Progress in NMR Spectroscopy, 1980, Vol. 14, pp. 27–66.

0079-6565/80/0301-0027$05.00/0

DIGITISATION AND DATA PROCESSING IN FOURIER TRANSFORM NMR

J. C. LINDON and A. G. FERRIGE
Department of Physical Chemistry, Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, U.K.

(*Received* 17 *December* 1979)

CONTENTS

1. INTRODUCTION

1.1. *Aims of the Review*

In this article we have attempted to describe the process which is undertaken to obtain a plotted, data reduced Fourier transform NMR spectrum. Our aim has been to highlight the pitfalls which can cause a decrease in information content in a NMR spectrum which occurs simply because a digital signal processing domain is used. We have also used the opportunity to summarise the various steps involved in taking a free induction decay from a NMR spectrometer and turning it into a plotted, line-listed, integrated record on a chart. At each stage of this procedure a number of approximations and assumptions are made and these are examined to see if they are justified.

In the main, we have limited ourselves to the experiment of a pulse acquisition followed by Fourier transformation, and we have avoided except for comparison purposes other forms of NMR signal processing such as rapid-scan correlation NMR[1,2] and Hadamard transform methods.[3]

Some problems such as limited dynamic range or poor resolution can be overcome by designing new experiments such as the Redfield 214 pulse method[4] for the former and the use of spin–echo sequences[5] for the latter. For fuller details of these methods and other techniques for optimisation of NMR spectrometer settings the reader is directed to other articles.[6,7]

Similarly, we have avoided discussion of detailed computational methods for implementing different algorithms and for the precise nature of the computations involved in an NMR data system the book by Cooper[8] is most useful. At the other end of the scale an overview of NMR software has recently appeared.[9]

We have also only briefly described the whole area of two-dimensional Fourier transformation as undoubtedly this will be extensively reviewed in the near future. However, where special problems occur in two-dimensional NMR, such as for spectrum phase correction,[10] these are described.

In the section on digitisation, we have included the case of quadrature phase detection since this is now universally available on all modern research spectrometers.

Finally, readers of this article will find no mention of any specific commercial spectrometer. If this is a disappointment to some users who had hoped to find specific details of their system, we make no apologies for this since we wished to remain perfectly general.

1.2. *FTNMR Computer Systems*

All Fourier transform NMR spectrometers use a dedicated minicomputer to process raw data from a spectrometer. Some computer systems simply consist of a core-resident FTNMR manipulation program which can control the acquisition of data from a number of different experiments (e.g. inversion–recovery, spin–echo sequences, etc.), process the FID including Fourier transformation and plot and/or display the result. Data reduction to give peak position print-out and integration is also standard. The next stage of development involves the use of a backing-store device (today usually disc-based) to store spectra and programs and this requires the addition of the capability that the FTNMR control segment can create files onto, and read files from the backing store. With a disc-based system it is then possible to remove the computer control segment from the FTNMR program and let it stand alone as an executive program. Only the executive or monitor program is now permanently resident in memory and can read from backing store any program which the user requires. This may be a spectrum simulation and fitting program or any routine for general calculation. Sometimes PASCAL, BASIC interpreters, FORTRAN compilers, assemblers and editors are also available as are a number of computer games to while away the time until lunch!

With the development of larger minicomputer memories came the possibility of performing two separate operations (e.g. acquisition plus FT) at the same time. This time-sharing, or foreground–background working, using different memory blocks has the penalty that the foreground job is slowed down somewhat but that the background job, e.g. a continued acquisition, can continue concurrently. Up to this point, processing of NMR data precluded the acquisition of another spectrum and such time-shared working overcomes this. Alternative approaches are to use a microprocessor controlled acquisition scheme or to use a separate processor for the data collection, both operating in the background whilst data processing can be at the control of the operator. Most commercial spectrometer manufacturers now offer advanced schemes which allow the user to control a number of jobs

concurrently in different memory areas with all parameters such as pulse regime, spectrum width, observe and decouple frequencies, decoupler power, sample temperature and field homogeneity controls under computer control. Automated sequences of experiments are possible with updated values of a number of parameters taken from parameter lists.

The next major development allowed the use of a disc backing store as a "virtual memory" for the acquisition, processing and output of spectral files larger than the computer memory. This opened the way for very large transforms on data tables up to say 512k words, and similarly acquisitions into data tables many times larger than the available memory.

With the development of a disc-based operating system this has recently allowed the operation of the most elaborate FTNMR computer system to date. In this mode of operation, the computer memory can be split into three parts each of which is assigned a job number. Similarly within each job the memory can be further subdivided and in each of these subdivisions a separate operation can be performed concurrently. Each operation is not limited to the memory available to it but is read to and written from the backing disc as necessary.

For example in a 40k memory computer, it is possible to acquire concurrently a 64k data table and transform two previously acquired 64k FIDs. The options for effective simultaneous processing appear to be limited only by the number of peripheral keyboards, VDUs and plotters available and by the time penalty of what is still a time-sharing system.

Finally a number of aspects of the dedicated NMR computer are being changed by the advent of microprocessors. For example the sine look-up table for the Fourier transform in one computer is held in ROM (read-only-memory). In another, the magnet shim settings for each probe are held in a microprocessor.

These rapid developments in computer software have extended the scope of high resolution NMR and these advances will hopefully in the near future be applied to solid state NMR, ESR and other spectroscopic techniques.

2. DIGITISATION

2.1. *Signal-to-Noise Ratio*

2.1.1. *Signal Averaging.* The necessity of digitisation is a direct consequence of the search for a way to improve the sensitivity of the NMR technique. This is most readily achieved by summation of successive scans which reduces the level of the noise relative to that from coherent signals. The spectrum sensitivity is conventionally defined by the signal-to-noise ratio which is measured from a frequency spectrum as the amplitude of a given signal divided by the r.m.s. noise level. When this ratio is inadequate for spectrum interpretation it can be improved by summing scans in a superimposable fashion. A coherent signal will increase linearly with the number of scans and since noise is random, its amplitude increases only with the square of the number of scans. The ratio of signal-to-noise after n scans to that of a single scan is therefore

$$S/N \propto \sqrt{n} \tag{1}$$

The signal detected by the receiver of a pulse Fourier transform NMR spectrometer is a decaying, continuously varying voltage containing both the nuclear responses and noise. The noise can arise from a variety of sources,[11] the probe, receiver amplifier and analog-to-digital converter (ADC). These latter two sources become important when considering averaging of small signals in the presence of large ones. In this case with the amplifier output turned down to avoid overflow most of the noise can come from the amplifier and the ADC, and improving the probe noise figure achieves nothing.

Signal averaging is particularly important in NMR because of saturation. Many other spectroscopic techniques allow one to improve the signal-to-noise ratio by scanning slower with appropriate filters, although this does not remove low frequency distortions.

In order to carry out successful signal averaging and to manipulate the result in a computer, the received voltage is converted into a digital form by sampling the signal amplitude at equal time intervals using a stable gating frequency and an analog-to-digital converter.

Successful signal averaging, or more correctly, signal summation, can then be realised provided the following conditions are met:

(i) there is a constant digitisation rate.
(ii) the ADC can recognise both positive and negative signals. i.e. a bipolar ADC is used or in the case of a unipolar ADC, a d.c. offset is added and this can be subtracted later by software.
(iii) the receiver phase must be constant.

In the discussion which follows on the experimental aspects of digitisation a number of assumptions are made and these are stated explicitly here.

(i) Except when otherwise defined, the signal-to-noise ratio is taken to mean the ratio of the signal height to the peak-to-peak noise height. For a Gaussian distribution of noise amplitudes the peak-to-peak value is strictly infinity but operators have their own way of defining noise and the peak-to-peak value is usually about $2\frac{1}{2}$ to 5 times the true r.m.s. level.
(ii) An inversion technique is used at the digitisation stage to eliminate the effects of any d.c. offset in the analog signal. Such techniques of course are also designed to remove echo effects in single phase detection,[12] ghost peaks in quadrature detection[13] and coherent noise arising from switching signals in the ADC and spectrometer.
(iii) A single scan just fills the analog-to-digital converter. Of course, on a commercial spectrometer if one wishes to use 90° pulses, this may

not be possible since many receiver amplifiers have gain settings in powers of two.

2.1.2. *Signal-to-Noise in Time and Frequency Domains.* In general the signal-to-noise ratios in the two domains for a single peak are not equal. It is easy to form a definition of signal-to-noise ratio in the frequency domain where for a single resonance the peak height can be measured relative to the peak-to-peak noise value. For random white noise the observed peak-to-peak noise value is somewhat arbitrary since the distribution function extends to infinity; however, the r.m.s. deviation of the noise is calculable and the measured peak-to-peak level is about five times this value. In the time domain, in principle, the signal could extend throughout the FID but as a definition let us assume that it has decayed to within the noise level well before the end of the acquisition period. Then the signal-to-noise can be visualised as the height of the FID at the start relative to the noise level at the end of the acquisition period ($t = T$). The true sensitivity in the time and frequency domains is really a measure of the energy content of the signals and in the time domain is equal to the square of the signal voltage summed over the total acquisition. Similarly the true sensitivity in the frequency domain is the corresponding area but for purposes of this discussion in order to discover the extra dynamic range possible in the frequency domain we will describe signal-to-noise ratios by peak heights.

In an empirical sense the signal-to-noise in the frequency domain, S_f, and that in the time domain, S_t, will not be equal and will be related by the resonance linewidth, the acquisition parameters and errors introduced in the Fourier transform.

However some information can be predicted from simple qualitative arguments. The height of the FID at $t = 0$ will be a true measure of the signal area and will be independent of the linewidth, $[=(\pi T_2^*)^{-1}]$. However in the frequency domain, using our definition, S_f will be inversely proportional to the linewidth and hence S_f/S_t will also be directly proportional to the T_2^* of the resonance. Thus S_f/S_t may be much greater than unity for a sharp line.

Becker *et al.*[7] have shown that the signal-to-noise ratio in a FID is

$$S_{\text{FID}} = \frac{M_0(T_2^*)^{1/2}}{\bar{v}} \left(\frac{T_2^*}{T}\right)^{1/2} (1 - \exp(-T/T_2^*)) \quad (2)$$

where $\bar{v}$ is the r.m.s. noise level and T is the acquisition time. If an exponential filter with a time constant of (T_2^*/n) is applied this will reduce the noise at the right hand end of the FID relative to the signal (see Section 3.2) and then the time domain signal-to-noise ratio, S_t becomes:

$$S_t = \frac{M_0(2nT_2^*)^{1/2}}{\bar{v}(n+1)} \frac{(1 - \exp(-(n+1)T/T_2^*))}{(1 - \exp(-2nT/T_2^*))^{1/2}} \quad (3)$$

The signal-to-noise ratio in the frequency domain, S_f, relative to that in the FID, if a matched filter is applied (i.e. $n = 1$) prior to Fourier transformation, is

$$S_f/S_{\text{FID}} = \left(\frac{FT_2^*}{2}\right)^{1/2} (1 - \exp(-T/T_2^*))^{1/2} \quad (4)$$

where F is the spectral bandwidth.[7]

Using eqn (4) for a single phase detector with a FID consisting of a single line, off-resonance, the relative signal-to-noise ratios can be quite large. For example, if $F = 1000$ Hz and $T_2^* = 1$ s then $S_f/S_{\text{FID}} \simeq 22$.

As a practical demonstration, we have synthesised FID's corresponding to different linewidths and added noise to them such that $S_t = 8:1$ for peak-to-peak noise. The parameters used were 8192 data points, which if a spectral width of 1000 Hz is assumed gives an acquisition time of 4.096 seconds, and the peak was placed in the centre of the spectral region. We have assumed single phase detection and the measured values of S_f are given in Table 1 clearly showing a proportionality with T_2^*.

TABLE 1. Relative signal-to-noise ratios in time, S_t, and frequency, S_f, domains for various linewidths ($\Delta\omega$)

T_2^*(s)	$\Delta\omega$(Hz)[a]	S_t	S_f	S_f/S_t[b]	S_f/S_t(quad)[c]
2.048	0.155	8	(128)[d]	16	45.3
1.024	0.311	8	64	8	22.6
0.512	0.622	8	32	4	11.3
0.256	1.243	8	16	2	5.6
0.128	2.487	8	8	1	2.8
0.064	4.974	8	4	0.5	1.4

[a] $\Delta\omega$ is the full width at half height.
[b] This column contains the measured value for single phase detection with the signal off-resonance.
[c] This column contains the predicted values for quadrature detection with the signal at the carrier; i.e. the FID would have half the amplitude on resonance and the detection system also gives a $\sqrt{2}$ gain.
[d] This value cannot be measured because for a $T_2^* = 2.048$ seconds and an acquisition time of 4.096 seconds serious truncation effects would be observed. (see Fig. 12 for an example of the distortions introduced).

The dynamic range observable in the frequency domain can be several powers of two higher than in the time domain and allowance must be made for this when considering computer word lengths for Fourier transformation (we shall return to this in Section 4).

2.2. *Hardware Requirements*

2.2.1. *Sample-and-Hold.* When directly digitising the non-static, analog FID from an NMR receiver, serious errors occur because the signal changes markedly within the time required for the analog-to-digital conversion (see Section 2.2.3 for a calculation which gives the maximum allowed rate of change of an input signal before errors occur). In order to overcome this problem use is made of a sample-and-hold device whose function is to "sample" the analog signal and then "hold" the value constant for the duration of the analog-digital conversion. This is accomplished by charging a capacitor during the sampling period. A perfect sample-and-hold device would, upon the com-

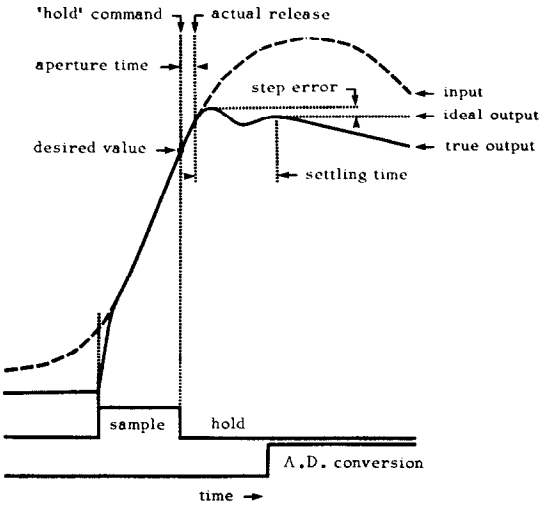


FIG. 1. Definition of the terms used to describe the operation of a sample-and-hold device. A full explanation of the sources of error is given in the text.

mand to sample, faithfully follow the input signal and on the hold command store the input signal at the instant the command was given. In practice, however, small delays and errors can occur, which, along with the principle of operation, are illustrated in Fig. 1.

On receiving the "sample" command the capacitor starts charging, but takes a finite time to track the input signal. This "acquisition time" is typically a few tens of nanoseconds which, in general, is only a small fraction of the overall sampling time. Usually the sampling period is allowed to occupy all the available time from the previous analog-digital conversion thus guaranteeing insignificant tracking errors.

The hold command does not take effect instantaneously and the time for the switch to operate (aperture time), is typically less than 50 nanoseconds. Since this value is constant within very close limits it effectively inserts a small time delay across the whole FID and can be considered as a negligible increase to the pre-acquisition delay normally inserted prior to data collection to allow amplifiers etc. to recover after the rf pulse.

After the "hold" command has taken effect the output voltage requires typically about one microsecond to settle. During this settling time the output voltage changes from its initial value, the difference being a step error. From this point the capacitor slowly discharges at a constant rate (droop rate) and it is during this period that the analog-to-digital conversion is carried out.

For these errors to be insignificant their combined effect must be less than the potential resolution of the digitiser for the duration of the conversion. Since good quality sample-and-hold devices have droop specifications of typically a few microvolts per microsecond and analog-to-digital conversions require only a few microseconds, this condition is easily satisfied, and errors from this source are insignificant.

The overall accuracy of a sample-and-hold device is typically about 0.01% or $1:2^{13}$ (but when signal averaging in the high dynamic range case and using an ADC of 16 bits, then a higher precision is also necessary on the sample-and-hold device (see section 2.5.5).

2.2.2. *Analog-to-Digital Conversion.* The output of the sample-and-hold circuit is fed to the input of the analog-to-digital converter (ADC). In modern spectrometers because of the high sampling rates necessary, this will be of the successive approximation type. Its mode of operation is shown schematically in Fig. 2. It works by comparing the input voltage with a voltage generated by the computer to be half full-scale and if the input is greater than this then the most significant bit in the ADC word is set.

The computer then generates a voltage corresponding to three quarters or one quarter full-scale according to whether the most significant bit was set or not and this is compared with the input voltage and the next most significant bit is set or not according to whether the input is greater or less than the generated value. The process is repeated for each bit of the ADC word. The detailed operation for a 10 volt full-scale, 5 bit ADC is as follows:

The most significant bit (bit 4) is turned on and the voltage corresponding to this level (5.000 volts) is obtained from a digital-to-analog converter (DAC). This is compared with the input signal using a comparator circuit, and if the input voltage is less than 5.000 volts the bit is switched off; if the input voltage is greater than 5.000 volts, bit 4 is left on. Next bit 3 is turned on and its corresponding voltage (2.500 volts),

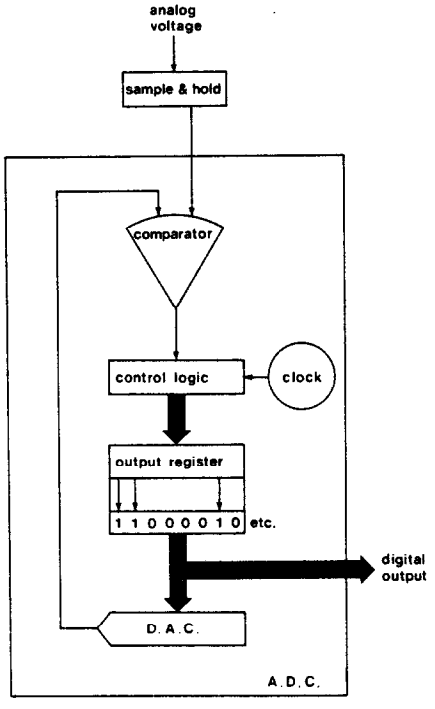


FIG. 2. Diagrammatic representation of the mode of operation of a successive approximation ADC.

added to the previous 5.000 volts if bit 4 was left on, is compared with the input and again bit 3 is turned off or left on according to whether the input is less than or greater than the generated voltage. Next bit 2 is turned on and its voltage (1.250 volts) is taken from the DAC, added to that generated from any previous bits left on and compared with the input voltage. This process is repeated down to bit zero when the generated voltage will correspond to half the scale for one bit (in this case 0.3125 volts), and thus it is possible to distinguish an input level greater or less than the voltage corresponding to *half a bit*, and for this reason the possible resolution of an N bit successive approximation ADC is about $1:2^{N+1}$.

2.2.3. *ADC resolution and error sources.* This review is not the place for a detailed discussion of the hardware limitations of ADCs, but a few of the usual sources of error may be listed.

(i) Quantisation error: the minimum deviation of the digital result from the analog input and is only achieved in the absence of effects arising from non-linearity etc. At best this can be $\pm\frac{1}{2}$ LSB (the least significant bit).

(ii) Linearity: a measure of the extent to which the output code of bits versus the input voltage differs from a straight line (or actually a staircase). If the linearity is worse than $\pm\frac{1}{2}$ LSB, adjacent code levels may not change.

(iii) Non-static input: an additional error can arise if the input changes during a conversion. A successive approximation ADC chooses a value for the most significant bit at the beginning of the conversion and hence can be in error by the total amount that the input has changed between then and the final least significant bit step. If the input is to be converted to an accuracy of $\pm\frac{1}{2}$ LSB then it must not change by more than the voltage corresponding to this amount during the total conversion and this problem is overcome by using the sample-and-hold device. It is interesting to carry out a calculation to determine the maximum frequency which can be digitised without a sample-and-hold. Consider a 10 bit, 10 volt full scale converter requiring 10 μsec per conversion. This will give an error greater than $\pm\frac{1}{2}$ LSB if the input changes by more than $\pm\frac{1}{2}$ $(10/2^{10}) \simeq \pm 5$ mV in the 10 μsec required for the conversion. That is, conversely, a rate of change greater than 5 mV/10 μsec will give an error greater than $\pm\frac{1}{2}$ LSB. The maximum rate of change of a sine wave of amplitude V_{max} and frequency ν occurs as it passes through zero.

$$\left(\frac{dV}{dt}\right)_{max} = 2\pi\nu V_{max} \tag{5}$$

$$\nu = \frac{1}{2\pi V_{max}}\left(\frac{dV}{dt}\right)_{max} \tag{6}$$

Substituting for a 10 volt sine wave gives a maximum frequency of ~ 8 Hz.

(iv) Resolution: as previously stated, in the absence of other errors or noise, the analog voltage is defined to one part in 2^{N+1} for an N-bit ADC. However, in a real spectroscopic case, signals smaller than this fraction can be digitised correctly either because the random nature of the noise added to a weak signal can break over the level required to set the least significant bit in the ADC (let us define this as one count) or in the presence of a strong signal, the small signal may modify the bits which are set in the ADC even though the absolute value of the small signal corresponds to much less than half the least significant bit. (In this case if the noise is less than one count then the digitised result for the small signal may be in error, see Section 2.6.1 and Figs. 7 and 8).

As a general rule as the ADC resolution increases the maximum sampling rate falls. Thus a 12 bit ADC may be capable of operating at 330 kHz but a 16 bit device may be limited to 50 kHz. But, when it is necessary to detect very small signals in the presence of large sample or solvent peaks, i.e. using high resolution ADCs, these may be the very situations in which wide sweep widths are required (e.g. for ^{19}F NMR), limiting the ADC resolution by virtue of the high sampling rates required. This problem is not alleviated by the use of quadrature detection, where two detectors 90° out of phase are used and the spectral width for each is only half that of single phase detection. This arises because although two detectors are used, it is conventional for the output signals from both to be digitised through a single ADC multiplexed to both channels so that as far as the ADC is concerned the same sampling rate is required for single and quadrature phase detection.

One possible method of overcoming this problem, so far not adopted by NMR manufacturers, but used in high dynamic range mass spectrometry, is to use two 12 bit ADCs overlapped by 8 bits. Here the NMR signal would have to be divided into two parts differing in amplitude by a factor of 16, each part being fed to a separate ADC. However, it would be necessary to ensure that both signals had identical phases immediately prior to digitisation. The two ADCs can now be termed as having high and low sensitivity. The high sensitivity device could be used for all conventional applications but in situations of high sensitivity and high dynamic range the top four bits from the low sensitivity ADC could be used to produce a 16 bit digitised result. The reason for overlapping the ADCs is to prevent the half least-significant-bit error appearing in the computer word from the low sensitivity device. The advantage is clearly one of conversion speed and hence an expansion in the maximum usable frequency width. In addition it may be simpler to interface two ADCs in this way rather than add the complication of switching two different ADCs with

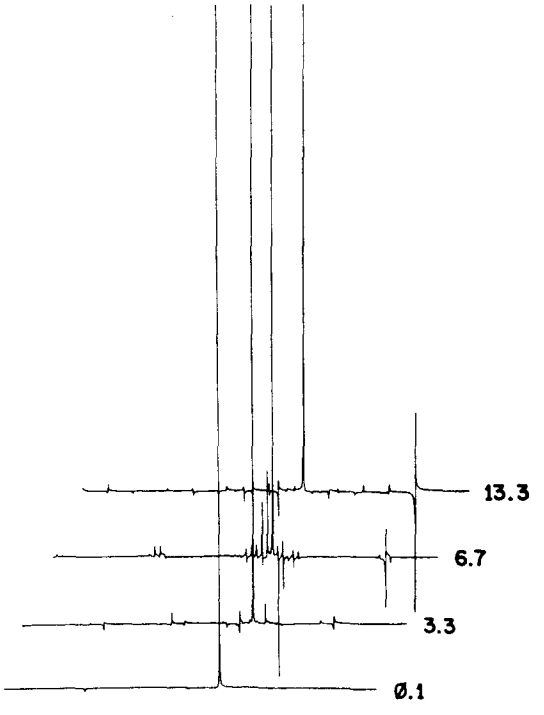


FIG. 3. The effect of timing errors on the digitisation of a single frequency. The diagram shows the result of Fourier transformation after sampling a 2500 Hz sine wave at 100 μs per point with a random timing error of $\pm$ the number of microseconds shown. (Reproduced from *Computers and Chem.* **1**, 55 (1976) with permission of the copyright holder).

entirely different characteristics for the different applications.

2.3. *Sampling Rates and Timing*

The speed at which the voltage must be sampled is defined by the Nyquist theorem and to prevent "aliasing" or "foldover" this must be at least twice the highest frequency of interest. Most textbooks on FTNMR carry a diagram which illustrates this principle.[14]

The constancy of the sampling rate is also most important because if this varies it will distort the frequency information in the final spectrum giving noise and/or harmonics. Figure 3 shows the result of sampling a 2500 Hz sine wave every 100 μsec with a random timing error of ± 0.1, ± 3.3, ± 6.7 and ± 13.3 μsec. The Fourier transform of the digitised result clearly shows a great deal of spurious information as the timing jitter is increased.[15]

2.4. *Quantisation Errors*

Any digitisation of an analog signal will only be an approximation to the truth and artefacts will be introduced simply by Fourier transformation of a digitised decaying pure sine wave, and we give here an analysis of the types of distortion which can occur and how serious they are. Two sources of such frequency artefacts are possible if we ignore, for the present, what happens if the signal of interest is less than one count in the digitiser (i.e. the voltage necessary to set the least significant bit).

If there is an exact number of data points per cycle of the sine wave any misrepresentation of the signal voltage will occur at integral multiples of the signal frequency and therefore give rise to signals at higher harmonics.

If, as in general, there is a non-integer number of data points per cycle, errors will have a periodic nature and will reappear at periods of $N(f - f_0)$, where f is the FID frequency and f_0 is the nearest frequency which has an integral number of data points per cycle. The latter effect is illustrated in Fig. 4, which shows the distortion introduced by placing a large signal at various offsets from the centre of the frequency spectrum. Clearly the solution, at least for a single scan, is to place the carrier frequency such that the large peak appears exactly at the centre of the spectrum. Then for single phase detection there will be exactly four data points per cycle, removing the second cause of distortion, since by definition $(f - f_0)$ is zero. Also the first harmonics will appear at the edges of the spectrum removing the first kind of distortion.

These two sources of quantisation errors will have an effect at any given ADC resolution, but it is important to realise that both will get progressively worse as the ADC resolution is decreased.

Figure 5a shows that even for a single peak in the centre of the spectrum, quantisation errors increase as the ADC resolution is lowered, the increased uncertainty in the digitisation of the input voltage being manifested as an increase in the noise level being worst close to the peak. If the signal is placed away from the centre of the spectrum then the harmonics and other distortions as in Fig. 4 reappear and these also get progressively worse as the ADC resolution is lowered (Fig. 5b).

The effects of having a non-integer number of data points per cycle will give spurious lines in cases where $(f - f_0)$ is greater than a few Hertz but where $(f - f_0)$ is small an increased noise level is observed.

An important question arises as to why, as is observed, the harmonics and other peaks, and of course the quantisation induced "noise", all are reduced by multi-scan averaging. If all the scans were truly identical with regard to *signal* then surely the quantisation errors would be the same? The many scans added together in a normal experiment are not all the same because small changes in d.c. offset, changes in pulse width, or any very low frequency component ensures that the *signal* digitisation is different for each scan. For example using a 70° pulse and a 12 bit digitisation a 0.1° change in the pulse angle causes, at a minimum, the least significant bit to be altered. Also another source of variability in a real spectrum will arise through small changes in spinning speed during a long acquisition. Figure 8 shows such a decrease in quantisation errors as a result of multiple

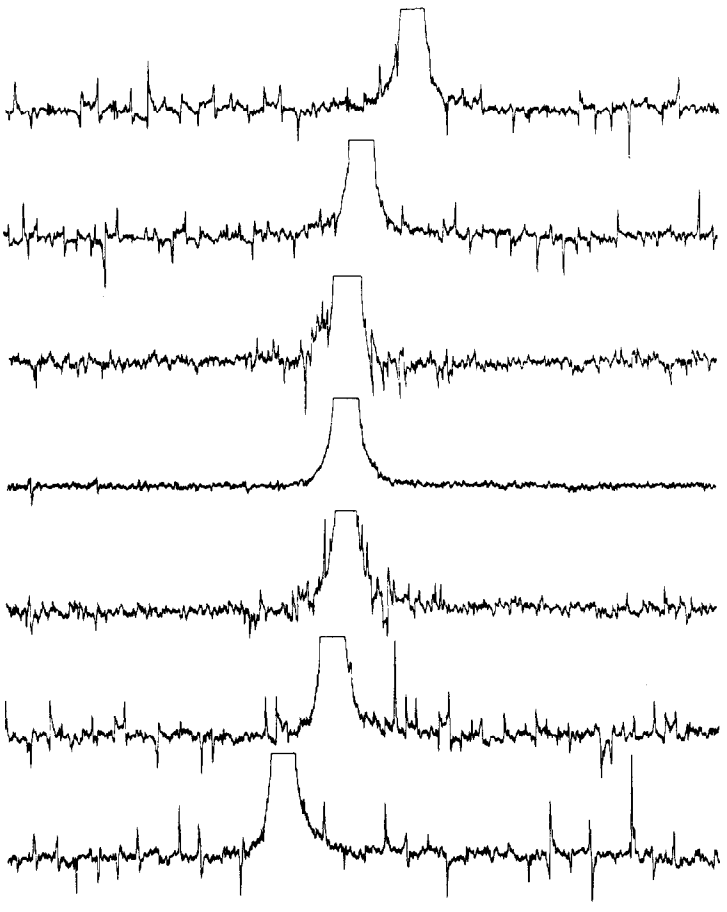


FIG. 4. The quantisation distortions introduced by placing a signal at various offsets from the centre of the spectral window. The offsets are, from the top, −25 Hz, −5 Hz, −1 Hz, 0 Hz, +1 Hz, +5 Hz and +25 Hz. Spectral width 1250 Hz, 8k data points with a plotted width of 275 Hz.

scan averaging. In addition, use of phase alternating pulse sequences[12] brings additional variability to a multi-scan experiment. Simply adding noise to an already digitised signal cannot affect the quantisation errors, but merely degrades the signal-to-noise ratio.

Clearly the more stable the spectrometer system, the less will harmonic and other quantisation errors be reduced by multiple scan averaging. The problem is greater at highest fields where, for superconducting magnets, the stability is high and more importantly the sensitivities are up to twenty times higher. Because of this decrease in noise, higher resolution ADCs are a pre-requisite.

2.5. *Signal Overflow*

2.5.1. *Introduction.* Since it is signal summation which is normally carried out, eventually at least one computer word in the data array will contain all ones and thus any further signal input would cause overflow. This cannot be allowed to happen in Fourier transform NMR data systems since the contents of any one location in the time domain spectrum will affect the appearance of all channels to some extent after Fourier transformation. The effect of allowing overflow to occur in the FID from a single line is shown in Figure 6. This shows the result after Fourier transformation of allowing the synthesised FID from an exponentially weighted sine wave to overflow at the beginning for the number of addresses shown.

2.5.2. *Possible number of scans.* For a noise-free spectrum the number of scans which can be added before overflow occurs depends on the resolution of the ADC and the computer word length. For a computer with a word length w and an ADC with a resolution of N bits then this maximum number of scans is 2^{w-N} and values are given in Table 2 for various w and N values.

Cooper[15] has calculated the possible number of scans which result as a consequence of allowing the spectrum to have a finite signal-to-noise ratio. For an initial signal-to-noise ratio of s counts of signal to one count of noise, the total input is $s + 1$ counts and therefore the proportion of signal to fill an N-bit ADC is $2^N s/(s + 1)$. The total number of scans T to fill memory will give a signal $T2^N s/(s + 1)$ and a noise value of $T^{1/2}2^N/(s + 1)$. If the maximum number in a computer word of length w is 2^w then when total signal plus noise is equal to 2^w overflow will occur.

$$2^w = 2^N\left\{\frac{Ts}{s+1} + \frac{T^{1/2}}{s+1}\right\} \tag{7}$$

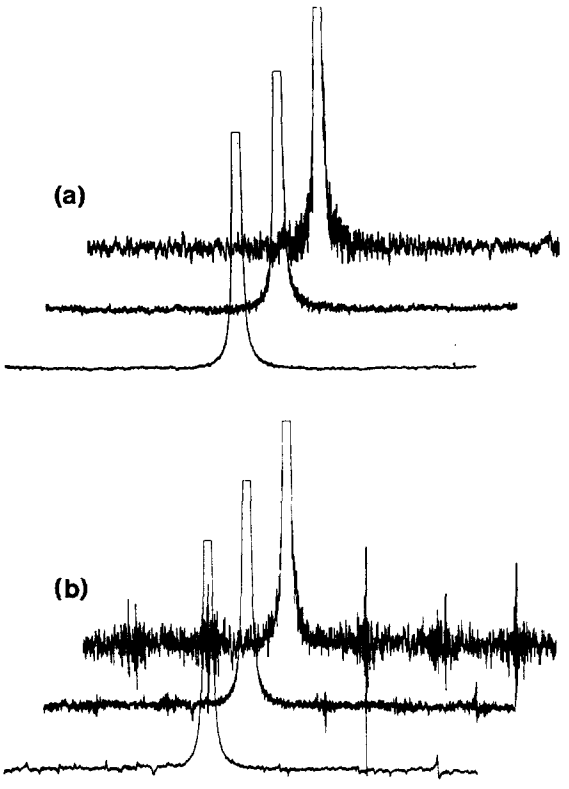


FIG. 5. The effect on the quantisation errors of reducing the ADC resolution for a signal placed (a) at the centre of the spectrum and (b) at an offset of −50 Hz.

$$\sqrt{T} = \frac{-1 + [1 + 4s(s+1)2^{w-N}]^{1/2}}{2s} \qquad (8)$$

This result has been calculated by Cooper[15] for various initial signal-to-noise ratios and $w - N$ values and these are given in Table 3.

In this table the initial signal-to-noise ratios will also depend on the ADC resolution because of quantisation errors and the final signal-to-noise ratios are those pertaining to the time domain.

For a very low initial signal-to-noise ratio such as in ^{13}C NMR it has not been considered necessary to use such a high resolution ADC as 12 bits and just a few bits have been suggested as sufficient. However this is an over-simplification because even in the presence of noise any coherent signal will still be subject to quantisation errors (Section 2.4). Also the uncertainty from the half least significant bit of the ADC will be present and will be a significant proportion of the total.

As is clear from Fig. 6 it is necessary to be able to detect overflow just before it occurs otherwise increased signal averaging just leads to a greater number of artefacts. Commercial NMR spectrometers are programmed to test for overflow and to halt acquisition if overflow is imminent.

If a satisfactory signal-to-noise ratio has been obtained prior to overflow then there is not any problem and a Fourier transformation can be performed, but it is interesting to note some of the computer parameters which may be required in a real case. Cooper[15] has performed this calculation for the situation which may occur in ^{13}C NMR and here a total scan time of three days is about the maximum which anyone would devote to one spectrum, and if a transient is recorded every second then about 2^{18} scans would be added. Thus 2^{18} scans gives a 2^9 improvement in signal-to-noise and for a final value of 4:1 in the time domain the initial signal-to-noise is about 2^{-7} or 0.0078. Substituting into eqn (8) gives $w - N \simeq 11$. This implies a maximum of a 5 bit ADC for a 16 bit computer word but this maximum may be too low for the reasons given above. However if, with the computer word full, the final signal-to-noise ratio in a spectrum is not good enough a number of options are possible and these are described in the next paragraphs.

2.5.3. *Scaling of memory and ADC.* When overflow is about to occur, one method of enabling the acquisition to continue involves dividing the memory contents, the incoming signal and the ADC resolution by two. This process can be repeated each time

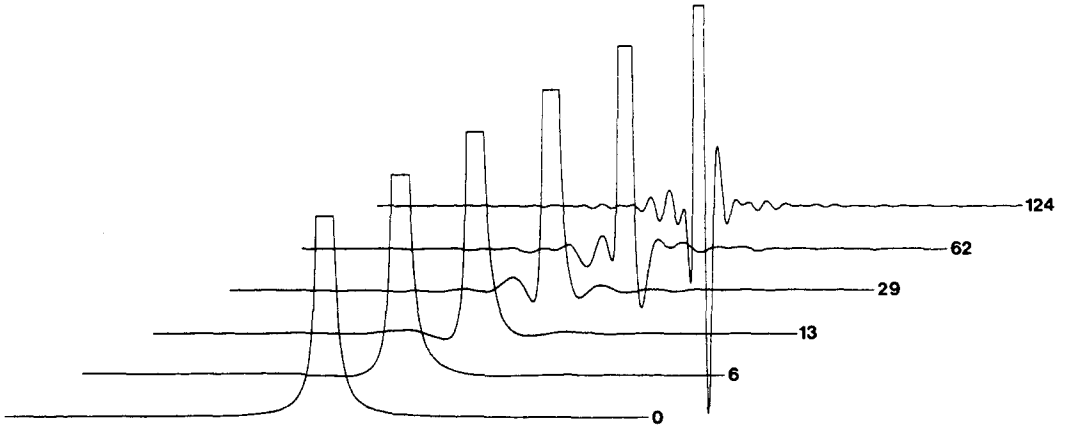


FIG. 6. The effect after Fourier transformation of allowing overflow by the number of channels shown at the beginning of the FID. The peak was generated by exponentially weighting a digitised sine wave and plot scaling was adjusted to give identical areas in all cases.

TABLE 2. Maximum number of scans for various computer word lengths and ADC resolutions assuming a noise-free signal

ADC resolution (bits)	Computer word length (bits) 16	20	24	32
6	1024	16384	262144	67108864
8	256	4096	65536	16777216
10	64	1024	16384	4194304
12	16	256	4096	1048576
14	4	64	1024	262144
16	1	16	256	65536

overflow is imminent until an ADC resolution of about 6 bits is reached. The method works for cases of low dynamic range but its main disadvantage is that in general small signals in the presence of large ones will eventually be inadequately digitised because the modulation caused by a small signal on a large FID will not be detected above the quantisation error. Also it is true that quantisation errors will increase as the acquisition proceeds.

Let us consider an example of an FID with an initial signal-to-noise ratio of 0.1 digitised using a 12 bit ADC into a 16 bit computer word. From Table 3 it is seen that 84 scans may be summed before overflow occurs at which point the signal-to-noise ratio in the time domain will be $\sqrt{84} \times 0.1 = 0.91$, clearly not adequate. The memory, input, and ADC are all divided by two; now only 15 bits of the computer word will be full and the number of scans required to fill the last bit is half that required to fill an empty word, and since now $w - N = 5$, a further 104 scans can be accumulated giving a signal-to-noise ratio of $0.1 \times \sqrt{188} = 1.37$. This process can be repeated until the required signal-to-noise is obtained. In the example given here, scaling to a 10 bit ADC will occur after 188 scans giving 1.37 as a signal-to-noise figure, scaling to 9 bits after 430 scans giving 2.07, scaling to 8 bits after 969 scans giving a signal-to-noise of 3.11, to 7 bits after 2135 scans giving a figure of 4.62 and to a 6 bit ADC after 4600 scans giving a final signal-to-noise ratio of 6.78.

For a case with an initial signal-to-noise ratio of 0.1 the noise will ensure that even for very low ADC resolution the signal will appear on top of the noise even though the full square root gain will not be obtained because the noise from the half least significant bit error is an appreciable part of the total. However, if the actual division of the input is accomplished by right shifting data digitised by a high resolution ADC, retaining the amplifier settings, then this method removes any uncertainty caused by the half least significant bit error of the ADC which would otherwise increase as the ADC resolution is decreased.

2.5.4. *Normalised averaging.* Cooper[15] has described the method of normalised averaging for preventing memory overflow. If the average signal in the memory after p scans is A_p and the signal from each scan is s_i then

$$A_p = \frac{1}{p} \sum_{i=1}^{p} s_i \tag{9}$$

$$A_{p-1} = \frac{1}{p-1} \sum_{i=1}^{p-1} s_i. \tag{10}$$

Multiplying eqn (9) by p and eqn (10) by $(p-1)$ and subtracting gives

$$s_p = pA_p - (p-1)A_{p-1}. \tag{11}$$

TABLE 3. Maximum number of scans and final signal-to-noise ratios for various ADC resolution (N), computer word lengths (w) and initial signal-to-noise ratios (S/N).

	$S/N = 0.01$		$S/N = 0.1$		$S/N = 1.0$		$S/N = 10.0$	
$w - N$	Scans	Final S/N	Scans	Final S/N	Scans	Final S/N	Scans	Final S/N
4	200	0.14	84	0.92	26	5.2	17	41.5
5	661	0.26	207	1.44	56	7.5	34	58.8
6	1996	0.45	484	2.20	117	10.8	69	83.4
7	5507	0.74	1079	3.28	240	15.5	139	118.2
8	14016	1.18	2332	4.83	489	22.1	279	167.3
9	33428	1.83	4929	7.02	992	31.5	560	236.8
10	75878	2.75	10251	10.12	2003	44.8	1123	335.1
11	166093	4.08	21076	14.52	4032	63.5	2248	474.1
12	354182	5.95	42982	20.73	8101	90.0	4498	670.7
13	741293	8.61	87159	29.52	16256	127.5	9001	948.8
14	1531049	12.37	176028	41.96	32587	180.5	18008	1342.0
15	3132577	17.70	354494	59.54	65280	255.5	36025	1898.0
16	6366811	25.23	712455	84.41	130710	361.5	72062	2684.5

By subtracting A_p from both sides, and dividing by p gives

$$A_p = \frac{s_p - A_{p-1}}{p} + A_{p-1}. \quad (12)$$

Thus, for normalised averaging, the current average is subtracted from each scan and this resultant divided by the number of completed scans is added to the previous average. Usually the division is achieved by right shifting the computer word and the number of bits shifted corresponds to a division by that power of two and for convenience eqn (12) becomes

$$A_p = s\frac{s_p - A_{p-1}}{2^M} + A_{p-1}, \quad (13)$$

where M is the first power of two greater or equal to p.

Using this method overflow cannot occur, but the number of scans which can be added is 2^N (where N is the number of bits in the ADC) since as soon as $p > 2^{N-1}$, 2^M becomes 2^N making all differences $s_p - A_{p-1}$ equal to zero. Also as M increases, the detectable dynamic range is reduced so this method cannot be used for the high dynamic range situation.

2.5.5. *Double length averaging.* The method which enables the continuation of averaging on most commercial spectrometers is that described in Section 2.5.3, namely scaling of memory and ADC. However even this method will not always allow enough scans to be accumulated to give an adequate signal-to-noise ratio, and to illustrate this, let us choose a specific example. Assume that we have a spectrum with an initial signal-to-noise ratio of 0.1 : 1 in the time domain and that we can use a minimum of a 6 bit ADC, then from Table 3 the maximum number of scans is 10251 with a final signal-to-noise of 10.1 : 1. This time domain signal-to-noise may not be adequate for a particular application which could involve deconvolution with a resolution enhancement function and so it then becomes necessary to continue averaging. This can be done by allowing two computer words for each location and to continue averaging in a double length mode.

One practical method is to dump the FID on backing store, say disc, at intervals defined by Table 2 before overflow occurs and to allow the FID to occupy more than one word per location. It is also possible to accumulate double length in memory or to acquire in a floating point format although this latter method has the disadvantages of difficult programming and the computations necessary may not be possible in real time.

To return to our example let us now assume we need to perform 100,000 scans to give a final signal-to-noise of 31.6 : 1. On substitution of the number of scans and the initial signal-to-noise ratio into eqn (8), $w - N$ is found to be 13.19. This implies a 20 bit word for a 6 bit ADC or a 24 bit word for a 10 bit ADC. If the computer has a 16 bit wordlength then double length averaging is necessary (i.e. using up to 32 bits).

It is not necessary to perform a 32 bit Fourier transformation because at each location the signal value may be in error by ± 7–8 bits as a result of the noise. If, in our example above, the contents of the computer words were shifted to the right to occupy only 16 bits then there would remain ± 5–6 bits of noise at each location and the noise after Fourier transformation would still be very much greater than quantisation noise as a result of defining the signals less precisely by using fewer bits.

2.6. *Signal Averaging in the Presence of Large Signals*

2.6.1. *Number of scans and ADC resolution.* So far we have only considered the digitisation requirements when averaging signals in the time domain which have a low signal-to-noise ratio. If, in that category there is also a large dynamic range, we have seen that special precautions, such as possibly using a double precision acquisition sequence, may be necessary.

Let us now turn to the situation commonly found in FTNMR where there is a high signal-to-noise ratio in the time domain and no dynamic range problem. Here further averaging is only necessary to improve the noise level, possibly to enable efficient deconvolution routines to be used or for quantitative analytical determinations. As long as a high resolution ADC is used to keep quantisation errors to a minimum, normal signal averaging techniques as described in Section 2.5 are adequate.

The major difficulty in FTNMR arises when trying to recover small signals of interest from noise in a high dynamic range situation where the major signal, usually from a solvent resonance, has a very high time domain signal-to-noise ratio. In this case, special consideration must be paid to the ADC resolution, the computer word length and, as described in Section 4.3, the type of Fourier transformation routine used.

An N-bit successive approximation ADC will not detect changes in the input voltage much less than $1/2^N$ and this occurs because of the way in which it operates since any change in voltage less than $1/2^{N+1}$ will not cause the least significant bit to be altered, and voltages up to $3/2^{N+1}$ will only affect the least significant bit. The effect is described as due to *changes* in voltage because in the FID any small signal will appear as a modulation on the large signal, or noise, and this eventually enables the detection of voltage changes due to the small signal less than 2^{-N}. Cooper[15] has analysed the final time domain signal-to-noise ratios possible for various ADC resolutions (N) and computer word-lengths (w). Since a large signal is present in the high dynamic range case only 2^{w-N} scans are possible before overflow and hence the final signal-to-noise ratio will be that initially, multiplied by the square root of the number of scans up to 2^{w-N}, provided the dynamic range of interest is not greater than twice the digitiser resolution. Defining the signal intensities for the large and small signals as S_L and S_S, although S_S will appear as a modulation on S_L, then the final signal-to-noise ratio for the small signal (S_f) is

TABLE 4. The final signal-to-noise ratios and the maximum number of scans in the high dynamic range case for various values of the small signal intensity. (S_L = 1000 in all cases)

w[a]	N[b]	Max no. scans	S_t[c] $S_S = 10$	$S_S = 1$	$S_S = 0.1$	$S_S = 0.01$
16	6	1024	320	—	—	—
	9	128	113	11.3	—	—
	12	16	40	4	—	—
	16	1	10	1	0.1	0.01
20	6	16384	1280	—	—	—
	9	2048	453	45.3	—	—
	12	256	160	16	—	—
	16	16	40	4	0.4	0.04
24	6	262144	5120	—	—	—
	9	32768	1810	181	—	—
	12	4096	640	64	—	—
	16	256	160	16	1.6	0.16
32	6	67108864	81920	—	—	—
	9	8388608	28963	2896	—	—
	12	1048576	10240	1024	—	—
	16	65536	2560	256	25.6	2.56

[a] w is the computer word-length in bits.
[b] N is the ADC resolution in bits.
[c] Values omitted cannot be calculated from eqn (15) because $S_S < 1/2^{N+1}$.

$$S_t = 2^{w-N}S_S, \text{ for } S_L/S_S \leq 2^{N+1},$$

or defining A as

$$A = \frac{-S_L}{S_S} + 2^{N+1} \tag{14}$$

$$S_t = 2^{(w-N)/2}S_S(|A| + A)/2A. \tag{15}$$

The final values for S_t for various values of w and N, assuming S_L = 1000 are given in Table 4.

The values of S_t given in this table are over-estimates because contributions from quantisation errors have been ignored.

It has often been stated that it is impossible to detect signals smaller than those corresponding to half the least-significant-bit in the digitiser. However this conclusion assumes an infinite signal-to-noise ratio in the time domain and it is possible to calculate what occurs for the case of a finite signal-to-noise ratio and to show that noise can increase small signals to detectable levels provided enough scans can be acquired. Ernst[16] pointed this out first and performed the necessary calculations which showed that one count of noise in the ADC will allow the digitisation of signals much smaller. Marchal *et al.*[17] have extended these calculations for the case of a 9-bit ADC with the large/small signal intensity ratio varying up to 2^{10}. Figure 7 shows the results of similar calculations for a 6 bit ADC, used simply for speed of computation.

This diagram shows how, for example for the first location, the small signal intensity will appear after many accumulations for different r.m.s. values of the noise in counts. Two conclusions which can be gained from this figure are that as long as the r.m.s. amplitude of the noise is less than the voltage corresponding to about one count in the digitiser, the value of the small signal voltage at any one location will be distorted (both over- and under-estimates are possible). Secondly, when the r.m.s. value of the noise is greater

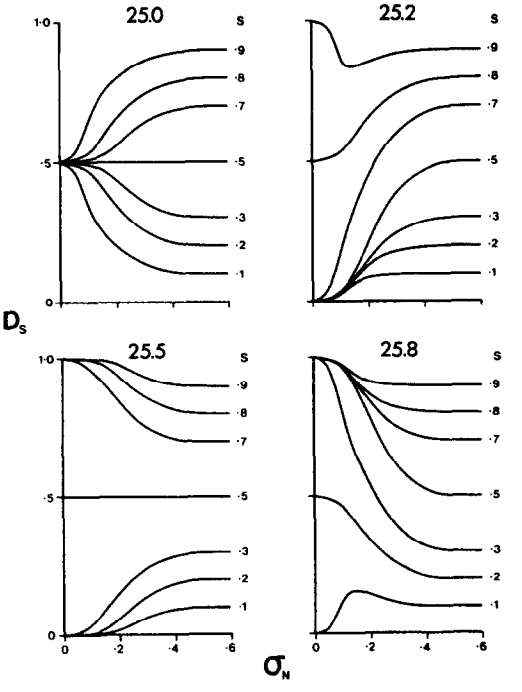


FIG. 7. The effect of noise on the digitisation of small signals (of amplitude S) in the presence of a large signal (of amplitudes 25.0, 25.2, 25.5 and 25.8) using a 6 bit ADC. D_S represents the digitised value of the small signal voltage after many scans taking into account the random nature of the noise with a standard deviation of σ_N counts.

than this level, the small signal is measured correctly. It should be noted that the small signal will eventually be defined adequately regardless of its initial size provided that the r.m.s. level of the noise is above about one count in the ADC. Because of the nature of a real exponential decay the small signal may well correspond to one count at the first location of the FID but it becomes progressively less during the acquisition. The amplitude of the large signal is only important in that it defines the voltage corresponding to one count.

From Figure 7 it can be seen that, at any one location, the value of the small signal voltage will be distorted if the r.m.s. noise level is zero. This does not necessarily mean that the small signal intensity will be distorted after Fourier transformation. This is because it is unlikely that the decays from the large and small signals will have identical time constants and therefore throughout the FID their proportions will change and both over and under-estimates of the small signal voltage may average out. Also, for the completely noise free situation, a small signal of less than one count will be detected by multiple scan averaging because no two scans are identical and this effectively adds in a variability which allows the small signal to be detected (Section 2.4).

An interesting question is whether signals smaller than half of one count in the ADC can be detected after Fourier transformation on a single scan in a noise-free situation. Intuitively it would appear that it is only necessary for the digitised result to be altered in a few locations by the small signal which may be just sufficient to cause a different bit to be set in those locations and hence sufficient frequency and intensity information would reside in the FID. Figure 8 demonstrates that this is so. Trace (a) shows the ^{1}H frequency spectrum of a mixture of acetone and benzene where the linewidths and hence the time constants in the FID are similar and with a ratio of the areas of 368:1. This trace results from Fourier transforming a single scan such that a 12 bit ADC was just filled. The noise in the time domain was less than one count on average arising only from the $\pm\frac{1}{2}$ LSB uncertainty in the ADC. Therefore, in the frequency domain the noise is dominated by quantisation errors and Fourier transform noise. Trace (b) is the result after Fourier transformation of the same FID after shunting the contents of the computer words to the right such that the four least significant bits are lost; now the $\frac{1}{2}$ LSB error has effectively disappeared. At this point the small signal corresponds to 0.7 count in the ADC at the beginning of the FID and progressively less as the FID proceeds. The benzene signal is still present at its correct relative intensity. Traces (c) and (d) show the result of dividing the same FID by a further factor of 4 and 16 respectively. In the latter case the small signal only contributes <0.04 of a count in the ADC and this is only at the first location of the FID. In principle the noise will still be present but will only amount to <0.004 of a count and can be ignored. After Fourier transformation the peak is still observable on a single scan with the correct intensity

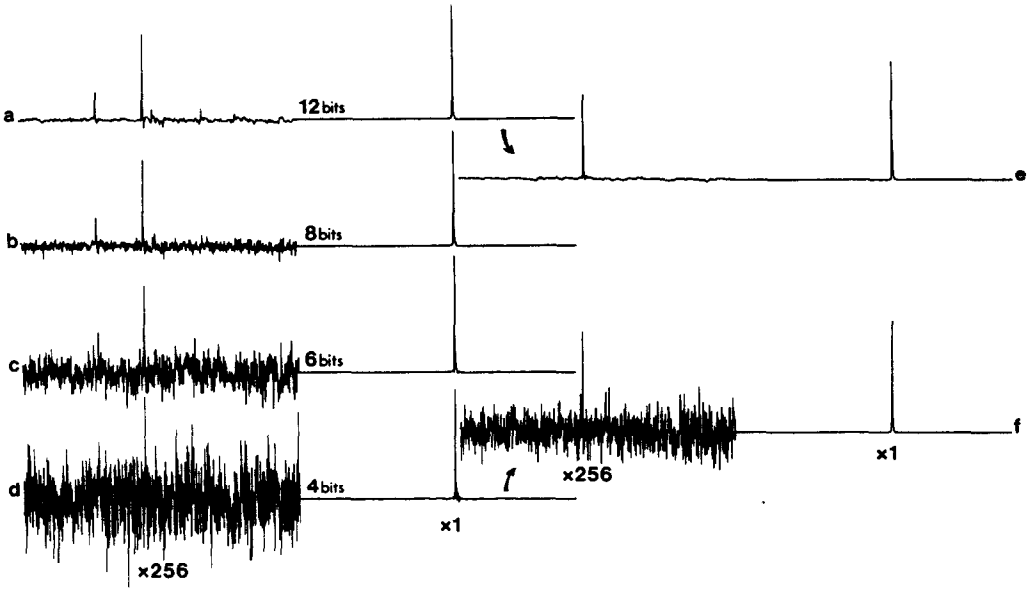


FIG. 8. Demonstration of a large signal causing the digitisation of a small signal in the absence of noise with the small signal intensity <1 count in the ADC. (a) shows the ^{1}H NMR spectrum of a mixture of benzene and acetone; the linewidths are similar and the area ratio is 1:368. This spectrum is the result of a single scan using a 12 bit ADC. The only noise in the time domain arises from the $\frac{1}{2}$ LSB of the ADC. Trace (b) is the result of taking the same FID and dividing by 16 so that the $\frac{1}{2}$ LSB noise effectively disappears. Now the small signal at $t = 0$ represents 0.7 counts. Traces (c) and (d) give the result of further divisions by 4 and 16 respectively. In the latter case the small signal is ≈ 0.04 count at the beginning of the FID. Trace (e) is the result of summing 16 scans using a 12 bit ADC and (f) is the result from transforming the sum of 16 separate noise free FIDs obtained by dividing 12 bit digitised scans to 4 bits.

within experimental error. The progressive increase in quantisation errors can be observed in the series (a)–(d).

Thus it is possible to observe two peaks which have a dynamic range of 368 : 1 at the start of the FID using a "noise free" signal and ADC with a resolution of only 4 bits. The presence of the small signal can be observed more clearly using multiple scan averaging. Trace (e) shows the result of adding 16 scans acquired using a 12 bit ADC and trace (f) the result of summing 16 separate FIDs acquired using a 12 bit ADC and scaled to 4 bits before summation. A gain in the signal-to-noise ratio is observed in both cases although the magnitude of the gain depends on the relative magnitudes of the noise introduced by quantisation errors and by the Fourier transform. The multiple scan results also show the reduction in the quantisation harmonics which are achieved because no two scans are identical.

All calculations so far have relied on determining the signal-to-noise ratio in the time domain. As shown in Section 2.1.2 the signal-to-noise ratio in the frequency domain can be several powers of two higher even allowing for rounding or scaling errors in the Fourier transform. For example a spectrometer, operating at a magnetic field of about $2T$ with quadrature detection, will produce a frequency domain signal-to-noise ratio for 100% H_2O of about 100,000 (or 2^{17}) for a single 90° pulse. From eqn (4) with a spectral width of 1000 Hz and a T_2^* of 1 second this gives a signal-to-noise ratio in the time domain of about 2^{11}. If the large solvent peak is off-resonance the FID amplitude and signal-to-noise ratio is doubled. This means that a 12 bit digitiser may only be barely adequate to give one count of noise. At the highest available magnetic fields, where the signal-to-noise ratio is about 4 bits greater, then the same questions of adequacy apply to a 16 bit ADC. However, in the high dynamic range case when using very high magnetic fields ($\approx$350 MHz) the large resonance may be artificially broadened because of the phenomenon known as radiation damping.[18] Here the large induced signal voltage causes the generation of a back emf which tends to destroy the original signal leading effectively to a shorter T_2^*. This has a further implication when considering the relative signal-to-noise ratios of large and small signals in that it effectively lowers the dynamic range in the frequency domain.

2.6.2. *Spectroscopic methods of improving dynamic range.* It is possible to reduce the demands on computer word length and ADC resolution by a variety of spectroscopic methods. However, almost all either take considerably longer for a required signal-to-noise ratio or alternatively only produce partial spectra.

(i) Solvent peak irradiation giving saturation of the largest resonance.[19]
(ii) Nulling of solvent resonance with a long T_1, such as water, by utilising the inversion-recovery sequence and collecting data when the solvent is just nulled.[20]
(iii) Synthesised excitation whereby the power spectrum of the transmitted pulse is designed not to include the solvent resonance frequency.[21]
(iv) The Redfield soft pulse method, which uses specially designed weak pulses to excite only the resonances of interest and to be nulled over as great a frequency range as possible.[4]
(v) Rapid-scan correlation NMR in which the large peaks are simply not included in the scan.[1,2]
(vi) Use of a notched filter[22] to suppress the detection of a particular frequency.
(vii) Later, we discuss the noise introduced by Fourier transformation of a large signal and show that the rounding error noise is proportional to the number of spectral channels which are full. Therefore placing the large signal near the carrier will reduce the number of cycles which the cosine voltage will pass through before it decays significantly. Therefore this may reduce slightly the amount of noise introduced by the Fourier transformation.[15]
(viii) Finally, a recent method of experimentally overcoming the dynamic range problem is to use the selective excitation technique of a train of very narrow pulses.[23]

2.6.3. *Block averaging.* This is the most widely used technique on commercial instruments to enable the continuation of signal averaging once overflow is imminent. The FID is averaged for a certain number of scans such that no overflow can occur. The spectrum is then Fourier transformed, phase corrected, stored in a separate area of memory or on backing store and the acquisition repeated after zeroing the data area. After a second block of data is Fourier transformed this is added to the first result and the process is repeated until a sufficient signal-to-noise ratio is obtained. The large solvent peak will have overflowed many times in the frequency spectrum but hopefully the small peaks of interest will remain.

Several difficulties and pitfalls are associated with the block averaging technique, the one of most practical importance is ensuring that the small signals of interest do not overflow and this is difficult to prevent without constant attention to the spectrum. This disadvantage disappears if the user has access to a disc-based FTNMR computer package because in this case each block of transformed data can be stored separately and only those in which there has been no homogeneity degradation or drift need be summed.

Cooper[15] has also pointed out that a Fourier transform program may scale two blocks of data quite differently even if they look similar to the eye and this could introduce errors if for example a block given a greater weighting were one in which the resolution had degraded.

In addition, the errors introduced by the Fourier transform process as noise may be greater than the

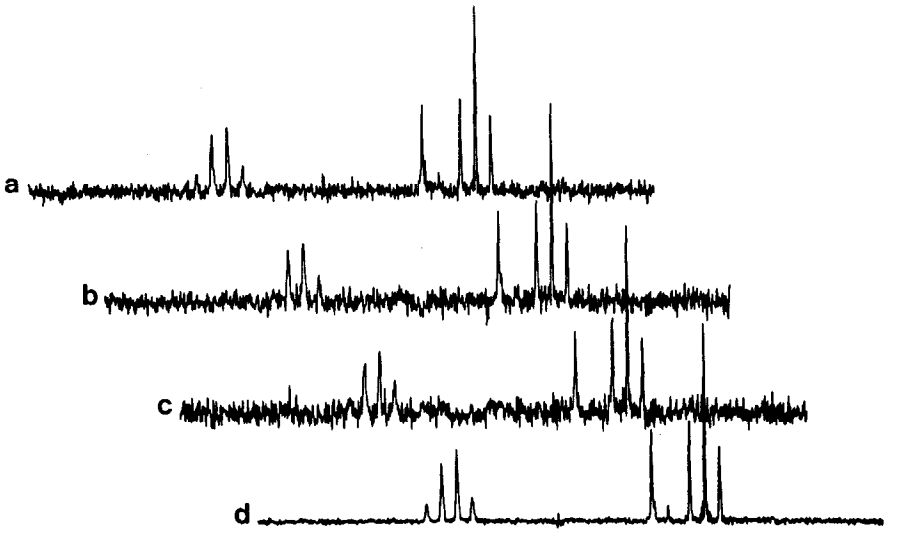


FIG. 9. A demonstration of the inadequacy of allowing input scaling to occur as a means of increasing the number of scans available when in the high dynamic range case. The sample consisted of a dilute solution of ethylbenzene in benzene with a small amount of C_6D_6 for field-frequency lock. Only the aliphatic region is shown. (a) The result of Fourier transformation of 64 scans using a 10 bit ADC into a 16 bit computer word. (b) The result from 64 scans at 10 bits plus 192 scans at 8 bits resolution. (c) The result of 64 scans at 10 bits plus 192 scans at 8 bits plus 768 scans at 6 bits resolution. (d) The result of 1024 scans at 10 bit resolution summed on backing store, and normalised to 16 bits prior to Fourier transformation.

small signals of interest. Each accumulated FID will be slightly different because of the random nature of the noise and it might be that this noise is the major component of the noise after the Fourier transform or that the transform noise is completely coherent so that block averaging gives no improvement. Cooper[24] has tested this point by simulation and shown that such transform noise is coherent and is not reduced by block averaging. If the transform noise is larger than the small signals of interest in the frequency domain then further block averaging will not improve the signal-to-noise ratio of interest and the only alternative is to perform a high dynamic range Fourier transform by which means the Fourier transform induced noise will be minimised (see Section 4.3).

2.6.4. *Double length averaging.* In a low dynamic range situation when the available computer word length is such that overflow is about to occur during an acquisition, it is possible to continue averaging by scaling down the memory contents, the digitised input and the ADC resolution, usually by a factor of two or four.

However, in the presence of a large signal from, for example, a solvent this is not possible because at a certain point the *noise* level will drop below one count in the ADC and if the small signals of interest give rise to modulations of the large decay which are the same order of magnitude as the noise, then at this point they will be distorted by the increasing quantisation errors.

It is possible to overcome this problem by storing the accumulated FID just before the point of overflow on backing store or in memory. Successive sets of accumulated FIDs can be added into the same area in a double length format, thus effectively doubling the word-length of the computer.

In order to realise the increased dynamic range it is not always necessary to perform a double length Fourier transform (see Section 4.3.2) because the least significant bits may only contain noise and in that case the total word can be normalised to single length.

Figure 9 shows an example of signal averaging in a high dynamic range situation. The sample consisted of a small amount of ethylbenzene in benzene solution with some C_6D_6 to provide a lock. Only the aliphatic region is shown. The top trace (a) represents the Fourier transform after 64 scans using a 10 bit ADC into a 16 bit computer word. At this point overflow would be imminent so the ADC resolution is divided by four as is the input signal and the computer memory. This now allows a total of 256 scans to be accumulated and this result is shown in (b). Finally the ADC resolution, input signal and memory contents are again divided by four to allow a total of 1024 scans to be measured. The result of this is shown in trace (c) and clearly now the signal-to-noise ratio is beginning to deteriorate. The bottom trace (d) is the result obtained by accumulating 1024 scans using a 10 bit ADC and transferring the data to backing disc every 64 scans when overflow became imminent. After 1024 scans the computer word contains about 20 bits due to the large signal. However the signal-to-noise ratio is such that at least the five least significant bits contain noise information (assuming one count per scan) and hence the data can be scaled to give a 16 bit word whereupon it is Fourier transformed in the normal way.

If the time domain dynamic range (which is usually less than that in the frequency domain) is such that normalisation of the data back to the length of the computer word causes loss of signal information, it is

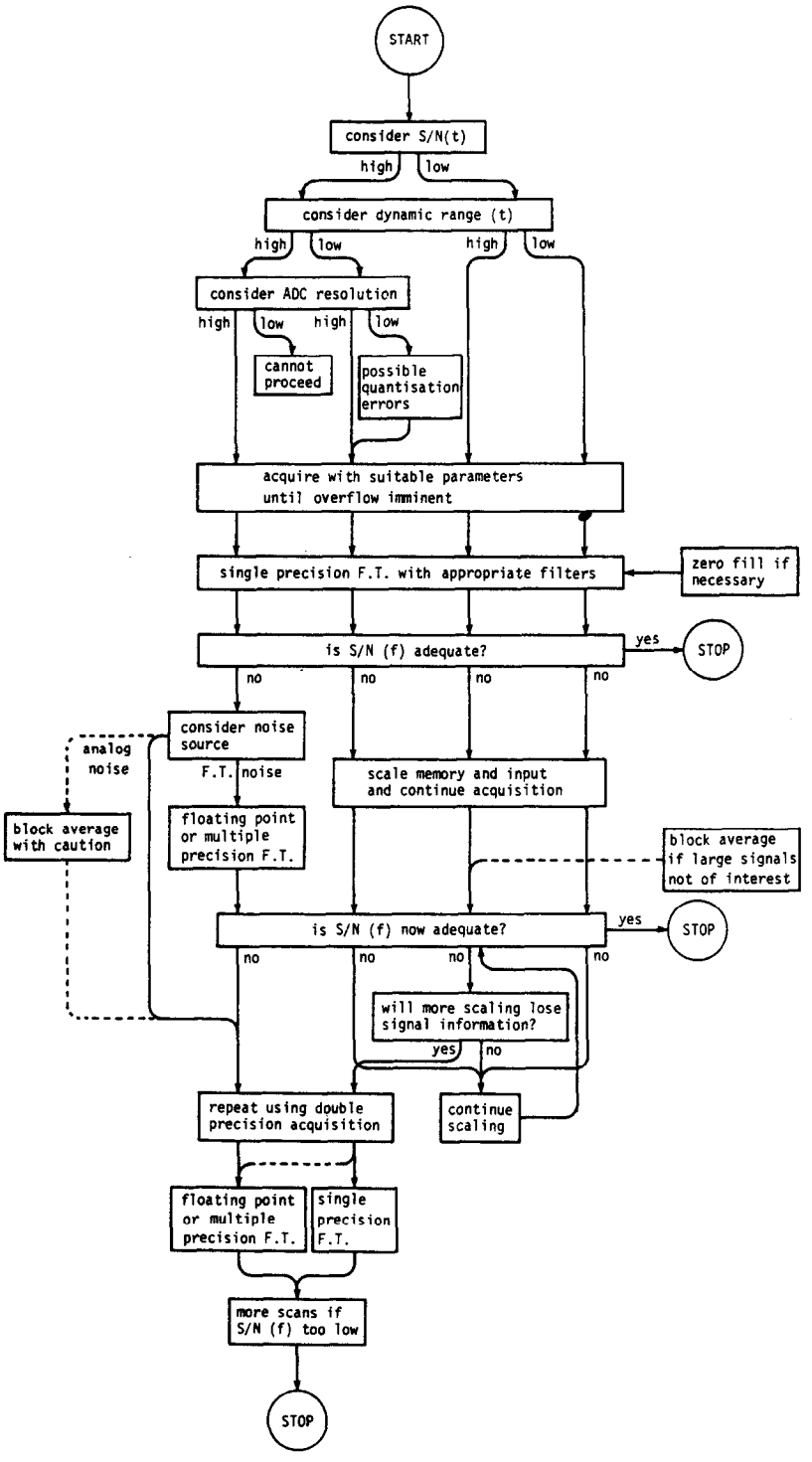


FIG. 10. Summary of the situations encountered for signal averaging in FTNMR and the techniques necessary for a successful result.

then necessary to perform a double length integer or floating-point format Fourier transform.

2.7. *Summary of Averaging Methods*

The choice of ADC resolution, computer word length and type of Fourier transform all need to be considered in the light of the dynamic range to be detected for samples with high or low sensitivity in the time domain.

Figure 10 shows a flow-chart which summarises the possible situations that can occur when attempting signal averaging in FTNMR. By following any particular line through this chart, this allows one to decide

those techniques which are probably suitable to give a successful result.

3. MANIPULATIONS PRIOR TO FOURIER TRANSFORMATION

3.1. *Zero Filling*

Until very recently, the question of what information can be gained by adding zeroes to the end of a free induction decay prior to Fourier transformation had been left in a rather imprecise and unsatisfactory state.

Bartholdi and Ernst[25] have shown from the principle of causality (i.e. that the FID is zero for $t < 0$ and real for $0 < t \leq T$) that for a free induction decay lengthened by an equal time period of zeroes independent information in the FID is obtained on Fourier transformation.

In practical terms, in order to resolve two peaks with separation $F_{\min}$ which will appear in the FID as a beat of period $F_{\min}^{-1}$, the acquisition time and hence the number of data points should be such that one observes at least half a cycle of this period i.e. $T = (2F_{\min})^{-1}$ or the time domain resolution is $F_{\min} = (2T)^{-1}$. To prevent aliasing it is necessary to sample at twice the highest frequency in the spectrum and this gives a resolution in the frequency domain of T^{-1}, where only half the data points are real. Adding to the FID an equal number of zeroes will double the point resolution in the frequency domain to be identical with that in the time domain, and as Bartholdi and Ernst[25] have shown these additional points are independent of the original set. For this reason the integrated signal-to-noise ratio in the frequency domain will be $\sqrt{2}$ times higher, but the peak height which is defined by the existing points will remain the same. However, if the new points are shifted by a frequency of $(2T)^{-1}$ so that they lie on top of the original set, this will result in an overall improvement in signal-to-noise by $\sqrt{2}$ although now the interpolated lineshape is degraded back to the original result. This is not the complete story, however, because it can be easily shown that adding more zeroes to an FID can give an additional improvement in peak lineshapes and positions. Bergland[26] has coined the term "picket-fence" effect to show why additional zero filling can improve the information content in a spectrum. In the frequency domain the spectrum consists of a series of sticks (hence the "picket-fence"), T^{-1} Hz apart, the intensities defined by the Fourier coefficients. However, because of the nature of the digitisation in the time domain which also consists of a series of delta functions the Fourier transformation should not be considered as a series of sticks but a series of sinc functions, T^{-1} Hz apart, with amplitudes given by

$$\frac{\sin\left[(\nu - n/T)\,T\right]}{(\nu - n/T)\,T}$$

where n is the harmonic number ($1/T$, $2/T$, $3/T$, etc.) of each frequency. Thus zero filling gives additional sinc functions which are interleaved between the original points in order to describe the intensities more accurately.

TABLE 5. Effect of Fourier interpolation on the definition of peak heights

No. of data points (multiples of original)	Point of overlap of sinc functions	Fractional value of Fourier coefficient midway between functions
1	$\pi/2$	0.637
2	$\pi/4$	0.900
4	$\pi/8$	0.975
8	$\pi/16$	0.994
16	$\pi/32$	0.998
32	$\pi/64$	1.000

Table 5 shows the fractional value of the Fourier coefficients exactly half way between adjacent sinc functions for various degrees of zero filling.

When attempting to resolve two peaks, the best possible situation would arise where a data point came exactly at the top of each resonance and here zero-filling by a factor of two would just give the potential time domain resolution but with no data points separating the peaks. For the general case Table 5 shows that to obtain greater than 99% of the intensity information at the most eight times the original number of data points will be required for the most favourable case. Also, in the most adverse case, a further doubling of the number of data points will give about the same proportion of the full intensity. i.e. an upper limit of 16 times the number of original data points is the maximum required.

As a practical demonstration we have measured the doublet arising from the protons of two non-equivalent methyl groups in dimethylformamide, with a separation at 90 MHz measured as 11.4 Hz. The FID was acquired into 16k data points and exponentially weighted such that the doublet was just resolved. (Knowing the separation, the line broadening to just merge the peaks is merely a factor of $\sqrt{3}$; this assumes Lorentzian lineshapes and can be obtained by double differentiation of the lineshape function for two overlapping lines). The FID was then truncated to 256 data points such that half of one cycle of the beat frequency was still observed. In order to demonstrate a zero-filling of $\times 64$ we had to choose a system which gave half a cycle of the beat frequency of interest and which decayed within 256 data points. The former was accomplished by adjusting the spectral width and the latter was realised using exponential weighting. Figure 11 shows the result of Fourier transformation of this 256 point FID with various amounts of zero filling. The appearance of the doublet does not improve significantly beyond a zero filling of $\times 32$.

In some cases the available computer data memory is insufficient to allow zero filling and although this could be overcome by using the disc-based Fourier

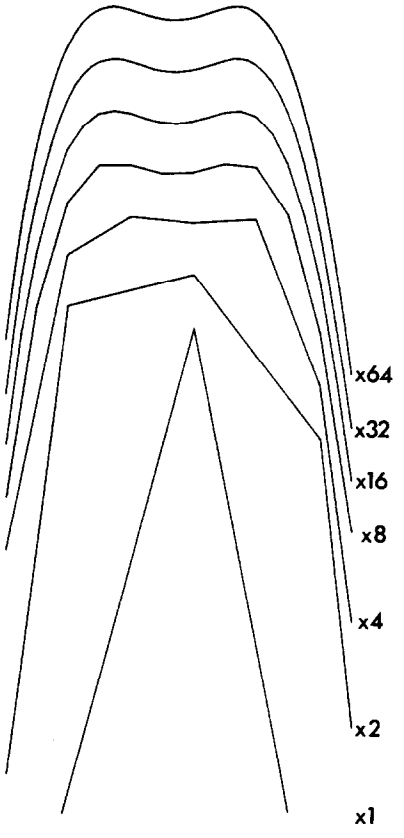


FIG. 11. The effect of zero filling a 256 point FID by the factors shown in order to improve the definition of a doublet. The spectrum displayed is the broadened methyl doublet of dimethylformamide such that one half cycle of the beat frequency arising from the doublet separation came within 256 data points and also decayed to within a few bits of zero.

transform which has recently become available, this new development has not yet been implemented in many user data systems.

An alternative method of obtaining the same information as from zero filling was first suggested by Pajer and Armitage,[27] and this involves complex interpolation of segments of the total spectrum. Bartholdi and Ernst[25] were the first to propose this general method although their complex interpolation function would be difficult to program. Additional data storage is only needed for the interpolated points chosen and also if only temporary data storage is required the complete spectrum may be retained on backing store.

The revised method for a single phase detection system which uses the conventional Fourier transform software, in which one conventionally performs Fourier transforms on only real data, is as follows:

(i) obtain a normal complex frequency domain spectrum consisting of $N/2$ real and $N/2$ imaginary points from an N point FID.
(ii) select a complex segment of the original; this must be a sub-multiple of two of the original if the fast Fourier transform routine is to be used (e.g. N' complex points) and set the imaginary part to zero.
(iii) take the inverse Fourier transform to obtain an N' point real pseudo-FID. This can be thought of as a normal FID sampled at twice the segment spectral width for the original acquisition time.
(iv) resolution or sensitivity enhancement functions may be applied to this pseudo-FID.
(v) it is now possible to zero-fill this decay by N' or more points and to Fourier transform in the usual way to produce an interpolated segment which can be phase corrected and processed as a normal spectrum.

For a spectrometer system using quadrature phase detection which yields a complex FID it is more suitable to take the complex frequency spectrum segment and to take the complex inverse transform to produce a complex pseudo-FID. After this, it is possible to apply sensitivity or resolution enhancement functions, zero-fill with complex points and take the normal complex Fourier transform to give a complex interpolated segment.

The resolving power of the method is the same as normal zero-filling although mathematically distinct. Oscillations are often seen on the wings of the peaks as a result of the pseudo-FID not being zero at the end. An extra constraint is that the chosen segment must not contain any peaks which extend past the edges. It is necessary to realise that apart from the usually observed absorption spectrum there is a dispersive component and in fact peaks in this representation extend much further than the absorption components. The dispersive tails extend throughout the whole spectrum but do decay such that it is possible to ignore any inaccuracies caused by such truncation.

For cases of narrow peaks superimposed on wide lines in narrow spectral widths, the truncation errors introduced make straight-forward segmental interpolation impossible. This problem may be overcome by applying a resolution enhancement to the original FID thus narrowing the lines in the frequency spectrum before selection of the segment to be interpolated. Any distortions introduced into the line shape can be removed at the stage of the pseudo-FID by applying an inverse function. For example if the original FID is multiplied by $[1 - \exp(-t/k)]$ to remove broad wings, this being the convolution difference method, then at the stage of the pseudo-FID this can be multiplied by $[1 - \exp(-t/k)]^{-1}$.

Since it is not always possible to choose a spectral segment (this has to be a fraction $1/2^n$ of the original if the discrete Fourier transform is to be used) which does not have a resonance near the edges, a suggestion has been made to apodise the segment at the ends using a cosine function[28] prior to inverse Fourier transformation.

To summarise therefore, the addition of more than N zeroes to an N point FID can yield more spectral

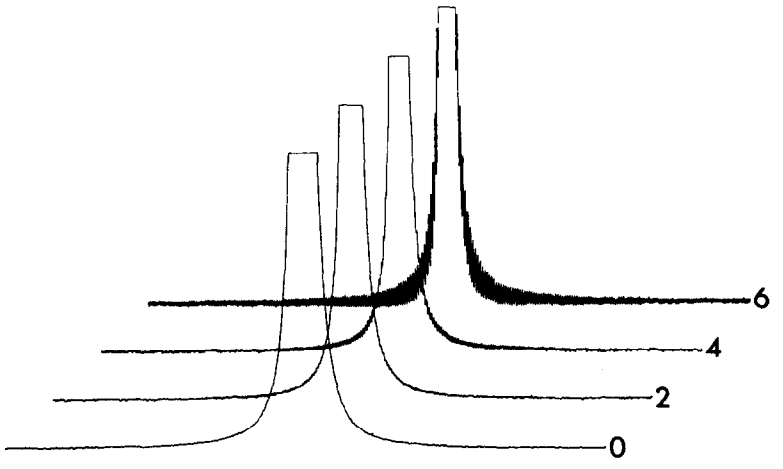


FIG. 12. The effect of allowing the FID from a single line to have a finite value at the ends of the acquisition. A 12 bit ADC was used and the diagram shows the Fourier transformed result for a level of 0, 2, 4 and 6 bits at the end.

information. True, the extra points over the first N added zeroes do not after Fourier transformation give independent information but the original points have the correct values and the extra points provide a (sin x)/x type interpolation. In some cases, these points serve to define a line better or to give a peak position more precisely, but modern software packages often now incorporate methods in the plot routines to provide some sort of interpolation on output (see Section 5.5).

3.2. *Sensitivity Enhancement*

3.2.1. *Introduction.* This section covers those techniques which have been applied to NMR data in order to improve the sensitivity of the final result. In general some compromise is reached on the balance between sensitivity and linewidth. A major review by Ernst[29] in 1966 elegantly and comprehensively set out the theoretical background both for CW-swept spectra and for the pulse-Fourier transform case. As a consequence of this monograph, we will only consider those techniques which have been applied on Fourier transform NMR data systems.

3.2.2. *Apodisation methods.* The earliest attempts at apodisation, which reflect the true use of the term, were simply to ensure that at the end of the acquisition period the FID had decayed essentially to zero, thus removing errors in the transformed result due to truncation. The effect of truncation is shown in Fig. 12 which illustrates the distortions introduced if the FID has non-zero signal intensity at the end of the acquisition. Also it should be noted that the linewidths are narrower in the truncated examples because of what are essentially longer T_2^* relaxation times.

The simplest way to ensure a zero in the last channel is to multiply the FID, channel by channel, by a linear function with an initial value of one and a final value of zero (see Fig. 13b). Since most of the signal information is in the early part of the FID, this method improves the final signal-to-noise ratio in the spectrum, but the lineshapes are no longer Lorentzian and can be badly distorted by "wiggles" on either side of the resonances, according to the degree of apodisation.

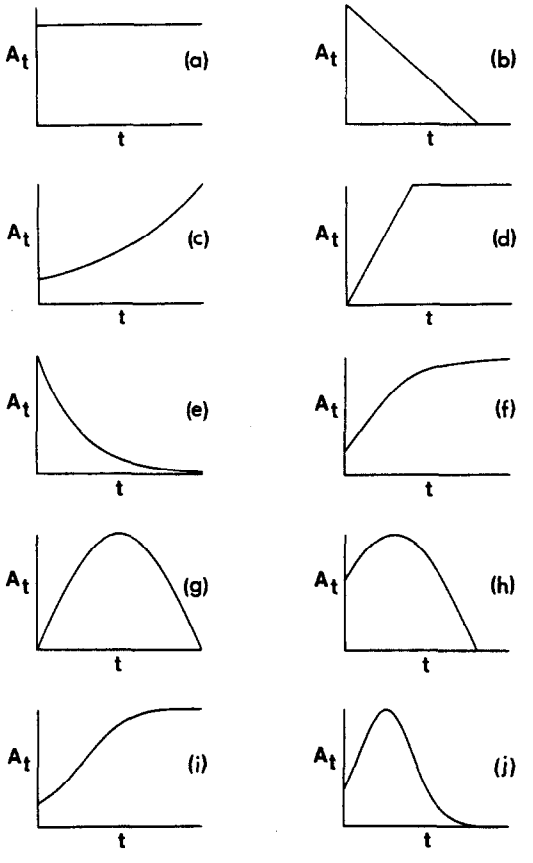


FIG. 13. The form of the various apodisation functions used to improve either the sensitivity or resolution of an NMR spectrum. (a) An unweighted FID. (b) Linear apodisation. (c) Increasing exponential. (d) Trapezoidal resolution enhancement. (e) Decreasing exponential. (f) Convolution difference. (g) Sine-bell. (h) Shifted sine-bell. (i) LIRE. (j) Gaussian transformation.

If however, the universally used negative exponential weighting function (Fig. 13e) is applied to the FID, the lineshape remains Lorentzian after transformation and all the linewidths increase by the same amount. This method improves the sensitivity and does not introduce any lineshape distortions.

If the signal-to-noise ratio in the final spectrum is the only criterion then there will be an optimum value of the broadening, depending on the linewidth of interest. The optimum or matched filter is simply an exponential weighting with the same time constant as the experimental decay. Multiplication of the decay by an exponential with the same time constant results, after transformation, in a line with twice the width of that resulting from an unfiltered decay.

It is possible to calculate the effect that various values of the exponential filter have on the signal-to-noise ratio in the frequency domain and on the noise content of the FID.[7] First it is necessary to define the signal-to-noise ratio in the frequency domain in a quantitative fashion. Let us assume for simplicity a single line on resonance detected using single phase detection, although this is not a realistic situation. The time domain signal $f(t)$ is:

$$f(t) = M_0 \exp(-t/T_2^*). \tag{16}$$

Since $f(t) = 0$ for $t < 0$ or $t > T$, the acquisition time, then the signal intensity in the frequency domain after Fourier transformation from Parseval's theorem is:

$$S(f) = M_0 \int_0^T \exp(-t/T_2^*)\, dt \tag{17}$$

$$= M_0 T_2^* (1 - \exp(-T/T_2^*)). \tag{18}$$

The r.m.s. value of the noise $N(f)$ is given by

$$N(f) = \left\{ \int_0^T [n(t)]^2 \, dt \right\}^{1/2}, \tag{19}$$

where $n(t)$ is the instantaneous noise fluctuation. This can be replaced by an r.m.s. average if $n(t)$ has a high frequency compared to the acquisition time T; then

$$N(f) = \bar{n} T^{1/2}. \tag{20}$$

Thus

$$S/N(f) = \frac{M_0}{\bar{n}} (T_2^*)^{1/2} \left(\frac{T_2^*}{T}\right)^{1/2} [1 - \exp(-T/T_2^*)]. \tag{21}$$

A number of well known experimental conclusions are clear from eqn (21). Since the signal-to-noise ratio depends on T_2^*, a long T_2^* gives a sharp line after Fourier transformation and hence a higher signal-to-noise ratio in the frequency domain than a line with a shorter T_2^*. If the acquisition time T is short compared with T_2^* then $S/N(f)$ will be increased, but the truncation of the FID as a result of T_2^* being of the same order as T will make this an unacceptable situation. However, a long acquisition time gives a higher proportion of noise and hence a lower $S/N(f)$.

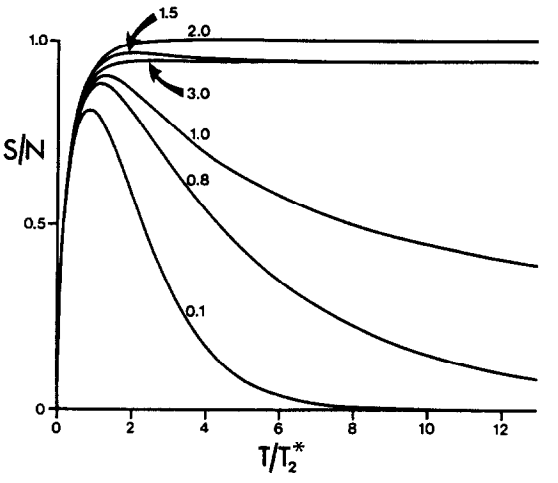


FIG. 14. The signal-to-noise ratio obtainable in the frequency domain as a function of the acquisition time T. (To make this parameter unit free the ratio T/T_2^* has been chosen). The curve marked 1.0 is for an unfiltered FID. The figures 1.5, 2.0 and 3.0 correspond to an exponential weighting giving that factor of line-broadening; 0.8 and 0.1 represent the result of applying an increasing exponential to give line narrowing to those fractions of the original.

Now, turning to the question of multiplying a FID with an exponential function having a time constant of T_2^*/k, then eqn (18), eqn (20) and eqn (21) become:

$$S(f) = \frac{M_0 T_2^*}{k+1} [1 - \exp((-k+1)T/T_2^*)] \tag{22}$$

$$N(f) = \bar{n} \left[\frac{T_2^*}{2k} (1 - \exp(-2kT/T_2^*)) \right]^{1/2} \tag{23}$$

$$S/N(f) = \frac{M_0 (2kT_2^*)^{1/2}}{\bar{n}(k+1)} \times \frac{(1 - \exp(-(k+1)T/T_2^*))}{(1 - \exp(-2kT/T_2^*))^{1/2}}. \tag{24}$$

As k tends to zero, eqn (24) approaches eqn (21).

The effects of truncation and exponential weighting on the signal-to-noise ratio in the transformed spectrum are shown in Fig. 14 with curves calculated from eqn (21) and eqn (24) for various degrees of exponential weighting.

The unfiltered response ($k = 0$) gives a maximum signal-to-noise ratio if $T \simeq 1.2T_2^*$. A matched filter corresponds to $k = 1$, and here the signal-to-noise ratio is essentially constant for T greater than about $2T_2^*$.

In a real spectrum it is not possible to choose the optimum filter for lines of different widths, and in fact it may not be desirable to apply the optimum filter because of the resulting increase in linewidth, and so some compromise is usually reached.

Exponential weighting functions are described on commercial systems either by a line broadening parameter in Hz, LB, or a sensitivity enhancement parameter in seconds, SE. The former is preferred since it

relates to the effect on the final spectrum irrespective of acquisition parameters; LB is simply $(\pi SE)^{-1}$.

Having obtained in eqn (24) an expression for the signal-to-noise ratio after Fourier transformation as a function of T_2^* and the exponential filter value, it is possible to divide this by the signal-to-noise ratio in the FID (the height at $t = 0$, i.e. M_0 divided by $\bar{n}$, the r.m.s. noise level) to obtain the ratio of the signal-to-noise values in the two complementary domains.

This will be essentially the same as eqn (24) but including a term $F^{1/2}$ where F is the spectral width of interest. This is because the fluctuating noise voltage at any point in the time domain is sorted out according to its frequency components by the Fourier transform process and thus any one point will only contain $1/F$ of the amplitude in the time domain or for r.m.s. voltages the noise level will be F times lower in the frequency domain. Thus

$$\frac{S/N(f)}{S/N(\mathrm{FID})} = \frac{(2kT_2^*F)^{1/2}}{(k+1)} \times \frac{(1-\exp(-(k+1)T/T_2^*))}{(1-\exp(-2kT/T_2^*))^{1/2}} \quad (25)$$

This expression can be simplified if we ensure that $T \gg T_2^*$ and then

$$\frac{S/N(f)}{S/N(\mathrm{FID})} = \frac{(2kT_2^*F)^{1/2}}{(k+1)} \quad (26)$$

This expression applies if single phase detection of a single line on resonance is used. It would be $\sqrt{2}$ times larger for quadrature detection and a factor of two smaller for single phase detection with the signal off-resonance when the offset frequency will cause the FID to have both positive and negative excursions and S/N(FID) will be apparently twice as high. In that case

$$\frac{S/N(f)}{S/N(\mathrm{FID})} = \left(\frac{kT_2^*F}{2}\right)^{1/2}\left(\frac{1}{k+1}\right). \quad (27)$$

k can be redefined in terms of a sensitivity enhancement parameter $E(E = T_2^*/k)$, or a line broadening parameter L, $(L = k/\pi T_2^*)$.

$$\frac{S/N(f)}{S/N(\mathrm{FID})} = \left(\frac{T_2^*E}{T_2^*+E}\right)\left(\frac{F}{2E}\right)^{1/2} \quad (28)$$

$$= \left(\frac{T_2^*}{\pi T_2^*L+1}\right)\left(\frac{\pi LF}{2}\right)^{1/2} \quad (29)$$

For the optimum filter

$$\frac{S/N(f)}{S/N(\mathrm{FID})} = \left(\frac{FT_2^*}{8}\right)^{1/2} \quad (30)$$

for the case of single phase detection off-resonance.

Equation (30) shows that the signal-to-noise ratio differences in the two domains can be substantially different, e.g. for $F = 1$ *kHz* and $T_2^* = 1$ second then this ratio is about 11 or for quadrature detection on resonance this ratio would be ≈ 32.

3.3. *Resolution Enhancement*

3.3.1. *Introduction and spectroscopic methods.* The techniques of resolution enhancement have been much less well documented than those of sensitivity enhancement and with the advent of digital signal processing regimes it is possible that some real gains in information content may be obtained in an analogous way to zero-filling.

Resolution enhancement falls into two main classes. Firstly special experiments can be designed which cause line narrowing. These, such as the spin–echo pulse sequence[5,30] are outside the scope of this review and have been summarised by Campbell.[31] The second method involves manipulation of the acquired data in some way and the various methods so far suggested are described in the next paragraphs and their relative efficiencies compared. These latter methods involve multiplying the FID by a function which increases the intensity at the latter end of the FID at the expense of the initial part. Thus they will cause a selective enhancement of some linewidths at the expense of others and this property can be utilised for baseline flattening. There will be a decrease in the observed linewidth with a corresponding decrease in the signal-to-noise ratio. All functions aim to improve the resolution with a minimum loss in signal-to-noise ratio and hopefully minor signal distortions.

It has sometimes been stated[35] that the information content of an FID and its Fourier transform are identical and that if there is no evidence of resolution in the frequency domain, then resolution enhancement cannot work. However, we have seen that zero filling an FID can yield new information and if one assumes that the NMR lines are Lorentzian then meaningful deconvolution can also be performed.

There is one special case in which the free induction decay is modulated by a low frequency beat, the characteristics of which are fixed by the spin system. Moniz and co-workers[32] have shown that selective Fourier transformation of the beat modulated FID allows one to separate the broad and sharp components. For example the FID arising from two lines close together will contain two frequencies, one indicative of the offset from the r.f. carrier and the other a function of the separation between the two lines. Clearly transformation of that part of the FID consisting of the initial decay before the first null point can give no information on the doublet separation but Fourier transformation of the beats will allow resolution to be obtained even though an absorption spectrum will not result. This can be obtained by adding in to the transformed beat spectrum a portion of the transformed initial decay. The method has been applied to ^{1}H-coupled ^{13}C NMR spectra where long range ^{13}C–^{1}H couplings have been resolved.

Finally, distinction should be made between resolution and linewidth because it is possible to improve the resolution in a spectrum by altering the lineshape but leaving the linewidth at half height unchanged.

3.3.2. *Convolution difference.* This method has been applied to both CW and FT NMR spectra, since it is usually applied to the frequency domain spectrum.[33,34] Here, a spectrum is obtained and copied into a separate area of memory. One copy is broadened by the application of a smoothing function and a fraction of this is subtracted from the original. The broader version of the spectrum obtained in an FT experiment, is simply a result of using a larger exponential weighting. Thus in the time domain the same result can be obtained by multiplying the FID by a function A_t given by

$$A_t = 1 - a \exp(-bt/T) \tag{31}$$

where a is the fraction of the broad component, b represents the time constant corresponding to the extra line broadening introduced, and T is the acquisition time (Fig. 13f).

3.3.3. *Increasing exponential.* This involves multiplying the FID by a positive exponential function (Fig. 13c)

$$A_t = \exp(bt/T) \tag{32}$$

where b is chosen empirically. In the limit b could exactly balance the natural decay time, T_2^*, of the FID and produce a non-decaying truncated interferrogram corresponding to a δ-function. However, then lineshape distortions are introduced and the noise increase is overwhelming. Usually some compromise is sought but often only very modest enhancements are possible before the signal-to-noise ratio decreases markedly.

A quantitative treatment of the linewidth reduction that is possible for a given signal-to-noise ratio in the frequency domain can be made by reformulating eqn (24) to take account of a positive exponential weighting. The effect on the final signal-to-noise ratio of applying an increasing exponential is shown in Fig. 14 as a function of T/T_2^*. Clearly, if the acquisition time is prolonged such that at the end of the FID only noise is being gathered, the final signal-to-noise ratio will be so low as to give a meaningless result. The diagram shows the degradation in the signal-to-noise ratio as a result of applying an increasing exponential for a 20% and a 90% reduction in linewidth.

3.3.4. *Trapezoidal function.*[35] This is an improvement on the simple exponential in that it reduces the beginning of the FID but leaves the latter part relatively unaffected (Fig. 13d). It is also essentially the same as the convolution difference method except that more lineshape distortions are introduced.

In this case the FID is multiplied by the function

$$\begin{aligned} A_t &= bt/T \\ A_t &= 1 \qquad (t \geq T/b) \end{aligned} \tag{33}$$

The parameter $b(\leq T)$ can be varied empirically to obtain the desired result.

3.3.5. *Sinebell function.*[36] This method involves multiplying the FID by a sinewave of zero phase and period of twice the acquisition time

$$A_t = \sin(\pi t/T) \tag{34}$$

In this method there are no empirical parameters to adjust (Fig. 13g).

The sinebell function affects both the beginning and end of the FID equally but usually the decay has vanished into the noise well before the end of the acquisition time and hence the convolution has the result of affecting the initial part of the FID most. Of course, an arbitrary parameter can be introduced into the sinebell routine just as in convolution difference where it is possible to vary the proportion of broad signals to be subtracted and this allows some control over the final sensitivity.

The shifted sinebell method is a modification of the original function in which a phase parameter has been added[37] (Fig. 13h). This then allows the maximum in the sinusoidal envelope to be adjusted to the component of interest in the FID

$$A_t = \sin(\pi t/T + \phi) \tag{35}$$

Recently Clin *et al.*[38] have shown that the sinebell routine can be achieved simply without the need to program the algorithm explicitly. The lineshape obtained after sinebell manipulation can be produced by adding two dispersion Lorentzian functions shifted by one computer address and 180° out of phase. In practical terms, the Fourier transform of an FID is calculated and, after phase correction to the dispersive component, is copied into a separate area of memory. This second block is left shifted by a single address and from this is subtracted the original version. A proportion of the original signal can be added back in, such that the final signal-to-noise ratio is not too seriously degraded, but of course this leads to a proportional degradation in the linewidth back towards the original.

Gueron[37] has summarised the similarities between the convolution difference and sinebell approaches and the methods have been generalised. He also showed that the sinebell is simply equivalent to the approach of Clin *et al.*[38] in that it represents the difference between two dispersive signals separated by one computer address. He has also shown that by adding a phase parameter to the sinebell routine, that this is equivalent to adding in a proportion of the original absorption signal to the dispersion difference spectrum. This lessens the troughs on either side of the resonance and improves the sensitivity with the concomitant compromise on resolution.

The sinebell function can be written as $\sin(2\pi Ft)$ where $F = (2T)^{-1}$ but larger values of F cause single narrow peaks to split into two.

3.3.6. *LIRE Function.* Akitt[39] has discussed the various forms of distortion introduced into a spectrum by the use of different deconvolution functions. He has proposed a new function which he calls LIRE and which reduces the rate of decay of an FID (Fig. 13i)

$$A_t = \frac{a}{(a-1)\exp(-t/T) + 1}. \tag{36}$$

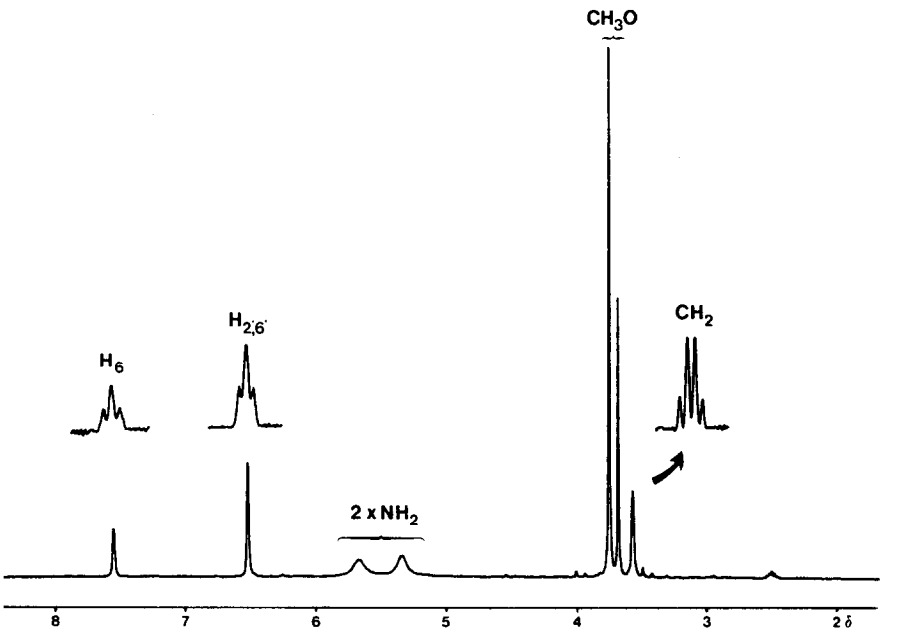


FIG. 15. The effect of Gaussian deconvolution on the 90 MHz ^{1}H NMR spectrum of trimethoprim in DMSO-d_6. Application of the function in eqn (37) enables the resolution of the long range coupling constants. 1250 Hz width, 8k data points zero filled to 16k before Fourier transformation, 600 Hz displayed.

This sigmoid function for $a > 10$ at first follows closely exp (t/T) but eventually falls back to a constant value. At this stage the weighting function gives a rather unfavourable signal-to-noise ratio but this can be overcome by further apodisation of the FID using a function which converts the lineshape to that arising from a power function of n ($n = 1$ corresponds to a Lorentzian, $n = 2$ to a Gaussian). Consequently with this apodisation the method resembles Gaussian deconvolution (Section 3.3.7).

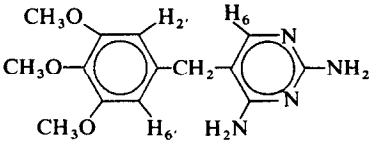
Another method analogous to the LIRE function has been suggested by Semendyaev[40] which involves a two stage manipulation of an FID. First the linewidths are narrowed by multiplying the FID by a positive exponential. Then the signal-to-noise ratio of the spectrum is increased towards that of the original by multiplying by a cosine weighting function.

3.3.7. *Gaussian Transformation.*[41] In this case the FID is multiplied by a combination of a positive exponential to cancel the natural negative decay of the component of interest followed by a negative squared exponential with an appropriate time constant to produce a Gaussian line of the desired reduced width.

$$A_t = \exp\,(at/T - b(t/T)^2) \qquad (37)$$

Here a and b are parameters which can be calculated for a single line and which are chosen empirically in a complex spectrum (Fig. 13j). This method gives the best resolution enhancement for a given signal-to-noise ratio and consequently has now been adopted by the majority of instrument manufacturers.

An example of this technique is given in Fig. 15 which shows the 90 MHz ^{1}H NMR spectrum of the widely used antibacterial compound trimethoprim.



Normally couplings between the methylene and the various aromatic protons are not observed, but optimum application of the Gaussian deconvolution reveals the expected coupling patterns.

3.3.8. *Enhancement of absolute value spectra.*[42] The absolute value or magnitude mode (i.e. $(v^2 + u^2)^{1/2}$) instead of the conventional absorption or v-mode has the disadvantage that each transition has extremely wide wings as a consequence of the fact that the dispersion or u-mode extends further from resonance.

$$v(\omega) = \frac{1}{1 + (\omega - \omega_0)^2 T_2^{*2}}\,;\; u(\omega) = \frac{(\omega - \omega_0)}{1 + (\omega - \omega_0)^2 T_2^{*2}} \qquad (38)$$

That is, from the form of the Lorentzian lineshape the v-mode signal tails off proportional to $(\omega - \omega_0)^{-2}$ but the u-mode only decreases as $(\omega - \omega_0)^{-1}$ for cases with large offset from the peak maximum, $(\omega - \omega_0)^2 T_2^{*2} \gg 1$. Also, the absolute value mode is not additive for the overlap of adjacent transitions since the left and right hand wings of the dispersive component of each line have opposite signs and tend to cancel each other.

Since the Gaussian decay is an even function the u- and v-modes have the same frequency dependence, and this makes the Gaussian deconvolution technique a convenient choice for processing absolute value spectra.

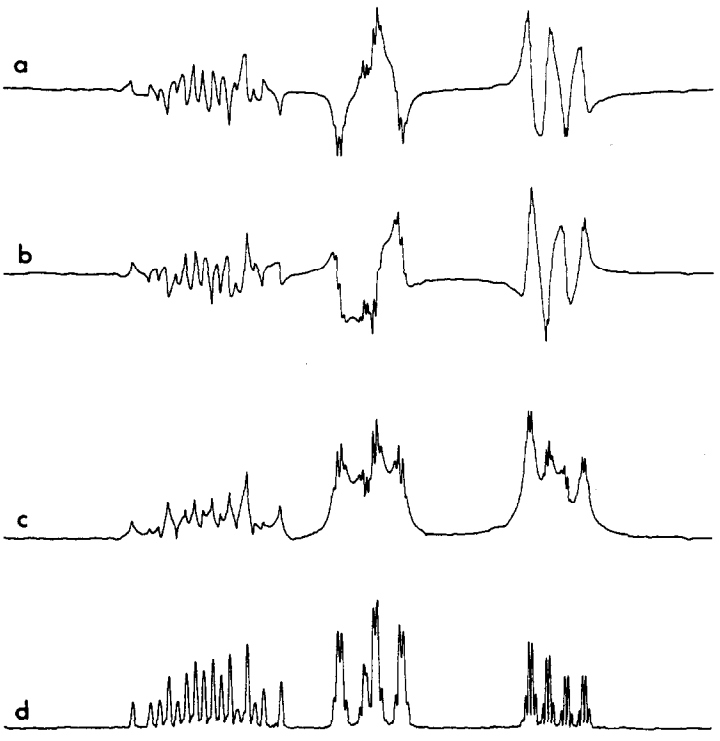


FIG. 16. The 90 MHz ^{1}H NMR spectrum of propylene oxide, excluding the methyl signal, obtained by Fourier transformation of the half-echo following a Carr–Purcell type A sequence with $\tau = 0.06$ sec. Trace (a) is the absorption mode and trace (b) the dispersion mode signal. Trace (c) is the absolute value presentation. Trace (d) shows the result of applying the double exponential weighting given in the text and plotted as absolute value.

The standard method of measuring spin–spin relaxation times is through the use of the Carr–Purcell spin–echo sequence. However, coupling between resonant nuclei gives rise to modulations of the spin echoes[43] which produce phase modulation of the resonances in the transformed spectra. For simple spin systems with first order couplings, even though they may be in complex molecules, the phase distortions can be analysed and even utilised to simplify spectra but in general very complex J-modulation results. In addition, J couplings which are not resolved also cause amplitude modulation of the signals. One way of overcoming these distortions is to present the spectra in absolute value mode but for anything but very simple spectra the distortions caused by overlap preclude intensity measurements. All these difficulties are overcome by deconvolution using the Gaussian method[41] and presenting the spectra in absolute value mode. As an example, Fig. 16 shows the results of Fourier transforming the second half of an echo formed by a "90°–τ–180°–τ– acquire" sequence for propylene oxide in DMSO-d_6 at 90 MHz. The methyl signals are not shown and the data was obtained with $\tau = 0.06$ sec. Trace (a) is the v-mode display and trace (b) the u-mode display, both showing severe echo modulation, whilst trace (c) shows the absolute value display indicating the non-additivity of overlapping transitions. After multiplying the FID as set out in eqn (37) with $a = 55$ and $b = 201.7$ and Fourier transforming, trace (d) is obtained. Here intensity measurements are quite feasible either by peak height or by integration and for a given resonance the intensity will still decay with a time constant of T_2^* for different τ values.

The newly expanding area of two-dimensional Fourier transform NMR[6,44] with the difficulties associated with phase correction has seen the other major use of absolute value presentation with the concomitant wide wings and non-additive overlap. Again Gaussian deconvolution with an absolute value presentation can give an enhanced appearance to the spectrum.

The Fourier transformation of the beats in a beat-modulated FID[32] (Section 3.3.1) leads to a non-absorption spectrum and, as an alternative to adding in part of the original absorption spectrum to produce a less distorted lineshape, presentation of the result as an absolute value spectrum would also remove most of the lineshape distortion.

3.3.9. *Comparison of the methods.* Figure 17 shows a comparison between most of the methods which involve computer manipulation. We have simulated a typical noise-free FID which on transformation gives the top left spectrum in the diagram. This can be considered as arising from a spectral width of 1000 Hz, with 8192 data points (an acquisition time of 4.096 seconds). The initial linewidth was 1.165 Hz corresponding to $T_2^* = 0.273$ seconds. The separation of the lines in the 1:4:6:4:1 quintet was 0.98 Hz. The

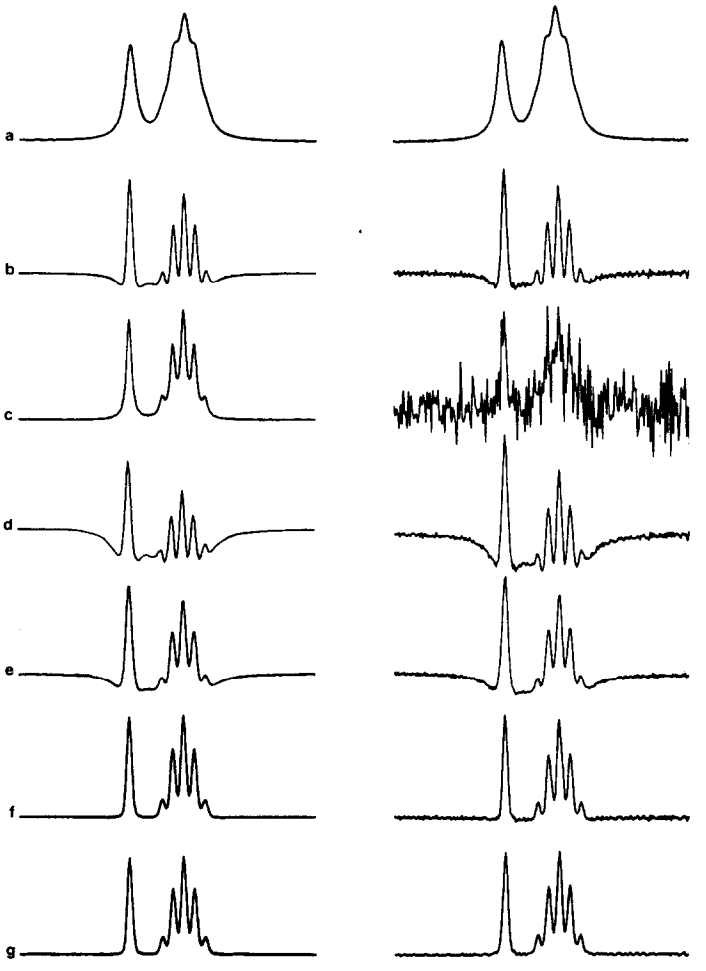


FIG. 17. A demonstration of some of the various computer methods for resolution enhancement. The left hand traces result from Fourier transforming a synthesised FID giving a single resonance and a 1:4:6:4:1 quintet. The right hand traces result from the same FID with noise added. (a) Initial spectrum, (b) trapezoidal function, (c) positive exponential, (d) convolution difference, (e) sine-bell, (f) LIRE, (g) Gaussian deconvolution. The enhancements were all applied to give the same linewidth reduction (to 45%) with optimum signal-to-noise ratio. The sine-bell provided an exception since because there are no adjustable parameters, a wider linewidth and consequently a better signal-to-noise ratio was achieved.

right-hand traces arise from the same FID with noise, sampled at the correct rate, added. A reduction in the linewidth to 0.45 of its initial value was selected as being a reasonably typical maximum and the parameters were adjusted to give the optimum signal-to-noise ratio.

Trace (a) is the initial spectrum, (b) is the result of the trapezoidal function, (c) increasing exponential, (d) convolution difference, (e) sine-bell, (f) LIRE, and (g) Gaussian transformation. From the left-hand traces it is obvious that the trapezoidal function, convolution difference and sine-bell all give serious base-line distortions, particularly for convolution difference when attempting large enhancements. In this case there is a compromise between signal-to-noise ratio and base-line distortion. The advantage of the rising exponential is that the Lorentzian lineshape is retained but the resolving power may not be so good because of wing overlap. The Gaussian method converts the Lorentzian lines into Gaussians which do not have the wide wings. In practice it is possible to over-enhance both for the Gaussian and LIRE methods such that any wings disappear, giving better resolving power.

The results on the right-hand side for an FID containing noise are really a comparison of the achievable signal-to-noise ratio for a given reduction in the linewidth. The increasing exponential gives such a large expansion of the noise at the end of the FID that, on scaling, the signal part of the FID is compressed into fewer bits leading to quantisation errors which are responsible for the periodic nature of the noise. The trapezoidal function, convolution difference and LIRE all give similar signal-to-noise ratios because they all involve multiplication of the FID by a function which increases and then becomes approximately constant. The LIRE method is the optimum function of these three because of the level baseline produced. A comparison of the sinebell and the

Gaussian methods shows them to have similar signal-to-noise ratios even though it was necessary to accept a smaller resolution enhancement for the sinebell because there are no adjustable parameters in this method. The sinebell technique gives a higher bandwidth of noise presumably because of the wider shape of the function. The Gaussian method although it involves empirical parameters is preferred because of the better baseline and signal-to-noise ratio for the same reduction in linewidth. The Gaussian and LIRE functions give almost identical results except that the signal-to-noise ratio is better for the Gaussian method. This is because Gaussian deconvolution cuts out noise at a point at which the LIRE function is constant except for apodisation in the latter case of the last few addresses.

The principal conclusion of these comparisons is that the Gaussian method gives the best signal-to-noise ratio for a given reduction in linewidth and also produces the minimum baseline distortion.

4. FOURIER TRANSFORMATION

4.1. *Continuous vs Discrete Transforms*

The normal integral equation which defines the Fourier transform pair is

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) \exp(-i\omega t)\, dt \tag{39}$$

where $F(\omega)$ and $f(t)$ are the frequency and time domain functions respectively. This equation can be replaced by a pair for the cosine and sine parts of the transform.

$$C(\omega) = 2\int_0^{\infty} f(t) \cos \omega t\, dt \tag{40}$$

$$S(\omega) = 2\int_0^{\infty} f(t) \sin \omega t\, dt. \tag{41}$$

In a digital computer eqn (39) must be replaced by a discrete Fourier transform and is recast as

$$A_p = N^{-1} \sum_{k=0}^{N-1} X_k \exp(-2\pi ipk/N) \qquad p = 0, 1, 2 \ldots N-1 \tag{42}$$

where A_p is the pth coefficient in the frequency domain and X_k is the kth value of the time domain signal.

4.2. *The Cooley–Tukey Algorithm*

Evaluation of the N values of A_p requires approximately N^2 multiplications and the Cooley–Tukey algorithm[45] was designed to speed up the process by cutting down this number to about $2N \log_2 N$ arithmetic operations. This was achieved in the following manner.

Let us define a time series of N points, X_k, where N is some power of 2; this series can be divided into two other series Y_k and Z_k containing only the even and odd numbered terms respectively i.e.,

$$Y_k = X_{2k} \quad Z_k = X_{2k+1} \quad k = 0, 1, \ldots N/2 - 1. \tag{43}$$

Each series will have its own discrete Fourier transform

$$B_p = \sum_{k=0}^{N/2-1} Y_k \exp(-4\pi ipk/N) \tag{44}$$

$$C_p = \sum_{k=0}^{N/2-1} Z_k \exp(-4\pi ipk/N). \tag{45}$$

The required result, A_p, can be written in terms of the even and odd numbered points:

$$A_p = \sum_{k=0}^{N/2-1} \{Y_k \exp[-4\pi ipk/N] + Z_k \exp[-2\pi ip(2k+1)/N]\} \tag{46}$$

$$= \sum_{k=0}^{N/2-1} Y_k \exp(-4\pi ipk/N) + \exp(-2\pi ip/N) \sum_{k=0}^{N/2-1} Z_k \exp(-4\pi ipk/N) \tag{47}$$

that is,

$$A_p = B_p + \exp(-2\pi ip/N) C_p. \tag{48}$$

For values of p greater than $N/2$ the Fourier transforms B_p and C_p repeat the values for p less than $N/2$. Therefore substituting $(p + N/2)$ for p gives

$$A_{p+N/2} = B_p + \exp[-2\pi i(p + N/2)/N] C_p \qquad 0 \le p \le N/2 \tag{49}$$

$$= B_p - \exp(-2\pi ip/N) C_p. \tag{50}$$

Therefore the first $N/2$ and the last $N/2$ points of an N point time series can be obtained by transforming separately two series each having $N/2$ points.

This method clearly allows further subdivisions so long as the new smaller time series contain a number of points which is a power of two. The Fourier transform of a single point is itself and so a Fourier transform of an N-point time series has been reduced to a number of complex multiplications and divisions. In general $N \log_2 N$ complex additions or subtractions and at most $(N \log_2 N)/2$ complex multiplications are required. The Cooley–Tukey approach also introduces fewer rounding errors than a conventional slow Fourier transform.

Since the first presentation of this algorithm a number of developments and variations have been documented and some have been implemented on various FT NMR data systems.[46–48] The most important of these is attributed to Bergland[48] which preserves the order and symmetry of the Cooley–Tukey process but which gives a factor of two reduction in computation and storage when the data is real.

4.3. *High Dynamic Range Fourier Transforms*

4.3.1. *Noise in the Fourier transform process.* We need only to consider this in the high dynamic range

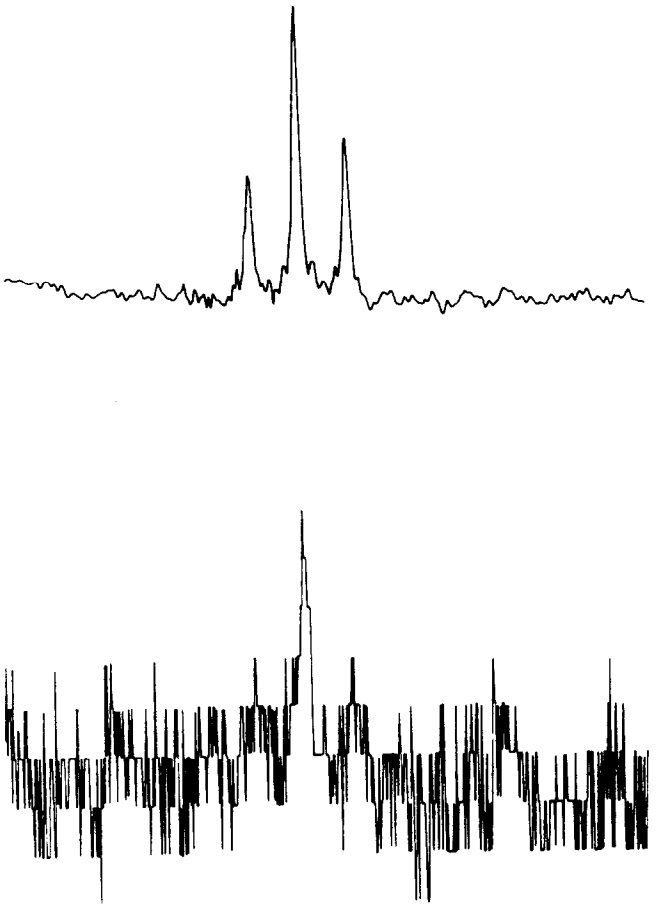


FIG. 18. The methyl triplet in the 90 MHz ^{1}H NMR spectrum of a very small quantity of ethanol dissolved in H_2O. The lower trace is the result of Fourier transforming the FID using a 16 bit computer. The upper trace is the result from a 32 bit transform in the same 16 bit computer, with the consequent decrease to half the digital resolution.

case where, for example, a large solvent peak fills the ADC and the noise at the digitisation stage is limited to that introduced by the $\pm\frac{1}{2}$ LSB of the ADC. A small signal of interest defined by only a few bits of the computer word may then be obscured after Fourier transformation by noise introduced by the transform itself. Figure 18 gives an example of this which shows the methyl triplet from a trace of ethanol in H_2O after Fourier transformation using a 16 bit and a 32 bit integer process.

Fourier transform noise arises from rounding errors introduced after any multiplication or addition which would cause the word-length of the computer to be exceeded; this overflow is avoided by scaling (i.e. division by two), and information in the least significant bits is lost.

Cooper[24] has adapted some theoretical work by Welch[49] to predict the maximum dynamic range observable after Fourier transformation, D_{max}

$$D_{max} = (5K/2)2^w 2^{-(M+3)/2}, \qquad (51)$$

where 2^M is the number of data points and w is the computer word-length. K is defined as

$$K = \frac{\text{r.m.s. } S/N \text{ after FT}}{\text{r.m.s. } S/N \text{ before FT}}. \qquad (52)$$

Equation (51) does in fact predict the general proportionality of D_{max} as a function of computer word length but over-estimates the observed dynamic range by a factor of about 100. This is because the method assumes a single, non-decaying, sine wave in the time domain. Since eqn (51) over-estimates the dynamic range, it lends credence to the supposition that a higher dynamic range is possible when the data is in the form of a non-decaying response such as that obtained when using the Hadamard transform technique.[3] This has been questioned recently[50] but from eqn (52) it would appear that if the signal-to-noise ratio before Fourier transformation is less for the Hadamard method compared to the pulse method, then a dynamic range gain may result when using a pseudo-random sequence and the Hadamard technique.

The over-estimation of dynamic range using eqn (51) occurs because real data in a pulse experiment, having an exponential decay, will not fill the computer word at each location throughout the acquisition. Cooper[24]

TABLE 6. Observed variation in dynamic range as a function of the linewidth of the large peak (16k transform, 1000 Hz bandwidth)[24]

Linewidth (Hz)	Observed dynamic range[a] 16 bit word	20 bit word
0.2	3,124	59,600
1.0	12,496	145,512
2.0	15,624	227,368
5.0	30,512	693,888
10.0	38,144	867,360
20.0	45,474	—

[a] A simulated spectrum was obtained by summing exponentially weighted sine-waves of various relative intensities and the dynamic range for a given width of the large peak determined by picking out the smallest observable peak in the Fourier transform generated noise.

has also simulated this latter case and shown that for a sharp line the dynamic range does increase with an increase in linewidth of the large peak (corresponding to less of the data area being full). The results are shown in Table 6.

The other variable in eqn (51) is the number of data points and the observed dynamic range should also depend on this parameter. The dynamic range will depend on the position of the large peak, as that may introduce quantisation errors (Section 2.4) but assuming that it is at the centre of the spectrum, then Cooper's simulated experiments show that the dynamic range may be slightly lower for larger transform sizes.

Similarly it has been suggested[24] that using zero-filling to increase the transform size may introduce noise and thus reduce the dynamic range. However, addition of, and multiplication by, zeroes cannot cause rounding errors and it is not likely that the dynamic range will decrease significantly. Much more noticeable is the fact that for large transforms for a given spectral width, the noise is increasingly better defined leading to an apparent increase. This is illustrated in Figure 19 for a 512 point data table arising from an exponentially weighted computer generated sine wave. The computer rounding errors corresponded to less than one count on average. The upper trace is the result of Fourier transforming the 512 point data table and the lower trace is the Fourier transform result after zero-filling to 16k points. The noise increases from a peak-to-peak value of 4 counts to a peak-to-peak level of 6 counts. Since Fourier transform noise is introduced by rounding and scaling operations any procedure which can minimise distortions caused by such calculations should lead to a higher dynamic range in the frequency domain.

Memory overflow is usually prevented by scaling all numbers such that they are always less than one fourth full scale. This arises because the Fourier transform process involves equations of the form

$$A_1 = A_1 + B_1 \cos\theta + C_1 \sin\theta \qquad (53)$$

As it is easiest in a computer to divide by 2 or 4, it is necessary to divide all numbers by 4 since three variables have to be added. Thus it is possible to lose one to two bits in dynamic range and for 16 bit

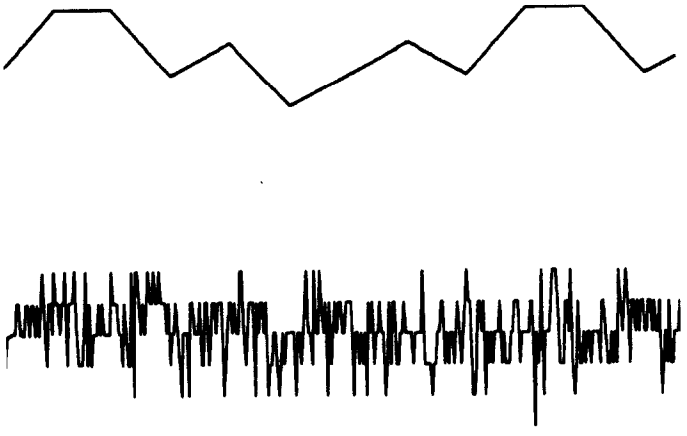


FIG. 19. The effect of zero-filling on the noise introduced by Fourier transformation of a 512 point, computer generated, exponentially weighted sine-wave. The main peak is not shown. The upper trace shows the result of transforming the 512 point data table and the lower trace, the result after zero-filling to 16k. The peak-to-peak noise increases from 4 to 6 counts.

TABLE 7. Computer wordlength requirements as a function of time domain dynamic range for various numbers of scans

No. of scans[a]	No. of bits filled	No. of bits of noise	Min. wordlength for FT[b]	$S/N(f)$ bits[c]	Min. computer wordlength[d]
1	12	1	12	17	18
4	14	2	13	18	19
16	16	3	14	19	20
64	18	4	15	20	21
256	20	5	16	21	22
1k	22	6	17	22	23
4k	24	7	18	23	24
16k	26	8	19	24	25
64k	28	9	20	25	26

[a] $1k \equiv 2^{10}$; this column also assumes a signal-to-noise ratio in the time domain of 2^{12} with adequate digitisation.
[b] this assumes that it is possible to scale the computer word by powers of 2 such that there is one bit of noise prior to Fourier transformation.
[c] this assumes a dynamic range gain in the frequency domain of 2^5 (this is not unreasonable after the application of an optimum filter).
[d] this assumes that no noise is introduced by the Fourier transform, and includes one extra bit for sign.

computers this may lead to large inaccuracies. One way to overcome this problem is only to scale the data when overflow has occurred having previously saved the numbers to be added. This sort of addition with a test at every operation would be so slow as to be useless in FTNMR systems. However, Cooper *et al.*[51] have shown that the simpler procedure of testing for arithmetic overflow (which is often part of the computer hardware) is relatively straightforward and can give a factor of 2–3 in dynamic range for 16 or 20 bit wordlength computers. Arithmetic overflow occurs when the sum of two numbers, having the same sign, has the opposite sign. This occurs because the most significant bit in a computer word is reserved for sign information, i.e. in a 16 bit computer two positive numbers (bit 15 = 0) may have bit 14 set. On addition bit 15 would be set and this indicates a change in sign.

4.3.2. *Double length and floating point format Fourier transforms.* If the data has been accumulated using a double precision acquisition routine then the above minor improvements in dynamic range may not be enough to give all the possible information present in the FID. To summarise this, therefore, a single precision integer Fourier transform will be sufficient in all low dynamic range cases. Here single precision is used to mean Fourier transformation using integer maths in a computer with a wordlength not greater than 24 bits. Note that a 16 bit word gives a resolution in the vertical axis of $\approx 1:2^{15}$ or 1:32768.

In a high dynamic range situation the desired wordlength is related to the output dynamic range and indirectly to the input ADC. Thus for a single pulse on 100% H_2O where a signal-to-noise ratio in the frequency domain of $>2^{17}$ is possible a longer wordlength than 16 bits is always required to define noise and small signals properly. For a 16 bit wordlength computer the Fourier transformation will either have to be in double length integer or in floating-point format.

It is illuminating to observe at which computer wordlength and time domain dynamic range, double precision working becomes necessary. Table 7 shows the expected results for the case of an FID with a dynamic range of 2^{12} (i.e. this implies at least a 12 bit ADC) which is typical of the dynamic range expected when measuring ^1H NMR spectra in H_2O. Column 2 shows the numbers of bits filled on a backing store assuming the availability of double precision acquisition software. Column 4 shows the minimum wordlength to be transformed without loss of significant signal information and if it is assumed that the dynamic range is a factor of 2^5 greater in the frequency domain then column 6 gives the minimum computer wordlength if a single precision integer Fourier transformation is to be performed.

One conclusion from this table is that for an initial time domain signal-to-noise ratio of 2^{12} if 4096 scans do not give a good enough signal-to-noise ratio on the small signals of interest, then a 24 bit integer Fourier transform will not be adequate. A practical demonstration of the inadequacy of performing a single length Fourier transform is shown in Figure 20.

Of the two forms, either double length integer or floating point, it is easier to envisage the double length integer method. Here two computer words are used to define each data point and normal integer maths is used. The disadvantages are that if double length zero filling is not possible because of lack of computer memory a further factor of two is lost in the frequency digital resolution. A second disadvantage is the extra time taken to perform the Fourier transform calculation, especially if the computer does not contain a hardware multiply/divide unit. In the FT software, again to ensure that overflow does not occur the

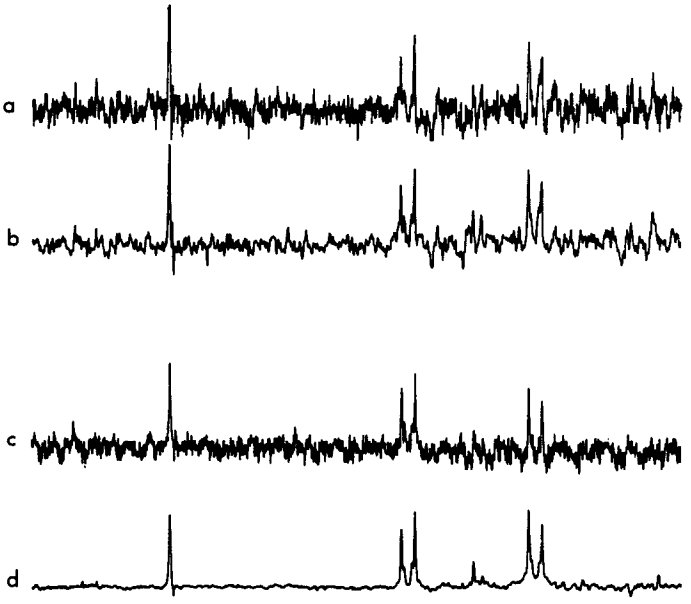


FIG. 20. A comparison of the effects of Fourier transformation using 16 bit integer and 32 bit integer routines. The 90 MHz spectra shown are of the aldehyde and aromatic resonances of 0.1% *p*-dimethylaminobenzaldehyde in H_2O. The top trace (a) shows the result of acquiring 20 scans and Fourier transforming using 16 bits. The corresponding result using a 32 bit Fourier transform is shown in trace (b). A significant improvement is observed showing that some of the noise is generated in the Fourier transform calculation. Trace (c) shows the result of transforming the sum of 700 scans using a 16 bit transform. Since little increase in the signal-to-noise ratio is observed most of the noise must arise from the transform. Trace (d) confirms this by processing the same data using a 32 bit transform.

facility for multiplying two double length numbers to give a quadruple length result before scaling must be included. If a disc-based acquisition sequence is available then the former disadvantage of lack of digital resolution disappears and if it is also possible to perform the Fourier transformation in a time-shared mode then the latter disadvantage of speed is also largely overcome.

Using a high-field spectrometer with a 16 bit digitiser which in the high dynamic range case could give a signal-to-noise ratio in the time domain of 2^{16} on a single scan, and also assuming a 2^5 gain in dynamic range after Fourier transformation, then a 40 bit integer transformation routine would still be more than adequate allowing up to about 16 million scans.

For a given wordlength computer, higher dynamic range data can be handled if the Fourier transformation routine allows the data to be converted to the floating-point format. For a single computer word the most significant bit is retained for the sign, and the rest of the word is split into two areas, one set of bits defining a mantissa (between 0 and 1) and the other defining an exponent (10^x). For example in a 16 bit computer word reserving 5 bits for the exponent and 10 bits for the mantissa allows the representation of all numbers between 0 and 2^{32} (since $32 = 2^5$) or 0 and 2.14×10^{10}, obviously a much greater range than for 16 bit integer in which the numbers are limited to between 0 and 2^{15}. Secondly, for the floating-point method all numbers (both large and small) are defined to the same absolute accuracy. In integer representation all numbers are accurate to one count which would be a large percentage for a small number. For a 16 bit or a 20 bit word use of the single word floating point method as defined above gives an increase in dynamic range over the integer method of between two and four bits.[46] Since in the single length floating point format the number of bits defining the precision of a number has been reduced by five to accommodate the exponent it is necessary to question whether the accuracy of the Fourier transform result is affected. For this point Cooper[51] has shown that the intensities are seriously affected for a 16 bit (10 bit mantissa) Fourier transform but that these discrepancies largely disappear for a 20 bit floating-point transform (14 bit mantissa). Cooper *et al.*[51] have also tabulated the times required for various types of Fourier transform using the same minicomputer in each case (PDP 11/40) and these are given in Table 8.

TABLE 8. Times (seconds) for various Fourier Transform methods using a PDP 11/40 computer[51]

Transform size (k)	Integer			Floating point
	Scale on each pass	Scale on hardware overflow	Scale on software overflow	
1	1.17	1.67	2.40	5.14
2	2.60	3.69	5.29	11.27
4	5.69	8.09	11.58	24.53
8	12.45	17.63	24.75	53.25
16	27.00	38.11	54.24	115.6[a]

[a] extrapolated value

However, a single word floating-point method cannot give the same dynamic range as a double length integer transform and so on commercial systems a floating-point representation is used in which each data point is defined by two computer words. As can be seen from Table 8 for a single length transform the floating-point representation has a time penalty and so if a double length transform is necessary, then as both double length integer and double length floating-point provide more than enough dynamic range it is preferable to choose the integer representation which is faster and easier to program.

One final point is that since the Fourier transform uses equations such as eqn (57), then the sin θ, cos θ terms should also be double length if the full dynamic range gain is to materialise.

4.4. *Two-Dimensional Fourier Transforms*

A large number of papers has already appeared on the many experimental aspects of 2D-NMR largely developed by the groups of Freeman[10] and Ernst[44] and we do not propose a discussion of them here, but we feel this an appropriate point to include an outline of the computation involved in obtaining an NMR spectrum which is a function of two frequencies. With the advent of disc-based operating systems allowing the use of virtual memory the time-consuming and cumbersome operation associated with attempting to transform large arrays in a limited memory should disappear.

In order to describe the computation involved let us consider the case of a 2D-spin echo experiment with t_1 as a variable time parameter defining the delay between the 90° pulse and the centre of an echo following the 180° pulse and t_2 being a second time variable running from zero to T, the acquisition time. The signal detected after a 2D-NMR experiment will be a function of both t_1 and t_2, $S(t_1, t_2)$. This is first transformed with respect to t_2 and then with respect to t_1 to produce the 2D spectrum $S(\omega_1, \omega_2)$. In the case given above ω_2 represents the normal frequency axis of NMR and ω_1 represents a frequency axis in which only J couplings with lines at the natural line-width appear.

To obtain $S(\omega_1, \omega_2)$ the following operations are performed.

(i) $S(t_1, t_2)$ is treated as a conventional FID. It can be exponentially weighted and Fourier transformed to give $S(t_1, \omega_2)$. This gives a spectrum with the correct frequencies but phases and intensities are modulated as a function of t_1.

(ii) The sine and cosine transforms of $S(t_1, \omega_2)$ are stored (S^s and S^c respectively).

(iii) The parameter t_1 is then altered and the whole experiment repeated, if necessary including multiple scan averaging, until a series of N transformed spectra $S(t_1, \omega_2)$ is obtained.

(iv) For all N spectra the scaling which occurs during the Fourier transform must be kept constant and thus the next stage in the computation is the overall normalisation of all the spectra taking into account the different scalings which must have occurred in the different spectra.

(v) The data at the stage of $S(t_1, \omega_2)$ are stored on backing disc as a series of spectra $S(\omega_2)$. These then have to be transposed to give a series of interferograms $S(t_1)$. The term interferogram has been suggested by Freeman[10] to distinguish the time function $S(t_1, \omega_2)$ from a free induction decay.

(vi) A second Fourier transformation is then performed to produce $S(\omega_1, \omega_2)$ the full two-dimensional frequency spectrum.

(vii) The usual requirements still hold for the sampling rate and the resolution is still governed by this and the number of data points. Also exponential weighting to improve sensitivity or resolution can be performed before either Fourier transformation.

(viii) The matrix $S(\omega_1, \omega_2)$ in fact consists of four parts: The cosine part of the cosine transform, the sine part of the cosine transform and the corresponding sine transforms. For spin–echo spectra all four quadrants must be retained if the signs of the frequencies need to be distinguished.

(ix) Phase adjustment of 2D J-resolved spectra may present difficulties (Section (5.1.4) and absolute value mode spectra are usually displayed instead.

Figure 21 shows an example of 2D FTNMR. This is the spectrum of a mixture of 45% ethyl chloride, 25% ethyl bromide and 30% ethyl iodide. The top trace (a) is the normal 60 MHz ^{1}H spectrum of the mixture. The centre block (b) is the result $S(\omega_1, \omega_2)$ of a Carr–Purcell type A spin-echo experiment (i.e. 90°–$\frac{1}{2}t_1$–180°–$\frac{1}{2}t_1$–acquire for t_2) with variation of t_1. Taking sections parallel to the ω_1 axis it is possible to pick out all three CH_3 triplets and all three CH_2 quartets in the correct (45 : 25 : 30) intensity ratios.

Since the spectrum of each compound is first-order, it is possible to take a projection on to the ω_2 axis which results in all the signals from a given multiplet falling on top of each other and giving a proton NMR spectra free from proton–proton coupling. This is shown in trace (c).[52]

5. MANIPULATIONS AFTER FOURIER TRANSFORMATION

5.1. *Phase Correction of Spectra*

5.1.1. *Introduction.* It has long been recognised that the Fourier transform obtained from a phase-sensitive detected free induction decay will in general consist not only of the v-mode spectrum and that phase correction is required to compensate for the transfer function of the detection system. This arises primarily

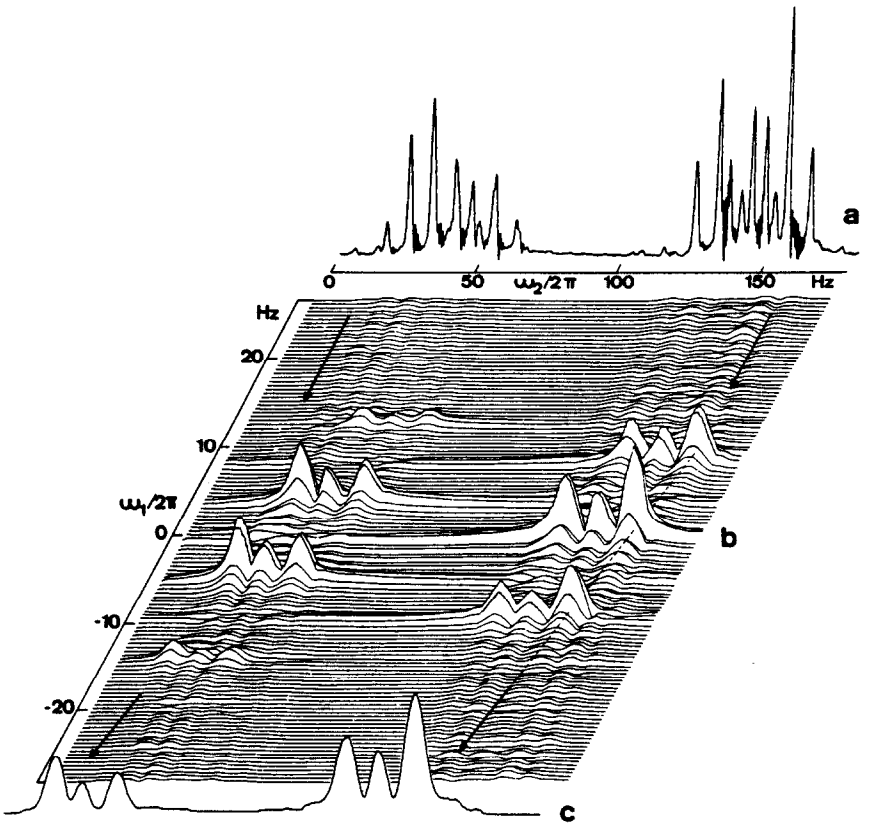


FIG. 21. An example of a two-dimensional J resolved spectrum obtained by double Fourier transformation following data collection using a Carr–Purcell type A pulse sequence[52] (a) A 60 MHz ^{1}H NMR spectrum of a mixture of 45% C_2H_5Cl. 25% C_2H_5Br and 30% C_2H_5I. (b) A 2-D NMR spectrum computed from 64 single echoes represented by 64 sample values. (c) A broad-band decoupled spectrum obtained by projecting the 2-D spectrum on to the ω_2 axis. A coarse digitisation of the 2-D spectrum was used to partially suppress the background signals. (Reproduced from *J. Chem. Phys.* **64**, 4226 (1976) with permission of the copyright holder).

from the hardware filter network used to discriminate against noise being folded back into the spectrum window and by the pre-acquisition delay necessary to avoid pulse break through and amplifier dead time. A transfer function, $h(\omega)$, of the type

$$h(\omega) = k \exp i(a_0 + a_1\omega) \tag{53}$$

is appropriate since it is obeyed by ideal Butterworth filters below cut-off.[53]

In commercial spectrometers phase correction is conventionally achieved for the ith point of the absorption (A) and dispersion modes (D) as

$$Y_A(i) = R_i \cos \theta_i - I_i \sin \theta_i \tag{54}$$

$$Y_D(i) = R_i \sin \theta_i + I_i \cos \theta_i. \tag{55}$$

R_i and I_i are the ith values of the real and imaginary parts of the Fourier transform spectrum. Two angles are required, a zeroth order (ϕ_0) and a first order (ϕ_1) and then

$$\theta_i = \phi_0 + 2i\phi_1/N. \tag{56}$$

The two angles ϕ_0 and ϕ_1 are converted by the computer into binary numbers (say 12 bits) and used as indices for a sine look-up table, which conventionally takes up some of the available computer memory.

After phase correction it is usual to display the absorption spectrum (v-mode) but the dispersion (u-mode) spectrum is usually also available. An alternative display mode is that of the magnitude or absolute value spectrum, $(v^2 + u^2)^{1/2}$. This approach, which is independent of phase angles, is useful when measuring spin–echo spectra of coupled systems and in many applications of 2-dimensional FTNMR.

Recently a number of papers have appeared on the analysis of the pattern produced when an absorption spectrum is plotted against its dispersion component.[54–57] This presentation apparently provides a means of distinguishing various line broadening mechanisms since for a pure Lorentzian the result should be a semi-circle. Deviations from the semi-circle allow a distinction between, for example, an unresolved spread in chemical shifts and chemical exchange broadening. More recently the effects of digitisation, noise, truncation and zero-filling on the accuracy of the results, have been reported.[56]

5.1.2. *Sine look-up tables.* It is only necessary to store the sine values between 0 and $\pi/2$ and then for a given setting of a phase angle potentiometer the knob value is read by an ADC and converted to the appropriate binary sequence, say b. Thus for an N point spectrum with $c(b)$ representing the contents of location b:

$$0 \leq b < N/4 \quad \sin\theta = c(b) \tag{57}$$

$$N/4 + 1 \leq b < N/2 \quad \sin\theta = c(N/2 - b) \tag{58}$$

$$N/2 + 1 \leq b < 3N/4 \quad \sin\theta = -c(b) \tag{59}$$

$$3N/4 + 1 \leq b < N \quad \sin\theta = -c(N/2 - b). \tag{60}$$

Consequently only $N/4$ sine table entries are needed for $2N$ points in the original time domain decay since half the points are imaginary after the Fourier transformation.

Interpolation between the entries in a sine look-up table is also possible and this leads to the question of how few entries are needed before distortions and artefacts are introduced. Cooper *et al.*[51] have performed calculations on a synthesised high dynamic range 16k data point spectrum with a sine look-up table varying in length from 4096 down to 16 entries. Linear interpolation was used between entries and no difference was detected in the resulting spectra for any length of the sine table after making allowances for errors detected in the exponentiation routine of the DEC-10 FORTRAN package used to program the calculation. In fact Kaiser[58] has analysed the errors expected and shown that they should be of the order of a few bits in a 16 bit word even for a sine table length of 32 entries for a 16k data table. The error, $\Delta(x)$, introduced by linear interpolation of a function $y(x)$ between values $y(x_n)$ and $y(x_{n+1})$ is a second order polynomial that is zero at x_n and x_{n+1}

$$(x) = \frac{y''(\bar{x})(x - x_n)(x - x_{n+1})}{2}, \tag{61}$$

where $y''(\bar{x})$ is the second derivative at some $\bar{x}$ between x_n and x_{n+1}. For an FID $f(t)$ with a frequency spectrum sine transform $F(\omega)$ that which is computed is the discrete form of

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)[\sin(\omega t) - \Delta(\omega t)]\, \mathrm{d}t. \tag{62}$$

The error in the Fourier transform is then within E where

$$E \leq \Delta_{\max} \int |f(t)|\, \mathrm{d}t. \tag{63}$$

Taking $h = x_{n+1} - x_n$ and for a sine table of M entries

$$\Delta_{\max} = \frac{h^2\pi^2}{256M}. \tag{64}$$

Thus

$M = 512$	256	128	64	32
$\Delta_{\max} = 2^{-20}$	2^{-18}	2^{-16}	2^{-14}	2^{-12}.

The integral $\int |f(t)|\, \mathrm{d}t$ is less than the sum of all lines in the spectrum. The calculation should be repeated for each of the $\log_2 N$ passes of the fast Fourier transform algorithm. Thus the total error E might be multiplied by $\log_2 N$ or if scaling occurs between all passes, by a factor of two. For a quarter length sine-table with 128 entries an error of the order of 2^{-15} of the true spectrum is expected whilst for a 64 point entry the error should be four times greater, but these errors are of the same order of magnitude as the noise introduced by the Fourier transform itself for a 16 bit word and may only be noticeable when using longer word length computers.

Akitt[59] has recently proposed a trigonometric based interpolation using a 64 entry sine table in order to save on memory space. This interpolation of course takes longer than the linear method but calculates sines and cosines in one pass. Accuracy is again to about 2^{-14} or 2^{-15} and the total program plus table only occupies 128 words.

5.1.3. *Automatic phase correction.* The possibility of performing automatic phase correction was first raised by Ernst[60] who demonstrated the results of using the Hilbert transform. Hilbert transformation relates the real and imaginary parts of the frequency response function of a linear system (e.g. a single pulse Fourier transform NMR experiment with the spin system at equilibrium). Because the Hilbert transform effectively provides a 90° phase shift, in order to gain some arbitrary phase change it is necessary to take a linear combination of the original spectrum with a proportion of the Hilbert transform. Since the dispersion component of a spectrum has zero integrated intensity then the proportions used in the linear combination of spectra are varied until a maximum intensity is found, giving a pure absorption spectrum. However, this method was found to be very unstable as a result of spectral noise and because of deviations from a level baseline.

Another possible method for fully automatic phase correction has been demonstrated by Neff, Ackerman and Waugh[53] using a method that requires precalibration of the detector transfer function. This method has the advantage of removing artefacts arising from mismatched channels when using quadrature phase detection.

5.1.4. *Phase correction in 2D-NMR.* In Section 4.4 we have given a description of the procedure involved in obtaining a two dimensional NMR spectrum which appears in four quadrants denoted by $S^{cc}(\omega_1, \omega_2)$, $S^{cs}(\omega_1, \omega_2)$, $S^{sc}(\omega_1, \omega_2)$ and $S^{ss}(\omega_1, \omega_2)$ where c, s refer to cosine and sine parts of each Fourier transform. Clearly any phase adjustment of a 2D spectrum is going to be more complicated than that for a 1D spectrum and to overcome this, in many cases it is usual to present an absolute value mode spectrum. However, to phase a 2D spectrum involves taking combinations of the four parts of the spectrum with angles which, as in normal NMR, involve two parameters, a constant term and a frequency dependent term for each frequency dimension ω_1 and ω_2.

Since absorption mode spectra have narrower line-shapes and less distortions from overlapping transitions, efforts have been made to alter the line-shapes such that phasing is possible. For spin–echo spectra obtained from linear combinations of the four quadrants, absorption lines show the phenomenon of "phase-twist" which results from a resonance posses-

sing different proportions of absorption and dispersion components in progressive slices through the line. Several experimental methods[65] have been proposed to modify the natural 2D-lineshape to give a double Lorentzian form

$$S^c(\omega_1, \omega_2) = A \cos(\alpha - \theta_1) \cos(\beta - \theta_2)/4D_1D_2 \quad (65)$$

where

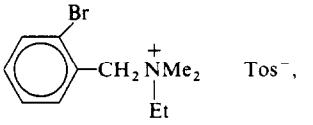
$$D_1^2 = \lambda_1^2 + (2\pi\Delta\omega_1)^2; \quad D_2^2 = \lambda_2^2 + (2\pi\Delta\omega_2)^2$$

$$\sin\alpha = 2\pi\Delta\omega_1/D_1; \quad \sin\beta = 2\pi\Delta\omega_2/D_2,$$

with λ_1, λ_2 as the exponential time constants for the two decays in the time domain spectrum. The angle $\theta = \theta_1 + \theta_2$ represents a phase error introduced by the instrument. In general this will be an unknown function of the two frequencies ω_1 and ω_2. Levitt and Freeman[61] have recently described a practical method for calculating the phase corrections required. However, the computations are lengthy and involved and if the sole reason for using an absorption presentation is to narrow the spectral lines then Gaussian deconvolution with the absolute value presentation provides an alternative.[42,65]

5.2. *Spectrum Subtraction*

In general spectrum subtraction can be a very useful technique when studying mixtures or for quantitative measurements on small molecules, which give rise to sharp lines, in the presence of macromolecules which give only relatively broad resonances. Also subtraction can be performed on both time and frequency domain spectra, the former having dynamic range advantages. This point is illustrated in Figure 22 which shows the methyl region of the spectrum obtained from the antihypertensive agent bretylium tosylate dissolved in water, containing a small amount of ethanol.



Trace (a) shows the total methyl spectrum with vertical expansions of × 32 and × 200. The large triplet arises from the methyl protons of the ethyl group in: and the small triplet is from the methyl protons in residual ethanol. In (b), the top trace repeats the highest expansion in (a) where the vertical bit levels are clearly seen in the noise. This results from using a 16 bit Fourier transform process which provides the dominant source of the noise. The centre trace results from subtracting a spectrum of the pure compound in the frequency domain. The bottom trace shows the results of subtracting the two time domain spectra acquired under identical conditions. This causes the time domain dynamic range to be lowered and hence the noise introduced by the 16 bit Fourier transform is reduced to within the spectral noise level. The traces in (c) correspond to those in (b) and give a comparison using a 32 bit Fourier transform routine. Here Fourier transform noise is insignificant in both cases.

5.3. *Baseline Correction*

5.3.1. *Introduction.* Most forms of baseline correction use a method which requires subtraction of a calculated function from a spectrum. This is a particular application of spectrum subtraction discussed in the previous section. The various methods documented in the literature are described in this section and some comparisons are presented.

5.3.2. *Definition of baseline.* Instrumental instability

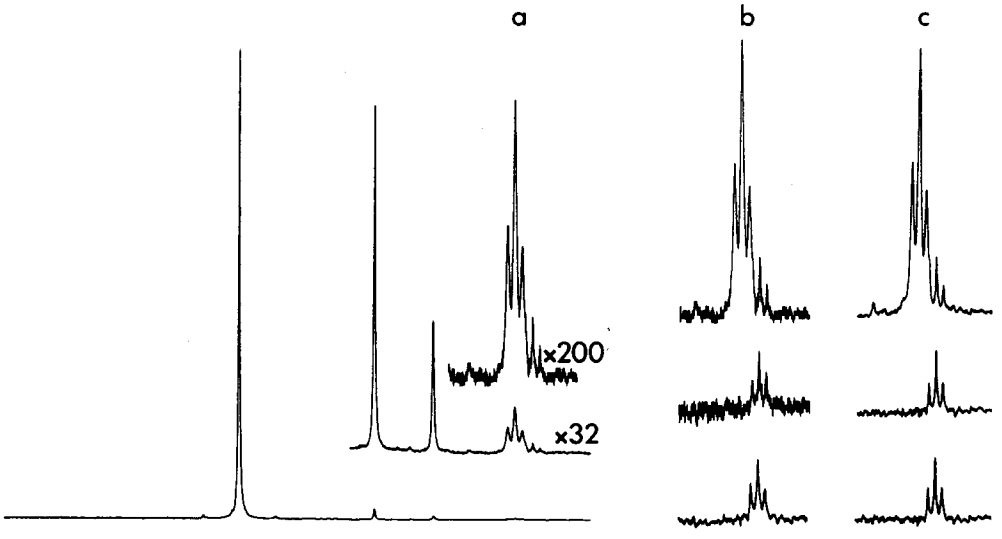


FIG. 22. Demonstration of spectrum subtraction in time and frequency domains using a sample of bretylium tosylate (see text) in H_2O containing a small amount of ethanol. Trace (a) is the total methyl spectrum where the small triplet is from ethanol. The top trace in (b) repeats that in (a) and shows the noise introduced by the transform. The centre trace shows the result of subtracting a spectrum of the pure compound in the frequency domain and the bottom trace the time domain subtraction. The traces (c) correspond to those of (b) and give a comparison using a 32 bit Fourier transform.

and spurious resonances (e.g. probe ringing) can give rise to broad variations in a spectrum baseline. In order to improve the accuracy of any subsequent data reduction it is desirable to remove these broad envelopes. A number of computational methods have been developed for Fourier transform NMR data packages but in general they all rely on the same procedure. Firstly the operator identifies regions of true baseline, secondly, a smooth curve is fitted to these regions and this computed curve is then subtracted from the original spectrum. This method has been used with much success in the authors' laboratory and in this procedure a series of channel addresses and their contents are selected by the operator. In case any of the channel addresses which are taken to be true baseline in fact contain an unnoticed spurious value, e.g. a large noise spike, it is possible to take the average value from the address chosen plus a previously specified equal number either side. Next, a curve is fitted to the selected averaged intensity values using the Lagrange interpolation method, where $Y_b(f)$, a polynomial of degree n, is the calculated intensity value at address f,

$$Y_b(f) = \sum_{i=0}^{n} \frac{Y_b(f_i)W_n(f)}{(f-f_i)W_n'(f_i)}, \tag{66}$$

and f_i are the $n+1$ chosen addresses.

$$W_n(f) = (f-f_0)(f-f_1)(f-f_2)\ldots(f-f_n) \tag{67}$$

$$W_n'(f_i) = \prod_{\substack{k\neq i \\ k=0}}^{n} (f_i - f_k). \tag{68}$$

The correction is performed by replacing $Y(f)$ in each channel f by $Y'(f)$ where

$$Y'(f) = Y(f) - Y_b(f). \tag{69}$$

For correction of very steep slopes, it is necessary to use only the address chosen and not to average over adjacent locations. Clearly, the calculation of a polynomial of degree n where n may be ≥ 20 would involve a great deal of computational effort and the usual approximation is to calculate a running average cubic between any four adjacent points: i.e. a cubic is fitted to the first four points chosen as baseline and values are calculated for the baseline between points two and three. Then point one is dropped and the fifth chosen baseline point is included, a new cubic is calculated and is used to give the y values for the baseline curve between points three and four. This process is repeated until a set of y values is obtained for the whole region of interest. These are then subtracted from the spectrum to give the baseline flattened result.

A much simpler method of baseline correction used in some commercial data systems has only two variable parameters, namely a d.c. offset and a slope correction.

Other methods of baseline correction are also available and examples of all the methods are shown in Figure 23. This diagram shows the ^{1}H NMR spectrum of benzene dissolved in the nematic mesophase, Phase V. The top trace (a) is the experimental spectrum showing the broad background from the liquid crystal solvent. By choosing points by eye to be baseline the polynomial curve in (b) can be calculated. The difference (a) − (b) is shown as trace (c). Trace (d) is the convolution difference result obtained by subtracting a proportion of a broadened form of (a) from itself. Trace (e) is the result of applying a Gaussian deconvolution function to severely over-enhance the sharp lines and then to display the absolute value mode spectrum.[42] Finally, the lowest trace (f) is obtained when the FID is left shifted by 15 locations to remove most of the intensity information apertaining to the broad components. Note that the spectrum contains an artefact, marked with an asterisk, which appears to have a dispersive phase. Magnitude deconvolution removes this peak presumably because its absorption component is broader than the lines arising from the benzene. This must be taken as a caveat that spectral information can be lost during baseline flattening exercises.

Attempts have been made to automate the first step of any baseline correction procedure, i.e. to determine the regions of true baseline without the intervention of the operator. Pearson[62] has suggested an iterative automatic method of baseline correction based on the assumption that a point in a digitised spectrum is considered to be part of the baseline if it lies within $\pm v\sigma_1$ of the x-axis where σ_1 is a "baseline standard deviation" and v is a positive constant (in practice between 2.0 and 3.0). The term σ_1 is the iteratively calculated standard deviation of the Y values which lie within $\pm u\sigma_1$ of the x-axis.

$$\sigma_1 = \left[\frac{\sum_i [Y(f_i)]^2 h(u\sigma_1 - |Y(f_i)|)}{1 + \sum_i h(u\sigma_1 - |Y(f_i)|)}\right]^{1/2}, \tag{70}$$

where $h(z) = 1$ if $z > 0$ or $h(z) = 0$ if $z < 0$. The value of u in practice is taken to be 4.0. The value of σ_1 converges to a well defined limit and then a smooth function $g(f)$ is fitted to this automatically derived baseline (i.e. $-v\sigma_1 < Y(f_i) < +v\sigma_1$) which is subtracted from the original spectrum to give the flattened result. The algorithm has the flowchart given in Figure 24. A correction was found to be negligible if $|g(f)| < 0.125\sigma_1$ for all values of f. In order to provide the most efficient use of the computer $g(f)$ is expressed as a linear combination of a finite ortho-normal set of functions.

$$g(f) = \sum_j C_j L_j(f) \tag{71}$$

i.e.

$$\sum_i L_j(f_i)L_k(f_i) = \delta_{jk}, \tag{72}$$

where i runs over all N f_i values in the spectrum.

$$C_j = \sum_{i'} Y(f_i)\, L_j(f_i) \tag{73}$$

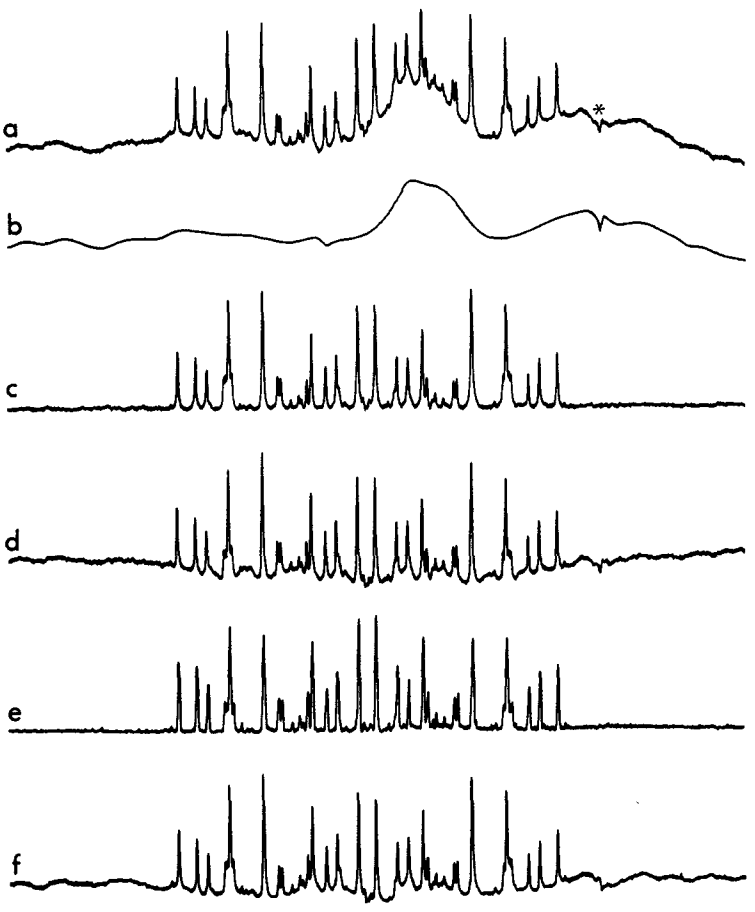


FIG. 23. A demonstration of the various methods of achieving baseline flattening. (a) The 90 MHz ^{1}H NMR spectrum of benzene dissolved in the nematic mesophase, Phase V; 6250 Hz spectral width, 24°C, 8192 data points. (b) A theoretical polynomial calculated by fitting points chosen to be baseline. (c) The result of (a) minus (b). (d) The convolution difference result obtained by subtracting a proportion of a broadened form of (a) from itself to yield approximately the same signal-to-noise ratio. (e) The result of applying the Gaussian deconvolution weighting function and presenting the absolute value spectrum. (f) Left shifting the FID by 15 data points before Fourier transformation.

where i' runs over only those values for which $|Y(f_i)| \leq v\sigma_1$. This method of defining the baseline has been used in conjunction with a cubic polynomial and an eight point Fourier coefficient function.[62]

5.4. *Spectrum Smoothing*

Occasionally, in order to save remanipulation of an FID or for reducing output during line-listing, it is advantageous to smooth a Fourier transformed spectrum and this is achieved through the use of an m-point moving average method; i.e. for a three point moving average each data point is taken to be the sum of half its own contents plus a quarter of those on both sides of it, the average being calculated progressively through the spectrum. Usually the operator specifies m, the number of data points for the average. In the authors' laboratory a complementary approach has been taken in that a triangular smoothing function has been used in which it is possible to specify the width of the triangle at half height in Hz. The smoothing for a particular point is then taken over the nearest odd number of channels $(2m + 1)$ below the specified width, the apex of the triangle being at the middle point (n). The smoothing function is

$$F'(n) = \frac{\sum_{f=n-m}^{n+m} F(f)\left[\frac{m-|n-f|}{m}\right]}{\sum_{f=n-m}^{n+m}\left[\frac{m-|n-f|}{m}\right]} \quad (74)$$

Mathematically, the operation represents convolution with a triangle and if repeated several times to provide greater and greater smoothing, the peaks will eventually resemble triangles.

5.5. *Integration*

Integration, the measurement of peak areas, is possible if the spectrum is presented in absorption mode, phase corrected and baseline flattened. The integration trace is simply the result, at any location, of adding together the contents of all previous locations.

Programs available on commercial systems usually allow the operator to improve an integral trace by

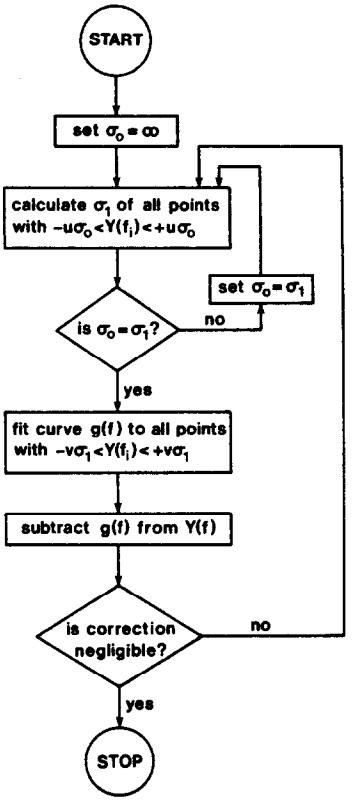


FIG. 24. Flowchart describing an automatic method for baseline flattening.[62]

applying a d.c. offset correction to remove any slope and a ramp correction to remove simple curvature. In most cases it is possible to print out integral values and at least one system allows a least squares fit of chosen integral data to predict the number of nuclei giving rise to each band.[63]

Many users only require the accuracy of integral data to be such that it is possible to distinguish the number of nuclei contributing to a particular resonance, but NMR is now increasingly used for quantitative assays and a discussion of the potential accuracy of integrals is of paramount importance. It is possible, if the conditions of the experiment are ill-chosen, for an integral value to depend on the number of computer word bits defining the peak height and on the number of spectral locations over which the integral is measured. Figure 25 shows the integral measured, I_M, compared to the true integral, I_T, as a function both of the number of counts of signal defining a peak and of the number of spectral channels, I_0, over a line.

The diagram was produced by digitising a Lorentzian line out to ± 10 times the linewidth at half height (at this point the band height has reduced to about 1% of the maximum) using up to 8 bits with various numbers of points over this frequency spread. The shaded areas show the regions into which the measured integral value would fall depending on the position of the steps in the frequency axis relative to the peak maximum.

A number of conclusions can be drawn from this figure. Firstly, in order to obtain an absolute accuracy of about 2% the signal must be defined by 6 bits. An error of approximately 0.7% results if the signal is defined by 8 bits. Of course, for a ratio of signal heights, provided that they are similar, the errors are much less. The number of data points over which the integral is taken can also be a factor contributing to an inaccurate result. For example, for a signal defined by 8 bits

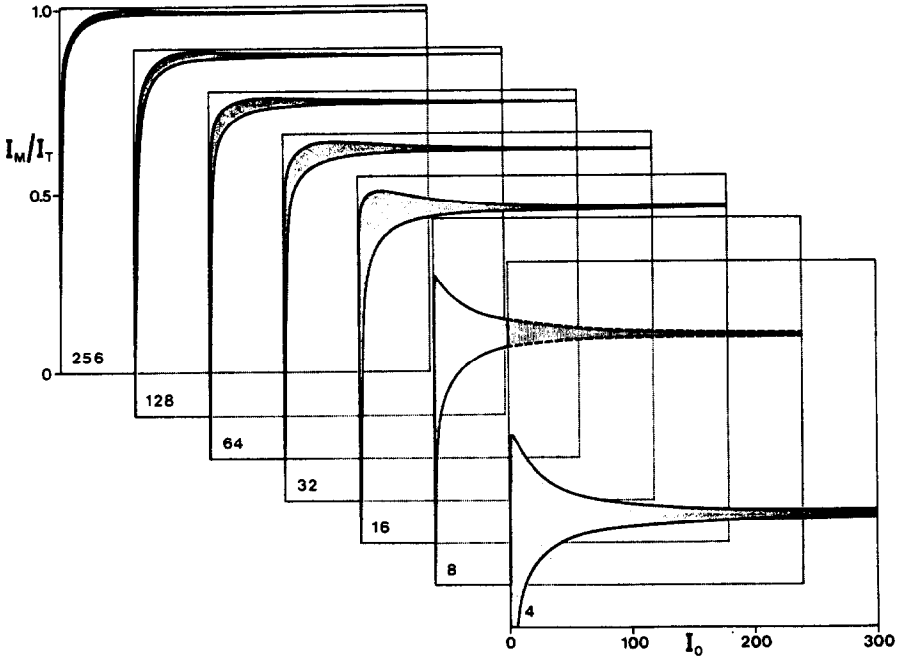


FIG. 25. The accuracy of digital integration as a function of the number of data points, I_0, over a Lorentzian peak and as a function of the number of counts defining the signal height. I_M/I_T represents the ratio of the measured to the true intensity.

integrated over 60 points the measured area will be low by up to 1.4% and for a ratio of two similar signal heights, up to 2.8%. If one wishes to measure the ratio of the areas of two signals with dissimilar heights, then the result may be in serious error. There will be a large error band on the measured small signal height when using only a few data points and therefore it is necessary to integrate over as large a number of points as possible. If this is achieved, then there will be an increasing error in the measured area ratio as the number of bits defining the small signal decreases. For example, if one wishes to measure the area of some small peak by comparing it with that from a large resonance, there could be a stringent dynamic range requirement. For a small peak only 4 bits high and a large peak of >8 bits the integral will be low by about 8% assuming moreover that there are several hundred points over the peaks. Using a 20 bit word computer for the same example reduces the errors to insignificant proportions. For a real spectrum containing noise there is an additional source of error in that it may be difficult to define regions of true baseline within the noise level.

5.6. *Plotting*

Output from the computer to provide a permanent record is usually to either a recorder stepped digitally in the x-direction with an analog voltage applied to the y-axis or to a full digital plotter with digital output in both x- and y-directions.

The analog voltage for a recorder is output via a digital-to-analog converter, DAC. The simplest design of the voltage output type, consists of a series of resistors connected to an operational amplifier, the resistors being weighted in a binary fashion and each one is connected either to the reference signal or to ground depending on the state of each bit in the computer word to be output. For details of the parameters which define the specification of a DAC, the reader is referred to suitable electronics textbooks.[64]

Usually, output to a recorder does not involve the full computer wordlength—typically a DAC will have a 12 bit resolution—so a computer word will be scaled or offset before output. The plot time is also set by software, the recorder either being stepped by pulses from the computer or by a decreasing voltage obtained via a DAC from a computer word in which the contents are monotonically stepped down.

Two modes of plotting are usually possible. The first method gives a constant speed in the x-direction. This causes the pen to speed up dramatically when it has to draw a large sharp peak, leading to poor peak definition and distortion. This is overcome by the "autoslew" or "constant pen speed" technique. Here the pen movement is slowed down directly in proportion to the distance it has to travel between adjacent y values. For example, in the authors' laboratory, when plotting in this mode the computer takes an average difference between the values in the vertical axis for the current point and an even number of adjacent points on either side. This running average compared to that for the previous data points serves to either speed up or slow down the pen travel by defining for point p a variable rate, R. For a calculation over $2N + 1$ points,

$$R \propto \frac{1}{2N}\left[\left|(y_{p-N} - y_{p-N+1}) + (y_{p-N+1} - y_{p-N+2}) + \cdots + (y_{p+N-1} - y_{p+N})\right|\right] \quad (75)$$

$$\propto \frac{1}{2N}\left[\left|y_{p+N} - y_{p-N}\right|\right], \quad (76)$$

where y_p is the currently plotted value. Point y_{p-N} is then dropped and point y_{p+N-1} included in order to plot point y_{p+1}.

Output is often arranged through a double buffering method such that, for example, 128 words are transferred to one output buffer, and whilst these are being plotted a second 128 words are entered into a second buffer. As soon as the first buffer is empty the second buffer is output and during this operation the next 128 words are read into the first buffer.

Another refinement found on many commercial spectrometers is a rapid up-down movement of the pen in order to get the ink flowing prior to commencement of plotting. Use of computer type graph-plotters has now caused this gimmick for punching holes in paper to wane.

The new, digital, computer plotters being very fast, are suited to the presentation of 2D-NMR spectra. Here a three-dimensional effect is created by stacking a set of spectra each offset in the x and y-directions by a small amount as in a plot for a T_1 determination. The presentation is greatly improved if it is possible to remove those sections of a trace which would otherwise be obscured "behind" another peak. Bodenhausen *et al.*[10] have described this process as "whitewashing" and it involves storing the y values of all previous traces and raising and lowering the pen at the appropriate points after making allowance for any interpolation in the plot output. Although information is not usually lost, this routine must be suppressed if negative peaks are to be observed. An example of "whitewash" is shown in Figure 21.[52]

Finally, Figure 26 shows different methods of improving the plotted output. In many cases the number of spectrum channels output will be many times less than the number of x steps on the recorder. In this case interpolation between the points to be output can lead to a better appearance of the spectrum and in some cases to a better definition of the peak maxima. If no interpolation is carried out a staircase of points is produced; this is improved by using a linear interpolation between points. An even better definition is obtained by fitting a cubic to any four points and outputting an interpolated curve between the middle two points. The process is repeated by dropping the

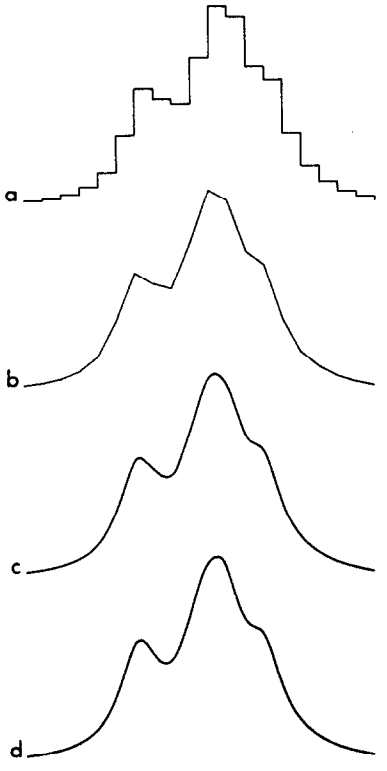


FIG. 26. The effect of interpolation on the appearance and accuracy of a plotted spectrum. (a) No interpolation, (b) linear interpolation between points, (c) cubic interpolation, (d) analog input.

first point and including a fifth. An example is shown in Figure 26 for (a) no interpolation, (b) linear interpolation, (c) cubic interpolation and (d) the analog input.

Finally, some mention should be made that consideration must be given to the number of bits defining both the x and y axes of any plotter. Usually about 12 bits are output in the vertical direction and 13 bits in the horizontal. For computer digital plotters where chart lengths of several metres are possible, spectral interpolation and output to much more than 13 bits would be necessary to provide a reasonable spectral appearance.

5.7. *Peak Print-Out*

All commercial programs have a facility for printing out the positions of peak maxima. The simplest is a straightforward list of those locations, scaled into frequency terms, which break above a given threshold; a peak is defined by the criterion that the locations on either side will have a lower y value. A more accurate method of assessing a peak maximum lies in calculating the centroid position from the highest point and its two adjacent locations, assuming a triangular top to the peak. Other more accurate interpolations have been developed and these can give better estimates of peak maxima (see for example the cubic interpolation for plotting, Section 5.5).

6. CONCLUDING REMARKS AND ACKNOWLEDGEMENTS

In this article we have attempted to highlight the criteria which should be considered before measuring an FT NMR spectrum, to show the limitations of using a digital Fourier transform and to perhaps illustrate how to get the best possible results. Extensive details are available[7,14] for optimising NMR spectral parameters and we have not discussed them here. We have also given quite extensive coverage to the difficulties presented by measuring spectra in a high dynamic range situation and discussed the requirements for successful signal averaging in this case. A number of different types of resolution enhancement functions have been described in the literature and we have attempted a realistic comparison. The implications of using double length integer Fourier transforms have also been set out. It is clear, however, that it is not possible to cover every aspect of even the small practical area of NMR signal processing and new developments will continue to be made particularly as they apply to 2-dimensional FTNMR.

We are grateful to Mr. Alan Strutt for extended discussions on the theory and operation of the electronic modules in a FTNMR spectrometer. Thanks are also due to Mr. A. J. Wyatt of Kratos–Instem Ltd., for the implementation in our FTNMR data system of many of the routines described here. We also acknowledge the encouragement of Dr. A. J. Everett during the preparation of this article and we thank the Wellcome Foundation Ltd., for permission to publish this material.

REFERENCES

1. J. DADOK and R. F. SPRECHER, *J. Magn. Reson.* **13,** 243 (1974).
2. R. K. GUPTA E. BECKER and J. A. FERRETTI, *J. Magn. Reson.* **13,** 27 (1974).
3. R. KAISER, *J. Magn. Reson.* **15,** 44 (1974).
4. A. G. REDFIELD and S. D. KUNZ, *J. Magn. Reson.* **19,** 114 (1975).
5. I. D. CAMPBELL, C. M. DOBSON, R. J. P. WILLIAMS and P. E. WRIGHT, *FEBS Lett.* **57,** 96 (1975).
6. D. I. HOULT, *Topics in Carbon-13 NMR,* **3,** 16 (1979).
7. E. D. BECKER, J. A. FERRETTI and P. N. GAMBHIR, *Analyt. Chem.* **51,** 1413 (1979).
8. J. W. COOPER, *The Minicomputer in the Laboratory: with examples using the PDP-11.* Wiley, London (1977).
9. D. L. DALRYMPLE, *Topics in Carbon-13 NMR,* **3,** 53 (1979).
10. G. BODENHAUSEN, R. FREEMAN, R. NIEDERMEYER and D. L. TURNER, *J. Magn. Reson.* **26,** 133 (1977).
11. D. I. HOULT, *Prog. in NMR Spectroscopy,* **12,** 41 (1978).
12. T. KELLER, U.S. Patent 3181650, Dec. 1973.
13. D. I. HOULT and R. E. RICHARDS, *Proc. Roy. Soc. (Lond),* **A344,** 311 (1975).
14. D. SHAW, *Fourier Transform NMR Spectroscopy,* Elsevier, Amsterdam (1976).
15. J. W. COOPER, *Computers and Chem.* **1,** 55 (1976).
16. R. R. ERNST, *J. Magn. Reson.* **4,** 280 (1971).
17. J.-P. MARCHAL, J. BRONDEAU and D. CANET, *J. Magn. Reson.* **33,** 471 (1979).

18. N. BLOEMBERGEN and R. V. POUND, *Phys. Rev.* **95,** 8 (1954).
19. A. G. REDFIELD and R. K. GUPTA, *Adv. Magn. Reson.* **5,** 81 (1971).
20. S. L. PATT and B. L. SYKES, *J. Chem. Phys.* **56,** 3182 (1972).
21. B. TOMLINSON and H. D. W. HILL, *J. Chem. Phys.* **59,** 2775 (1973).
22. A. G. MARSHALL, T. MARCUS and J. SALLOS, *J. Magn. Reson.* **35,** 227 (1979).
23. G. BODENHAUSEN, R. FREEMAN and G. A. MORRIS, *J. Magn. Reson.* **23,** 171 (1976).
24. J. W. COOPER, *J. Magn. Reson.* **22,** 345 (1976).
25. E. BARTHOLDI and R. R. ERNST, *J. Magn. Reson.* **11,** 9 (1973).
26. G. D. BERGLAND, *IEEE Spectrum,* **41,** (1969).
27. R. T. PAJER and I. M. ARMITAGE, *J. Magn. Reson.* **21,** 485 (1976).
28. P. GONORD, S. K. KAN and M. SAUZADE, *J. Magn. Reson.* **24,** 457 (1976).
29. R. R. ERNST, *Adv. Magn. Reson.* **2,** 1 (1966).
30. F. F. BROWN, I. D. CAMPBELL and P. W. KUCHEL, *FEBS Lett.* **82,** 12 (1977).
31. I. D. CAMPBELL, *Methods Biochem. Anal.* **25,** 1 (1979).
32. W. B. MONIZ, C. F. PORANSKI and S. A. SOJKA, *J. Magn. Reson.* **13,** 110 (1974).
33. I. D. CAMPBELL, C. M. DOBSON, R. J. P. WILLIAMS and A. V. XAVIER, *J. Magn. Reson.* **11,** 172 (1973).
34. I. D. CAMPBELL, S. LINDSKOG and A. I. WHITE, *J. Mol. Biol.* **90,** 469 (1974).
35. M. GASSNER, O. JARDETSKY and W. CONOVER, *J. Magn. Reson.* **30,** 141 (1978).
36. A. DE MARCO and K. WUTHRICH, *J. Magn. Reson.* **24,** 201 (1976).
37. M. GUERON, *J. Magn. Reson.* **30,** 515 (1978).
38. B. CLIN, J. DE BONY, P. LALANNE, J. BIAS and B. LEMENCEAU, *J. Magn. Reson.* **33,** 457 (1979).
39. J. W. AKITT, *J. Magn. Reson.* **32,** 311 (1978).
40. N. N. SEMENDYAEV, *Stud. Biophys.* **47,** 151 (1974).
41. A. G. FERRIGE and J. C. LINDON, *J. Magn. Reson.* **31,** 337 (1978).
42. J. C. LINDON and A. G. FERRIGE, *J. Magn. Reson.* **36,** 277 (1979).
43. E. L. HAHN and D. E. MAXWELL, *Phys. Rev.* **88,** 1070 (1952).
44. W. P. AUE, E. BARTHOLDI and R. R. ERNST, *J. Chem. Phys.* **64,** 2229 (1976).
45. J. W. COOLEY and J. W. TUKEY, *Math. Comput.* **19,** 297 (1965).
46. G. D. BERGLAND, *Math. Comput.* **22,** 275 (1968).
47. R. C. SINGLETON, *Comm. A.C.M.* **10,** 647 (1967).
48. G. D. BERGLAND, *Comm. A.C.M.* **11,** 705 (1968).
49. P. D. WELCH, *IEEE Trans. Audio. Electroacoust.* **AU17,** 151 (1969).
50. J. W. COOPER, *Transform Techniques in Chemistry,* Ed. P. R. GRIFFITHS, Plenum, New York (1978).
51. J. W. COOPER, I. S. MACKAY, and G. B. PAWLE, *J. Magn. Reson.* **28,** 405 (1977).
52. W. P. AUE, J. KARHAM and R. R. ERNST, *J. Chem. Phys.* **64,** 4226 (1976).
53. B. L. NEFF, J. L. ACKERMAN and J. S. WAUGH, *J. Magn. Reson.* **25,** 335 (1977).
54. A. G. MARSHALL and D. C. ROE, *Anal. Chem.* **50,** 756 (1978).
55. D. C. ROE, A. G. MARSHALL and S. H. SMALLCOMBE, *Anal. Chem.* **50,** 764 (1978).
56. A. G. MARSHALL and D. C. ROE, *J. Magn. Reson.* **33,** 551 (1979).
57. A. G. MARSHALL, *J. Phys. Chem.* **83,** 521 (1979).
58. R. KAISER, personal communication.
59. J. W. AKITT, *J. Magn. Reson.* **33,** 665 (1979).
60. R. R. ERNST, *J. Magn. Reson.* **1,** 7 (1969).
61. M. H. LEVITT and R. FREEMAN, *J. Magn. Reson.* **34,** 675 (1979).
62. G. A. PEARSON, *J. Magn. Reson.* **27,** 265 (1977).
63. The installation in the authors' laboratory.
64. A number of publications are clear and concise: those produced by component manufacturers are probably most pertinent. e.g. *Data Conversion Handbook,* D. S. BRUCK, Hybrid Systems Corp. (1974).
Analog-Digital Conversion Handbook, Ed. D. H. SHEINGOLD, Analog Devices Inc. This latter book contains a good bibliography.
65. P. BACHMANN, W. P. AVE, L. MULLER and R. R. ERNST, *J. Magn. Reson.* **28,** 29 (1977).