SwaN-MR: A Complete and Expansible NMR Software for the Macintosh

Giuseppe Balacco

Laboratori di Ricerca Chimica, A. Menarini Industrie Farmaceutiche Riunite s.r.l., Via Sette Santi, 3-50131 Firenze, Italy

Received March 1, 1994[®]

A new computer program is presented useful to any scientist involved in the field of FT-NMR spectroscopy. It comprehends a complete set of functions to manipulate up to four-dimensional spectra as well as an array of tools for annotating presenting these spectra. Principles that directed its design include the implementation of strict scientific language and extreme easiness of use and learning. Spectra derived from different models of spectrometers have been successfully transferred to the Macintosh and processed routinely. A six-month experience in two chemical laboratories raised favorable appreciations.

INTRODUCTION

Rapid progress in the field of personal computers allowed in the last few years the use of such an inexpensive platform for off-line processing of NMR spectra. Although many programs appeared for this purpose,[1,2] I felt the need of developing a new one. Besides processing two-dimensional spectra, the new program should have met the following requisites: use of a strictly scientific language; capability of reading spectra coming from any spectrometer; immediateness in the sense that the least possible number of rules and commands must be learned and memorized by the user; room for future expansions; easy interfacing with other programs; facilities to edit spectra, both graphically and numerically; and low cost. My colleagues and I felt, in fact, that none of the existing programs responded simultaneously to the above requisites. Furthermore, when I began to write this program, entitled SwaN-MR after a pop-music song,[3] I did not know of any NMR software running on a Macintosh, except Win-NMR[2] that processes only one-dimensional Bruker spectra and is not available in Italy.

The first problem facing anyone wishing to process their spectra on a PC is data transfer. This problem is more serious with old spectrometers. We have found low-budget solutions for three different spectrometers, so our conclusion in this respect is that a solution can always be found, though it may be painful if the spectrometer manufacturer does not cooperate and very expensive in the opposite case. I hope that an increasing request for data-transferring on PCs may elicit more attention from NMR instruments vendors and third parties.

EXPERIMENTAL SECTION

SwaN-MR was written in C using the Symantec Think C 6.0 compiler and the Think Class Library.[4] No other library, e.g., ANSI run-time libraries, was included. All the algorithms are original of the author; some of them, however, were inspired by a well-known book.[5a] Listing of these algorithms are available upon request to the author. Requests must precise which algorithm is wanted. To obtain maximum speed, a few calculation-intensive sections were written in assembly language and encapsulated in the C source. SwaN-MR requires a MC 68020 or higher processor and a

floating point unit. The program can read only files written in its own format, so a collection of converters has also been written. SwaN-32 converts spectra in the Bruker X-32 format. SwaN-Asp converts spectra in the older Aspect 24-bits format. A more sophisticated program, named AcchiappaSpettri, which is the Italian for spectra-catcher, is devoted to interface a Macintosh with a Varian Gemini system. It imports spectra coming out from the parallel port of the instrument and converts them directly in the SwaN-MR format. The same application drives the Hewlett-Packard plotter normally shipped by Varian with its spectrometers.

A brief explanation on how data-transfer was accomplished can be useful for the interested reader. It should be noted that our solutions are not unique: often it is possible, when a spectrometer is provided with a serial output, to transfer data using the Kermit file protocol.

Spectra from a Varian Gemini 200 were transferred from the parallel printer port to the Macintosh modem port, through a parallel-to-serial converter and a RS-232 cable. This hardware is available from computer supplies stores. A series of macros was written on the Gemini to make the overall process simple even to inexperienced users. These macros are eventually based upon the Varian commands HPRINT and IPRINT. Spectra from Gemini arrive as a text flow on the Macintosh, where this flow is read and interpreted by AcchiappaSpettri. Transferring a spectrum may take from 3–30 mins, so we normally prefer to wait until the night and then to transfer all the spectra acquired during the day.

A spectrum of human cerebro-spinal fluid was kindly supplied from Prof. G. Natile, University of Bari. He used the NMRLINK-PC program to transfer this spectrum from a Bruker AM-300 to an IBM PC, without data conversion. The IBM formatted file was then converted on the Macintosh with the Apple File Exchange application (supplied with the operating system to all Macintosh buyers). At this stage, the content of the file was still in the original 24 bits form. Finally, conversion to the SwaN-MR format was performed with the cited SwaN-Asp program.

Spectra acquired a Bruker AMX-600 and stored on tape were loaded on a Silicon Graphics Indigo workstation and transferred with an Ethernet cable to a Macintosh. Here the final conversion was carried out using SwaN-32.
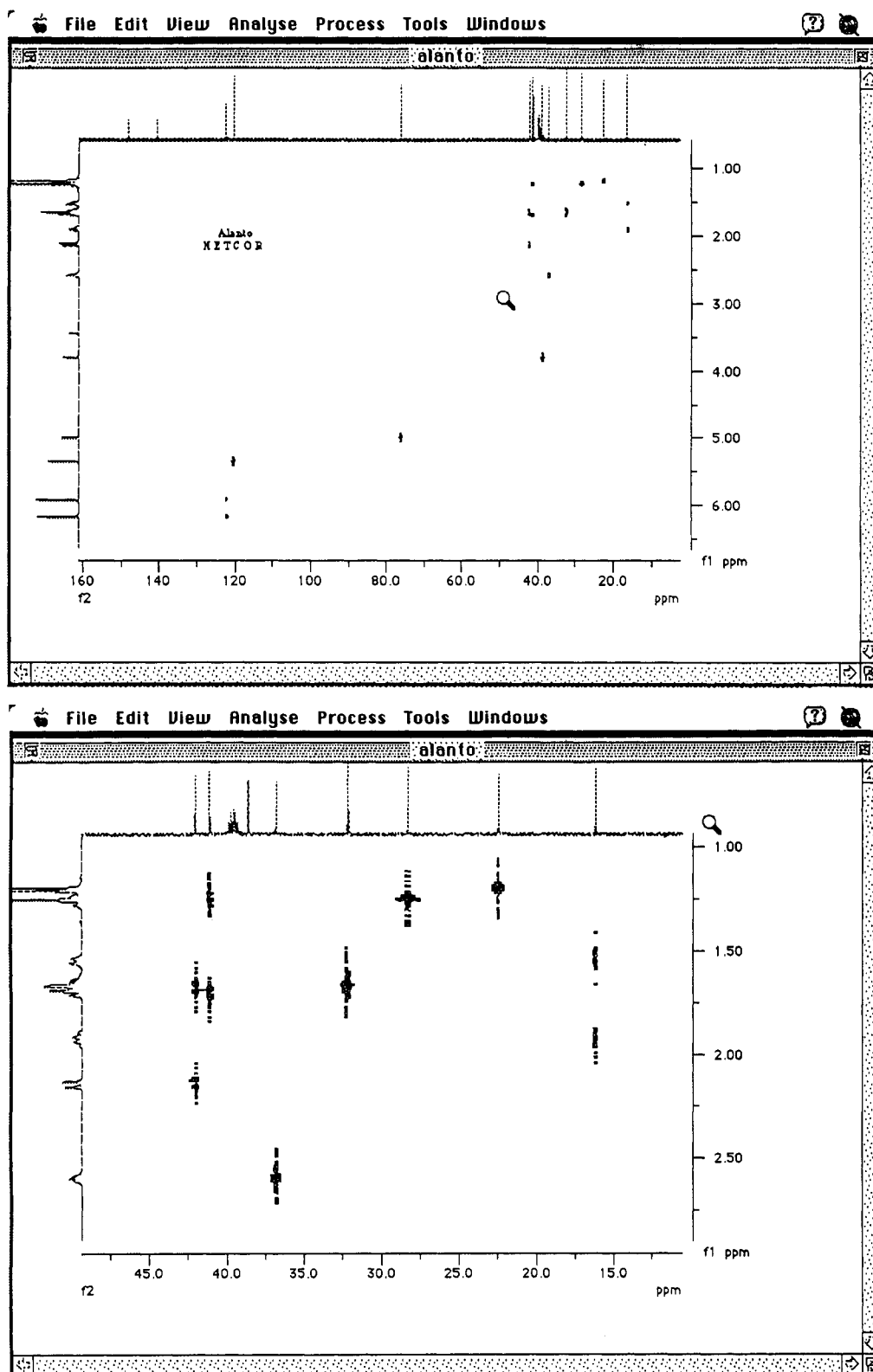
Figure 1. A 2D spectrum (a, top) before and (b, bottom) after a zoom operation with the mouse.

INTERFACES

The SwaN-MR user is always faced with the same menus, regardless if one is working with 1D or 2D data sets. Basic functions, like zooming, referencing, integration, etc... require the same operations. Newcomers to the program can therefore learn how to use all its functions in a very short time. Each spectrum is shown in its own window, together with accessory elements like scales, integrals, and comments. Every element, including the spectrum itself, can be dragged

with the mouse, so the user is given total freedom of constructing the image that will be seen on the screen or on paper.

A central concept in this respect is that of page. A page is a set of options created by the user. These options include plot limits, scale units, colors, the presence or absence of scales, integrals, annotations, etc... Up to 99 pages may be created for each file. Though all pages share the same data, they are seen by the operative system just as different pages
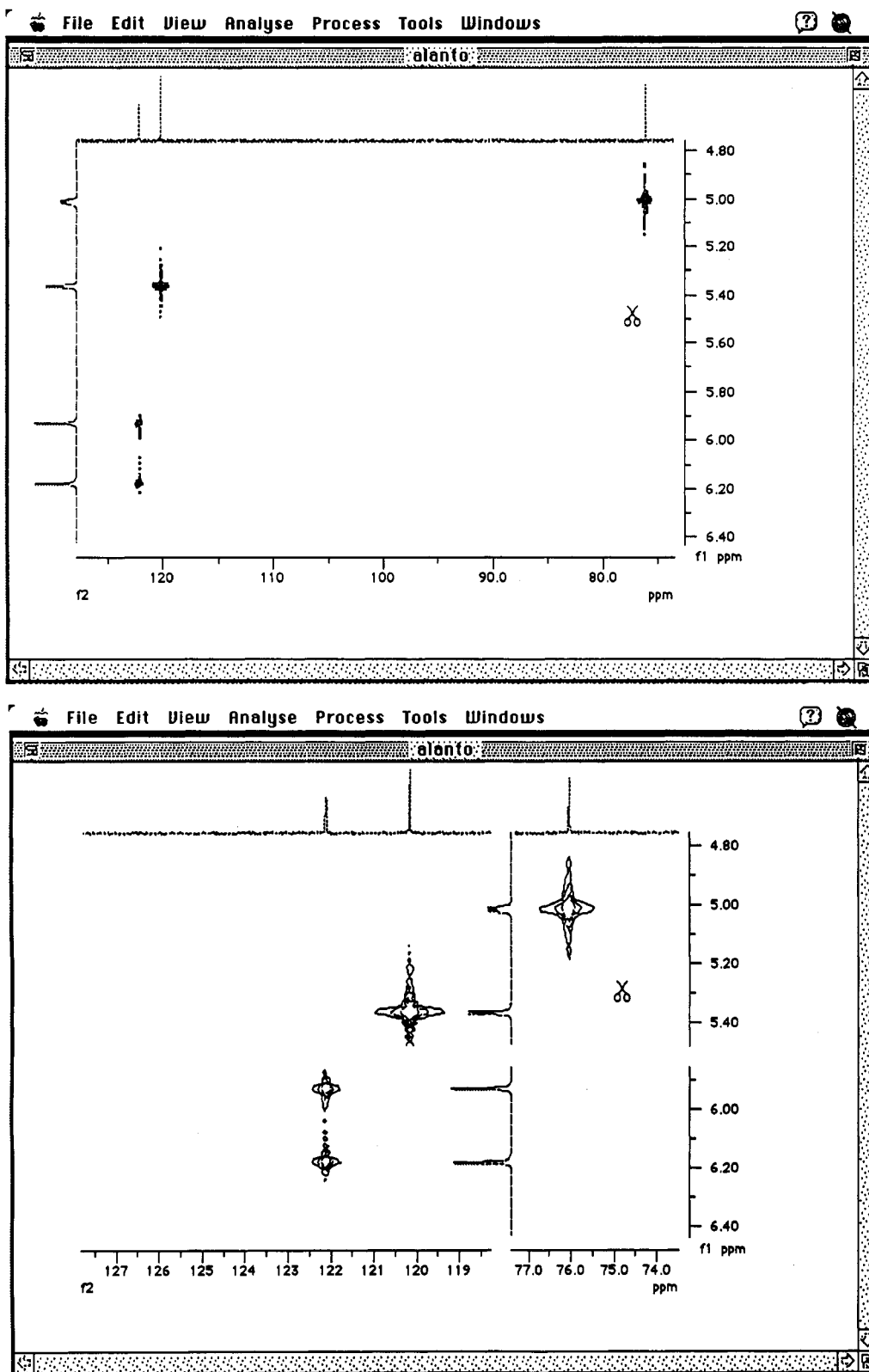
Figure 2. A spectral region void of peak is excluded: (a, top) before and (b, bottom) after the cut. As can be seen by projections and scales, four regions of the spectrum are simultaneously expanded.

of a document. This trick simplifies printing operations.

Very simple mechanisms allow direct annotation on the spectrum with text or figures. These notes are called "comments". The "peak-peaking" function of SwaN-MR writes down peak frequencies just in the same format. This uniformity enables the user to edit the peak-peaking output: an example could be deleting the frequencies of the solvent resonances. Comments may be grouped in sets to be shown alternatively, as for pages.

Many commands (e.g., expansions, integration, J coupling measurements, etc...) are issued with a single mouse operation. The user can select among a series of tools. Figure 1 shows the pointer transformed into a magnifying lens, ready to perform an expansion. A very useful and nice feature of the program is the possibility of expanding and showing simultaneously up to nine portions of a spectrum, after cutting it with a scissors-shaped tool (see Figure 2). There are cases when more precision is needed and someone finds it difficult
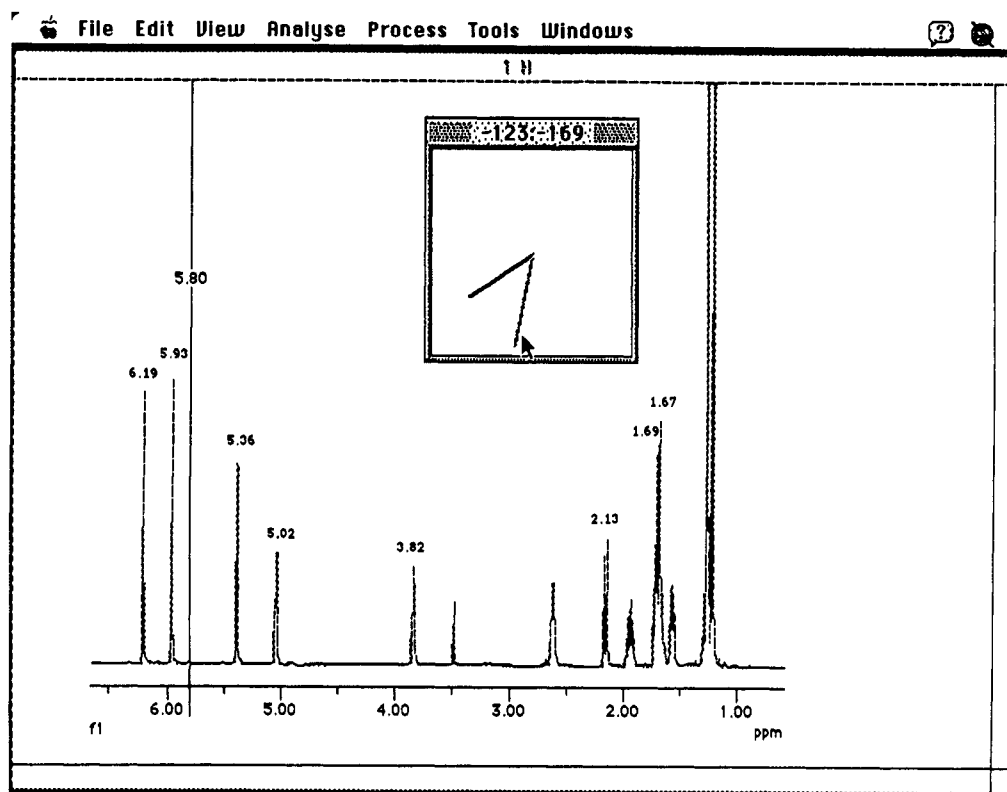
Figure 3. Manual phase correction with a pivot.

to use the mouse. An example is phasing. Many other programs associate a vertical displacement of the mouse to a first-order correction, while a horizontal displacement causes a zero-order correction. Personally I find it hard to adjust simultaneously to the zero- and first-order corrections, so I have adopted a clock with two hands. These have to be dragged independently with the mouse button held down. Figure 3 shows how this operation is carried out. The vertical mark at 5.80 ppm indicates the "pivot point". Marks such as the ones shown in Figure 2 can be created in any number, both horizontal and vertical. They indicate to the program a point in the spectrum on which a subsequent operation must be applied. Figure 4 shows the case in which the frequency scale along f_1 is referenced against a signal. Unlike the lens in Figure 1 or the scissors in Figure 2, marks are not a cursor that follow the mouse movement: they are elements in the window that can be dragged by the mouse.

Visualization is limited to 1D traces and 2D slices. In the latter case there is the usual choice among intensity, contour, stacked, and whitewashed plots. Any number of different spectra may be overlapped, in order to compare them easily. The user may suspend visualization if it slows the program.

Everything is implemented in pure Macintosh style. The newest features of System 7 have been adopted, like Help Balloons for on-line help and Apple Events for interaction with Finder. The clipboard is used to export and import text and pictures. It can also be used to transfer special data between SwaN-MR documents, e.g., integration limits or data points. Future applications may, of course, adopt the same format to exchange data with SwaN-MR. Data can also be exported to a digital plotter interpreting the Hewlett-Packard language. Another instrument for organizing and transferring data is a general purpose window named Results that receives the text output of all the SwaN-MR functions. Once in the Results window, text can be printed, selected, and copied or

saved as a file. This window can even be used to represent data numerically instead of graphically.

ALGORITHMS

SwaN-MR lacks concepts that are of central importance in non-Macintosh environments, like the command line and the macro commands. The strategy chosen to obtain flexibility and user-friendliness was to introduce a sort of "silent agreement" between the program and the user. A few examples can clarify this idea.

When a spectral file is opened for a time other than the first one, SwaN-MR asks the user if he wants a Fourier transform to be applied. If the spectrum has never been processed before, SwaN-MR awaits commands by the user because it is implicit that one has still to choose the parameters for weighting and transforming.

Normally, before processing along each dimension an entire 2D phase-sensitive data set, a spectroscopist processes a trace of it along that dimension in order to optimize apodization and phase-correction. Then one takes note of the best parameters and processes the whole 2D spectrum. With SwaN-MR there is no need to take notes, because the program takes care of memorizing the parameters.

The last example is discussed as follows. It was already said that SwaN-MR has a command to overlap spectra. If the user tries to overlap a 1D spectrum over a 2D one, the program behaves differently; it will show the former as a projection along the appropriate axes because in such a case this is what a spectroscopist normally wants.

These examples show how a lot of functions are available to the user through a minimum number of menu commands. Many specifications are implied. Undetermined situations would have risen unless three main simplifications were assumed before writing the program: (1) Many files may be simultaneously open, but for each of them it is never
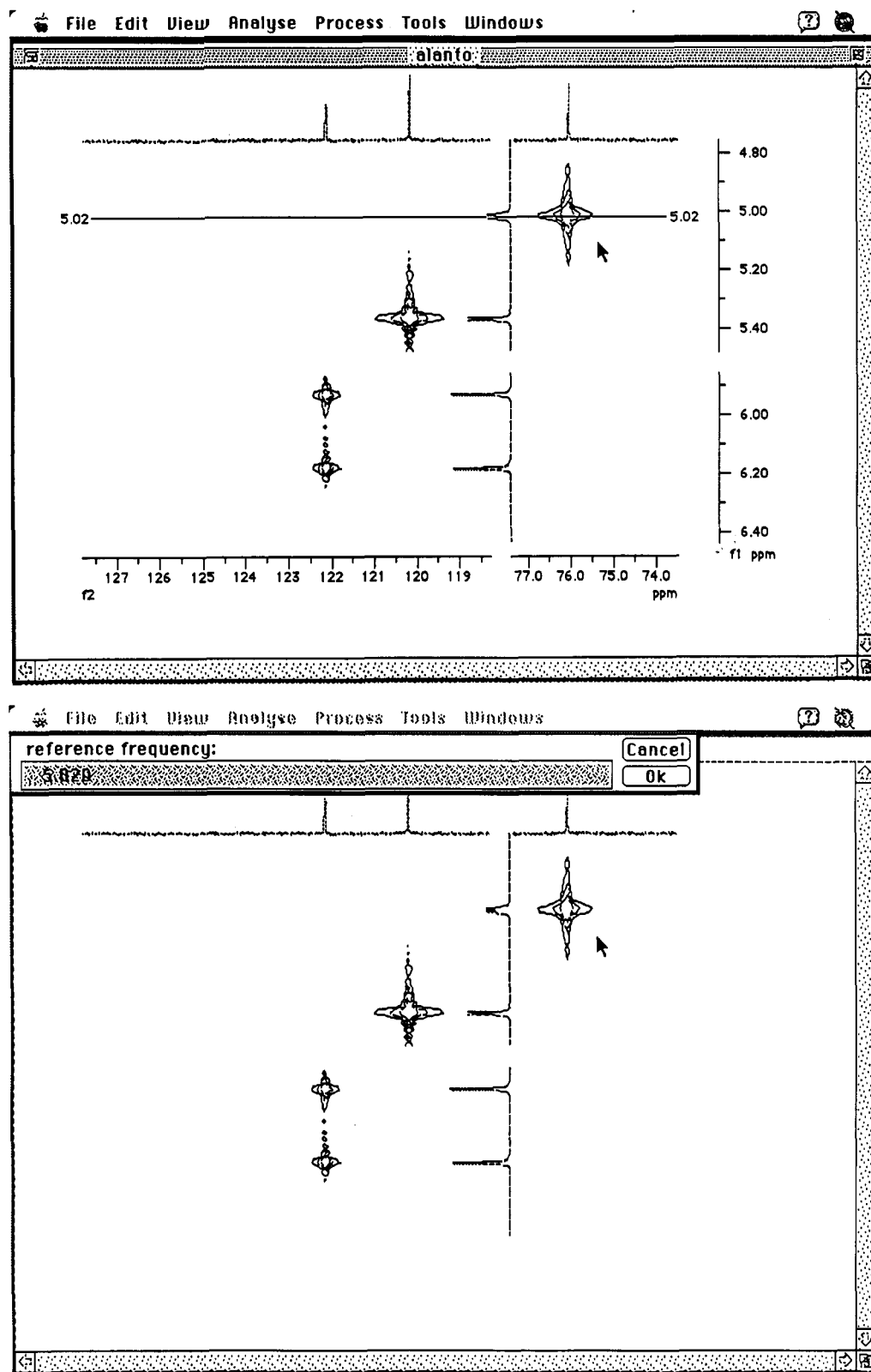
Figure 4. Referencing the f_1 scale. (a, top) The user has put a horizontal mark on the reference peak; (b, bottom) after issuing the Calibrate command a dialog appears; a new value can be entered.

possible to have more than one window simultaneously open. (2) The processing of nD spectra must always be performed starting from the directly revealed dimension toward the more indirectly revealed one. This order corresponds to the normal practice. (3) All spectra have four dimensions.

These simplifications, or pre-established limits, constitute the skeleton of the program. A spectrum is considered as a four-dimensional matrix, in which any dimension may be collapsed into a single point. It is possible to load it from

disc in its entirety or only a portion of it (any 4D submatrix). It is then possible to extract any portion of an already loaded spectrum in a different area of memory and so on, creating a "stack of extracts". Only the extract on top of the stack is accessible at any moment. After deleting it, the previous extract comes to the foreground. Processing and displaying parameters are shared by all extracts. With this mechanism the user can, at any stage, experiment quickly with a new operation on a small portion of the spectrum (or on a copy
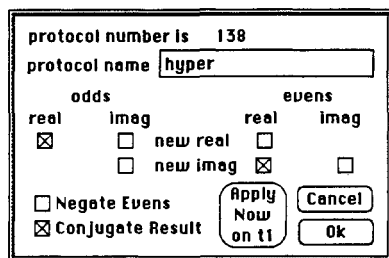
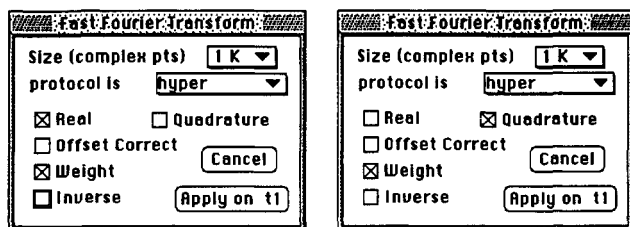**Figure 5.** Dialog for specifying special processing protocols.

Figure 6. Dialog for specifying main FT options.

of it) and, once one has found the best parameters, return to the main data set. Processing follows the usual flow of events; apodization, zero-filling, Fourier transform, and phase correction are grouped into a single command. Phase correction can also, of course, be performed separately. Many algorithms were attached to this skeleton. Their number is continuously growing. Below is depicted the situation at this writing time.

Weighting functions include the following: exponential line broadening and resolution enhancement, Gaussian transformation, sine and cosine bell (with shifted and squared variants), Hanning and Hamming apodizations, and polynomial smoothing.

Fourier transform includes many options: first point correction, offset correction, zero-filling, real and singlature transformation, conjugation, and frequency shifting. True nD transformation (not just a repetition of many 1D transformations) is executed in memory for maximum speed. This implies that transposition of matrices is never needed. If a spectrum is too large to be loaded in memory, it is fragmented in two or more pieces.

Figure 5 shows how compilation of special processing programs may be accomplished by the user. He has to check or uncheck eight different buttons. The example in the figure pertains to both the TPPI and the Ruben–States acquisition modes. The diagram means that, for each interferogram in the f_1 dimension, the real part comes from the real part of odd t_1 increments, while the imaginary part comes from the real part of even increments; finally the complex conjugate is taken of each point to invert the sign of frequencies. The number of possible protocols is limited: $2^8 = 256$; not all of them are really useful. Every single user is free to rename them in a way which is meaningful to him/her. Compatibility between different copies of the program is guaranteed by a numerical index attached to each protocol, which cannot be altered. Figure 6 illustrates where the two above methods (TPPI and Ruben–States) differ from each other. The former uses a real Fourier transform without quadrature detection. The dialog in this figure is also used for zero-filling and Fourier transform.

Linear prediction has been for a long time considered very time consuming even when performed on much more expensive platforms than the Macintosh.[7] Recently a new algorithm has appeared for QR decomposition of generic Toeplitz matrices.[8a,b] It is very fast (almost comparable to the FFT algorithm) and numerically stable, so, after writing a modified version for complex numbers, I based on it many useful operations. They are as follows: reconstruction of missing or altered portions of the spectrum; extraction of the signals' parameters directly from the FID; and data compression to save disc space. Linear prediction may also be elegantly combined with convolution in time domain. When Marion et al. proposed this scheme for solvent suppression,[9] they reconstructed the first and last points of the FID (which cannot be convoluted) with a simple linear extrapolation from adjacent points. LP is undoubtedly more correct end efficient, because it corresponds to the theoretic shape of signals.

Phase and base line corrections are feasible both in one and in two dimensions. Phase can be corrected either manually (see Figure 3) or with some known automatic methods. The simplex method proposed by Siegel[10] is implemented in a totally automatic fashion, while the implementation of the DISPA algorithm[11] requires the user to specify two reference peaks. The choice between the two solutions is left to the user. DISPA is more precise but needs two well resolved and symmetric lines. The other method is insensitive to peak crowding and non-Lorentzian line shapes but suffers from noise and distorted base lines. Magnitude representation of spectra is, of course, available.

Base line correction may also be performed with different degrees of automatism. The user can choose between two fitting functions: a polynomial (up to the fifth degree) and a trigonometric series (of up to 11 terms). Fitting between the spectrum and these functions is calculated with a Householder reduction, which is remarkably faster than the often preferred Singular Value Decomposition.[5b] Initially I had adopted the latter, but soon I discovered that, in the context of base line correction, when a singularity is met there is nothing that can remedy it. Efforts must be directed toward avoiding the singularity. The condition number of the design matrix can be easily decreased of orders of magnitude if all frequencies are normalized between 0.5 and 1.5 (when fitting to a polynomial) or among 0 and π (when fitting to a trigonometric series). All fitting operations in SwaN-MR (including Linear prediction) are performed with the highest numerical precision. Automatic sampling of the base line was inspired by a published paper.[12] The same algorithm is called by another automatic routine that integrates 1D spectra.

Fitting of peaks to Lorentzian and Gaussian absorption shapes is possible. The fitting algorithm is the one by Levenberg and Marquardt.[5c] There is no upper limit to the number of peaks that can be simultaneously fitted.

Integration is the simplest of algorithms, yet it is not obvious when the spectrum has two dimensions. The common tactic is to sum the intensities of points enclosed by a given rectangle. In this way, interference of neighbor peaks is marked. The opposite tactic would substitute the rectangle with an irregular curve following closely the peak shape. This alternative is not yet practicable, neither manually nor automatically. SwaNMR adopts, for the moment being, an oval shape, which is often near to the base of two-dimensional peaks.

SwaN-MR: NMR Software for the Macintosh

J. Chem. Inf. Comput. Sci., Vol. 34, No. 6, 1994 **1241**

Finally, an array of utilities is given for spectral editing, from simulation to symmetrization, passing through sum/difference/substitution/suppression of (portions) of spectra.

CONCLUSION

Many researchers in ours and in another company (Malesci-Firenze) have been using SwaN-MR during the last 6 months. All of them have found it superior in every aspect to the software of our spectrometers (which are Varian's). During this period, many new functions were added to the program without altering neither the program structure nor the format of spectral files. This is a measure of how easily the program can be expanded and maintained.

The software described in this article is licensed free from Menarini to any research group sending an explicit and signed request.

ACKNOWLEDGMENT

REFERENCES AND NOTES

(1) (a) Felix for Windows; Biosym Technologies, Inc.: 9685 Scranton Rd, San Diego, CA 92121-2777, 1992. (b) Grams/386; Galactic Industries Corporation: 395 Main St, Salem, NH 03079, 1993. (c) Nuts; Acorn NMR: 46560 Fremont Blvd, #418, Fremont, CA 94538-6482, 1994. (d) MacFID; TecMag: 6006 Bellaire Blvd, Houston, TX 77081, 1993. (e) NMR-286; SoftPulse Software: Box 504 Guelph, Ontario N1H 6K9 Canada, 1990. (f) LabCalc; Galactic Industries Corporation: 395 Main St, Salem, NH 03079, 1991.

(2) Win-NMR; Bruker Analytische Messtechnik GmbH: Silberstreifen, D-7512 Rheinstetten 4, Karlsruhe, Germany, 1991.

(3) Berry, W.; Buck, P.; Mills, M.; Stipe, M. Swan Swan H in Life Rich Pageant-R.E.M. IRS, 1986.

(4) Think C 6.0; Symantec: 10201 Torre Avenue; Cupertino CA 95014-2132, 1993.

(5) (a) Press W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. Numerical Recipes in Pascal—The Art of Scientific Computing; Cambridge University Press: New York, 1986. (b) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. Numerical Recipes in Pascal—The Art of Scientific Computing; Cambridge University Press: New York, 1986; Chapter 14.3, p 562. (c) Press W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. Numerical Recipes in Pascal—The Art of Scientific Computing; Cambridge University Press: New York, 1986; Chapter 14.4, p 574.

(6) Mannheim, M.; Hollender, L. NMRLINK-PC, Bruker Spectrospin: 34, rue de l'Industrie, F-67160 Wissembourg, France, 1991.

(7) Led, J. J.; Gesmar, H.; Application of the Linear Prediction Method to NMR Spectroscopy. *Chem. Rev.* **1991**, *91*, 1413−1426.

(8) (a) Cybenko, G. Fast Toeplitz Orthogonalization Using Inner Products. *Siam. J. Sci. Stat. Comput.* **1987**, *8*, 734−740. (b) Gesmar, H.; Hansen, P. C. "Fast" Linear Prediction and Its Application to NMR Spectroscopy. *J. Magn. Reson.* **1994**, *106*, 236−240.

(9) Marion, D.; Ikura, M.; Bax, A. Improved Solvent Suppression in One- and Two-Dimensional NMR Spectra by Convolution of Time-Domain Data. *J. Magn. Reson.* **1989**, *84*, 425−430.

(10) Siegel, M. M. The Use of the Modified Simplex Method for Automatic Phase correction in Fourier-Transform Nuclear Magnetic Resonance Spectroscopy. *Anal. Chim. Acta* **1981**, *133*, 103−108.

(11) Sotak, C. H.; Dumoulin, C. L.; Newsham, M. D. Automatic Phase Correction of Fourier Transform NMR Spectra Based on the Dispersion versus Absorption (DISPA) Lineshape Analysis. *J. Magn. Reson.* **1984**, *57*, 453−462.

(12) Dietrich, W.; Rudel, C. H.; Neumann, M. Fast and Precise Automatic Baseline Correction of One- and Two-Dimensional NMR Spectra. *J. Magn. Reson.* **1991**, *91*, 1−10.