gelius.
M.Y.Adam
J.Chem.Phys. **89**, (12) (1988).