Chapter 3

Practical aspects of high-resolution NMR

NMR spectrometers are expensive instruments, representing one of the largest financial investments a chemical laboratory is likely to make, and to get the best results from these, they must be operated and maintained in the appropriate manner. This chapter explores some of the fundamental experimental aspects of relevance to high-resolution, solution-state NMR spectroscopy, from instrumental procedures through to the preparation of samples for analysis. Later sections also deal with the basics of calibrating a spectrometer and assessing its performance.

3.1. AN OVERVIEW OF THE NMR SPECTROMETER

A schematic illustration of a modern NMR spectrometer is presented in Fig. 3.1. The fundamental requirement for high-resolution NMR spectroscopy is an intense static magnetic field that is provided, nowadays exclusively, by superconducting solenoid magnets composed of niobium–tin alloy wire. These are able to produce the stable and persistent magnetic fields demanded by NMR spectroscopy of up to 22.3 T with current technology, corresponding to proton frequencies of 950 MHz. The drive for ever-increasing magnetic fields, encouraged by the demands for greater signal dispersion and instrument sensitivity, has continued since the potential of NMR spectroscopy as an analytical tool was first realised over 60 years ago (Fig. 3.2) and still represents an area of intense competition between magnet manufacturers. At the time of writing it appears that proton observation frequencies beyond the 1 GHz ceiling will demand the use of newer high-temperature superconducting wires to carry the huge currents required, which not only presents substantial technical challenges for the manufacturer but will likely further increase magnet costs for the end user. However, these very high-field magnets are currently of most interest to areas such as structural biology whilst in the area of organic chemistry the need for such extreme signal dispersion may be less important, although enhancements in sensitivity remain welcome and valuable developments. In recent years, there have been a number of advances in probe technologies that provide the chemist with very significant sensitivity increases without the need for higher field instrumentation, and these will also be reviewed later in this chapter.

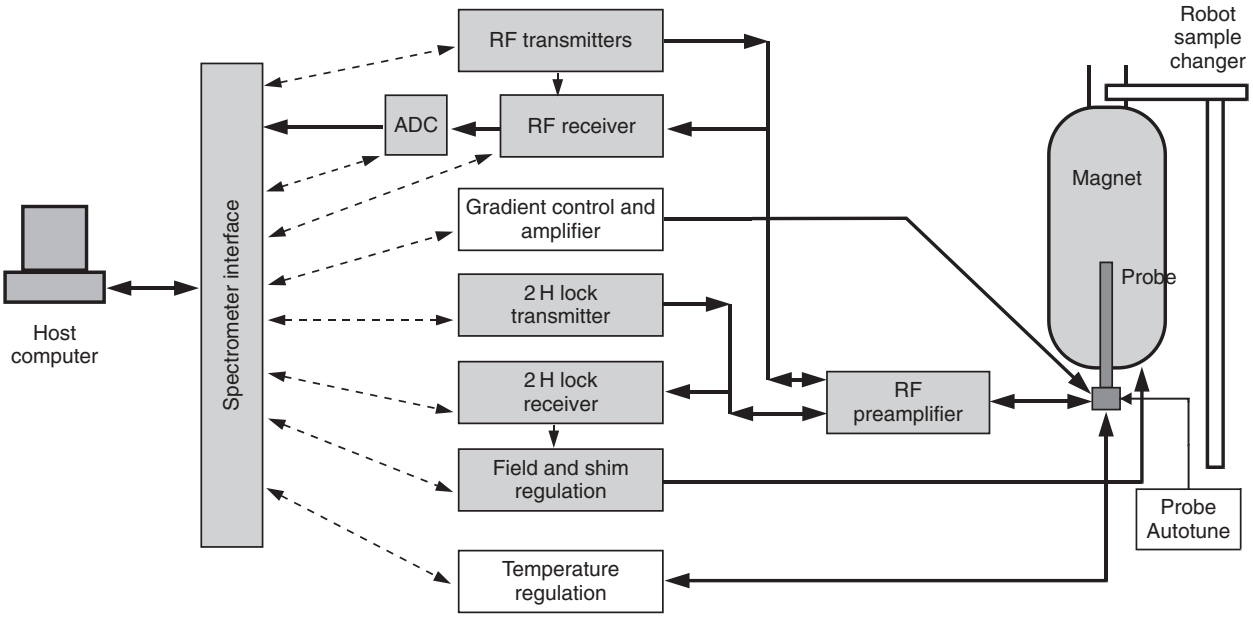


Figure 3.1. A schematic illustration of the modern NMR spectrometer. The greyed sections are essential components of the spectrometer, whereas the others may be considered optional features (although pulsed field gradients are essentially standard on research instruments).

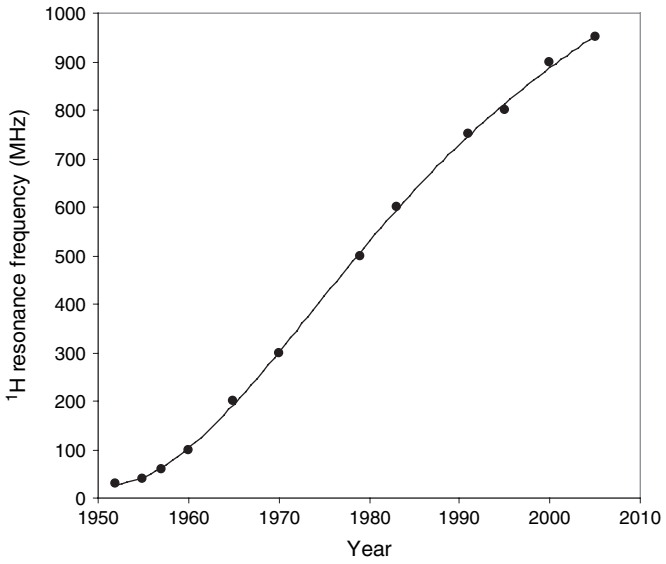


Figure 3.2. The increase in proton resonance frequency since the introduction of NMR spectroscopy as an analytical method (adapted with permission from reference [1] and extended Copyright 1995, Elsevier).

The magnet solenoid operates in a bath of liquid helium (at 4 K), surrounded by a radiation shield and cooled by a bath of liquid nitrogen (at 77 K), itself surrounded by a high vacuum. This whole assembly is an extremely efficient system, and once energised, the magnet operates for many years free from any external power source. It requires only periodic refilling of the liquid cryogens, typically on a weekly or bi-weekly basis for the nitrogen, but only every 2–12 months for helium depending on magnet age and construction. The central bore of the magnet dewar is itself at ambient temperature, and this houses a collection of electrical coils, known as the shim coils, which generate their own small magnetic fields and are used to trim the main static field and remove residual in homogeneities. This process of optimising the magnetic field homogeneity, known as *shimming*, is required for each sample to be analysed, and it is discussed in Section 3.4.5. Within the shim coils at the exact centre of the magnetic field sits the head of the probe, the heart of any NMR spectrometer. This houses the rf coil(s) and associated circuitry that act as antennae, transmitting and receiving the electromagnetic radiation. These coils may be further surrounded by pulsed field gradient coils that serve to *destroy* field homogeneity in a controlled fashion (this may seem a rather bizarre thing to do, but it turns out to have rather favourable effects in numerous experiments). The sample, sitting in a cylindrical glass tube, is held in a turbine or 'spinner' and descends into the probe head on a column of air or nitrogen. For routine 1D experiments, it is still common practice to spin the sample at 10–20 Hz whilst in the probe to average to zero field inhomogeneities in the transverse (x–y) plane and so improve signal resolution. Sample spinning is now rarely used for multidimensional NMR experiments on modern instruments, as it can induce additional signal modulations and associated undesirable artefacts. This requires that acceptable resolution can be obtained on a static sample, and while this is perfectly feasible with modern shim technology, older instruments may still demand spinning for all work.

Probes come in various sizes of diameter and length, depending on the magnet construction, but are more commonly referred to by the diameter of the sample tube they are designed to hold. The most widely used tube diameter is still 5 mm, other available sizes have included 1, 1.7, 2.5, 3, 8 and 10 mm (Section 3.4). Probes may be dedicated to observing one frequency (*selective* probes), may be tuneable over a very wide frequency range (*broadband* probes) or may tune to predefined frequency ranges, for example 4- or quad-nucleus probes. In all cases, they will also be capable of observing the deuterium frequency simultaneously to provide a signal for field regulation (the 'lock' signal). A second (outer) coil is often incorporated to allow the simultaneous application of pulses on one or more additional nuclei.

In many locations, it is advantageous to mount the whole of the magnet assembly on a vibration damping system as floor vibrations (which may arise from a whole host of sources including natural floor resonances, air conditioners, movement in the laboratory and so on) can have deleterious effects on spectra, notably around the base of resonances (Fig. 3.3). Whilst such artefacts have lesser significance to routine 1D observations, they may severely interfere with the detection of signals present at low levels, for example those in heteronuclear correlation or nuclear Overhauser effect experiments.

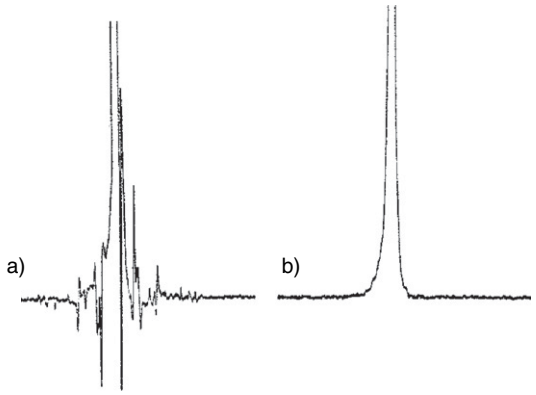


Figure 3.3. Floor vibrations can introduce unwelcome artefacts around the base of a resonance (a), which can be largely suppressed by mounting the magnet assembly on an anti-vibration stand (b) (figure courtesy of Bruker BioSpin).

Within the spectrometer cabinet sit the rf transmitters and the detection system for the observation channel, additional transmitter channels[1], the lock channel and the PFG transmitter. Most spectrometers come in either a two-channel or three-channel configuration, plus the lock channel. The spectrometer is controlled via the host computer (for which there has been a move towards Windows- or Linux-based PC systems in recent years), which is linked to the spectrometer via a suitable interface such as Ethernet. The electrical analogue NMR signals are converted to the digital format required by the host computer via the analogue-to-digital converter (ADC), the characteristics of which can have important implications for the acquisition of NMR data (Section 3.2). The computer also processes the acquired data, although this may also be performed 'off-line' with one of the many available NMR software packages.

Various optional peripherals may also be added to the instrument, such as variable temperature units that allow sample temperature regulation within the probe, robotic sample changers and so on. The coupling of NMR with other analytical techniques such as HPLC has become an established method through the development of flow probes and gained popularity in certain areas. The need for this will obviously depend on the type of samples handled and the nature of the experiments employed.

With the hype surrounding the competition between instrument manufacturers to produce ever-increasing magnetic fields, it is all too easy for one to become convinced that an instrument operating at the highest available field strength is essential in the modern laboratory. Whilst the study of biological macromolecules no doubt benefits from the greater sensitivity and dispersion available, problematic small or mid-sized molecules are often better tackled through the use of the appropriate modern techniques. Signal dispersion limitations are generally less severe and may often be overcome by using suitably chosen higher dimensional experiments. Sensitivity limitations, which are usually due to a lack of material rather than solubility or aggregation problems, may be tackled by utilising smaller probe geometries or cryogenically cooled probe heads (Section 3.4.2). So, for example, if one has insufficient material to collect a carbon-13 spectrum, one could consider employing a proton-detected heteronuclear correlation experiment to determine these shifts indirectly. Beyond such considerations, there are genuine physical reasons largely relating to the nature of nuclear spin relaxation, which mean that certain experiments on small molecules are likely to work less well at very high magnetic fields. In particular, this relates to the nuclear Overhauser effect (Chapter 8), one of the principal NMR methods in structure elucidation. For many cases commonly encountered in the chemical laboratory, a lower field instrument of modern specification is sufficient to enable the chemist to unleash an array of modern pulse NMR experiments on the samples of interest and subsequently solve the problem in hand. A better understanding of these modern NMR methods should aid in the selection of the appropriate experiments.

3.2. DATA ACQUISITION AND PROCESSING

This section examines some of the spectrometer procedures that relate to the collection, digitisation and computational manipulation of NMR data, including some of the fundamental parameters that define the way in which data are acquired. Such technicalities may

[1] Reference to the "decoupler channel(s)" is often used when referring to these additional channels, but this should not be taken too literally as they may only be used for the application of only a few pulses rather than a true decoupling sequence. This nomenclature stems from the early developments of NMR spectrometers when the additional channel was only capable of providing "noise decoupling", usually of protons.

not seem relevant to anyone who does not consider themselves a spectroscopist, but the importance of understanding a few basic relationships between experimental parameters comes from the need to recognise spectrum artefacts or corrupted data that can result from inappropriate parameter settings and to appreciate the limitations inherent in NMR measurements. Only then can one make full and appropriate use of the spectroscopic information at hand.

3.2.1. Pulse excitation

It is widely appreciated that modern NMR spectrometers use a 'short pulse' of rf energy to excite nuclear resonances over a range of frequencies. This pulse is supplied as *monochromatic* radiation from the transmitter; yet, the nuclear spin transitions giving rise to our spectra vary in energy according to their differing Larmor frequencies and so it would appear that the pulse will be unable to excite all resonances in the spectrum simultaneously. However, Heisenberg's Uncertainty Principle tells us that an excitation pulse of duration Δt has associated with it a frequency uncertainty or spread of around $1/\Delta t$ Hz and so effectively behaves as if it were polychromatic. The duration of the rf pulse Δt is usually referred to as the *pulse width*. A short, high-power pulse provides excitation over a wide frequency window whilst a longer, low-power pulse (which will provide the same net tip angle) is effective over a much smaller region (Fig. 3.4).

Consider a proton spectrum recorded at 400 MHz with the transmitter frequency placed in the centre of the spectral region of interest. A 10 ppm spectral window corresponds to 4000 Hz; therefore, we need to excite over $\pm$2000 Hz, meaning the pulse duration must be 0.5 ms or less. The observation frequency for carbon-13 on the same instrument would be 100 MHz, so a typical 200 ppm spectral width (SW) corresponds to $\pm$10 kHz, requiring a pulse of only 100 µs but with considerably higher power as its energy is now spread over a wider area. If the pulse were to be made very long (say tens of milliseconds) and weak, it would only excite over a rather small frequency range, giving rise to *selective excitation* of only part of the spectrum. As will become apparent in later chapters, the use of selective excitation methods is now commonplace in numerous NMR experiments. Long, low-power pulses that excite only a selected region of a spectrum are commonly referred to as 'soft pulses', whereas those that are of short duration and of high power are termed 'hard pulses'. Unless stated explicitly, all pulses in this book will refer to non-selective hard pulses.

Off-resonance effects

In practice, modern NMR instruments are designed to deliver high-power 90° pulses closer to 10 µs, rather than the hundreds predicted from the above arguments. This is to suppress the undesirable effects that arise when the pulse rf frequency is *off-resonance*, that is, when the transmitter frequency does not exactly match the nuclear Larmor frequency, a situation of considerable practical significance that has been ignored thus far.

As shown in Chapter 2, spins that are on-resonance have their magnetisation vector driven about the rf B_1 field towards the x–y plane during the pulse. Those spins that are off-resonance will, in addition to this B_1 field, experience a *residual* component ΔB of the *static* B_0 field along the z axis of the rotating frame for which:

$$\frac{\gamma \Delta B}{2\pi} = \gamma \Delta B = \Delta \nu \,\text{Hz} \tag{3.1}$$

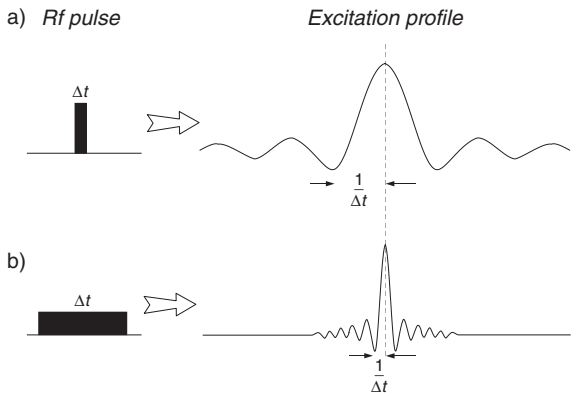


Figure 3.4. A single monochromatic radiofrequency pulse has an effective excitation bandwidth that depends inversely on the duration of the pulse. A short intense pulse is therefore able to excite over a wide frequency window (a), whereas a longer weaker pulse provides a more selective excitation profile (b).

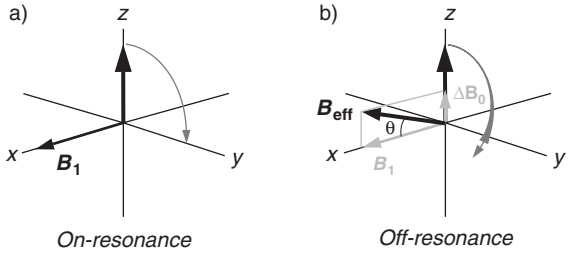


Figure 3.5. Excitation of magnetisation for which the rf is on-resonance (a) results in the rotation of the bulk vector about the applied rf field, B_1. Those spins which experience off-resonance excitation (b) are instead driven about an effective rf field, B_{eff}, which is tipped out of the x–y plane by an angle θ, which increases as the offset increases.

where $\Delta\nu$ represents the offset from the reference frequency. The vector sum of B_1 and ΔB is the *effective field* B_{eff} experienced by an off-resonance spin about which it rotates (or more correctly, nutates) (Fig. 3.5). This is greater than B_1 itself and is tilted away from the x–y plane by an angle θ, where

$$\tan\theta = \frac{\Delta B}{B_1} \tag{3.2}$$

For those spins further from resonance, the angle θ becomes greater and the net rotation towards the x–y plane diminishes until, in the limit, θ becomes 90°. In this case, the bulk magnetisation vector simply remains along the $+z$ axis and thus experiences no excitation at all. In other words, the nuclei resonate outside the *excitation bandwidth* of the pulse. Since an off-resonance vector is driven away from the y axis during the pulse it also acquires a (frequency-dependent) phase difference relative to the on-resonance vector (Fig. 3.6). This is usually small and an approximately linear function of frequency, so can be corrected by phase adjustment of the final spectrum (Section 3.2.8).

How deleterious off-resonance effects are in a pulse experiment depends to some extent on the pulse tip angle employed. A 90° excitation pulse ideally transfers magnetisation from the $+z$ axis into the x–y plane, but when off-resonance the tilt of the effective field will act to place the vector above this plane. However, the greater B_{eff} will mean the magnetisation vector follows a longer trajectory, and this increased net flip angle offers some compensation, and hence a 90° pulse is fairly tolerant to off-resonance effects as judged by the elimination of z-magnetisation (Fig. 3.7a). In contrast, a 180° inversion pulse ideally generates pure $-z$-magnetisation leaving none in the transverse plane, but now the increased effective flip angle is detrimental, tending to move the vector further from the South Pole. Thus, 180° pulses do not perform well when applied off-resonance (Fig. 3.7b), and can be a source of poor experimental performance. In practice, it is relatively easy to provide sufficiently short pulses (ca. 10 µs) to ensure that excitation and inversion is reasonably uniform over the relatively small frequency ranges encountered in ^{1}H NMR spectroscopy. However, for nuclei that display a much greater frequency dispersion, such as ^{13}C or ^{19}F, this is often not the case, and resonance distortion and/or attenuation can occur, and spurious signals may arise in multipulse experiments as a result. Experimental approaches to overcoming these limitations include the use of clusters of pulses, known as *composite pulses,* or of frequency-swept *adiabatic pulses*, both of which aim to compensate for these (and other) defects; see Chapter 10 for further discussion.

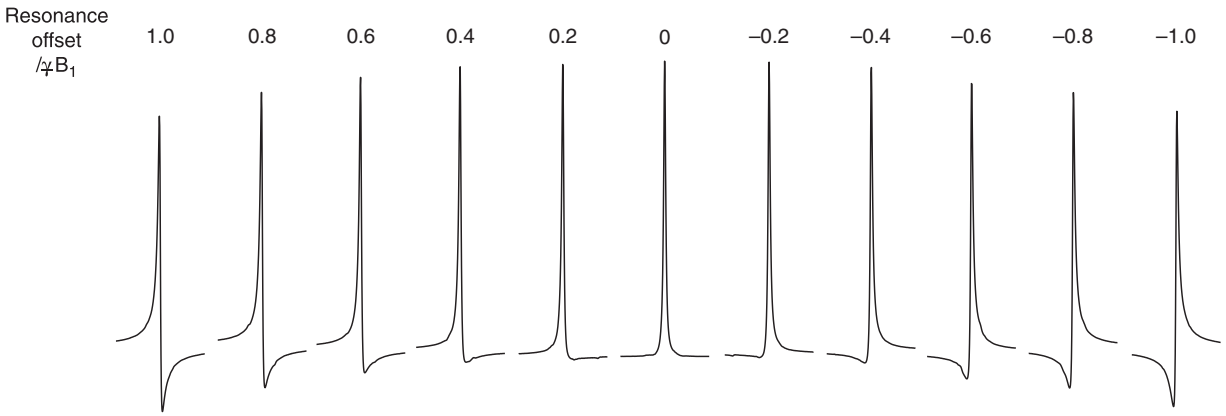


Figure 3.6. Experimental excitation profiles for a 90° pulse as a function of resonance offset relative to the applied rf field strength γB_1. Greater offsets introduce larger phase errors and reduce the amplitude of the resultant transverse magnetisation.

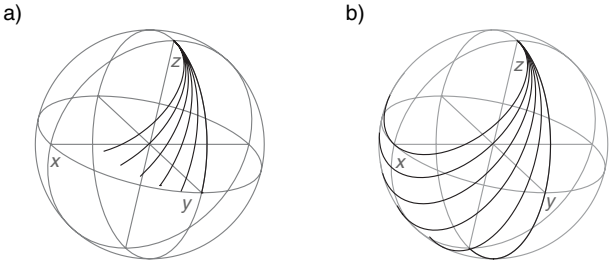


Figure 3.7. Excitation trajectories as a function of resonance offset for (a) a 90° pulse and (b) a 180°pulse. The offset moves from zero (on-resonance) to $+\gamma B_1$ Hz in steps of $0.2\,\gamma B_1$ (as in Fig. 3.6). The 90° pulse has a degree of "offset-compensation" as judged by its ability to generate transverse magnetisation over a wide frequency bandwidth. In contrast, the 180° pulse performs rather poorly away from resonance, leaving the vector far from the target South Pole and with a considerable transverse component.

3.2.2. Signal detection

Before proceeding to consider how one collects the weak NMR signals, we briefly consider the detection process that occurs within the NMR receiver. The energy emitted by the excited spins produces tiny analogue electrical signals that must be converted into a series of binary numbers to be handled by the computer. This *digitisation* process (Section 3.2.3) must occur for all the frequencies in the spectrum. As chemists we are really only interested in knowing the chemical shifts *differences* between nuclei rather than their *absolute* Larmor frequencies since it is from these differences that one infers differences in chemical environments. As we already know, these frequencies are typically tens or hundreds of megahertz whilst chemical shift ranges only cover some kilohertz. For example, a proton spectrum recorded at 200 MHz corresponds to a frequency range of only 2–3 kHz, and even for carbon (50 MHz on the same instrument), the frequency range is only around 10 kHz. Therefore, rather than digitising signals over many megahertz and retaining only the few kilohertz of interest, a more sensible approach is to subtract a reference frequency from the detected signal *before* digitisation, leaving only the frequency window of interest (Fig. 3.8). The resultant signals are now in the *audio* range, which, for humans at least, corresponds to frequencies less than 20 kHz (so it is also possible to hear NMR resonances).

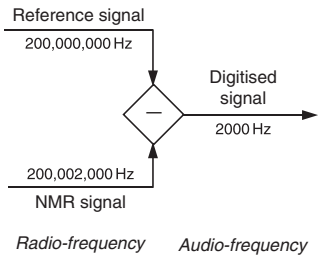


Figure 3.8. The NMR detection process. A fixed reference frequency is subtracted from the detected NMR signal so that only the frequency *differences* between resonances are digitised and recorded.

The reference frequency is usually chosen to be that of the original pulse used to excite the spin system and is supplied as a continuous reference signal from the transmitter (Fig. 3.9). This detection process is exactly analogous to the use of the rotating-frame representation introduced in Chapter 2 where the rotating frame reference frequency is also that of the rf pulse. If you felt a little uneasy about the seemingly unjustified use of the rotating frame formalism previously, then perhaps the realisation that there is a genuine experimental parallel within all spectrometers will help ease your concerns. The digitised FID you see therefore contains only the audio frequencies that remain after subtracting the reference, and it is these that produce the resonances observed in the final spectrum following Fourier transformation.

3.2.3. Sampling the FID

The Nyquist condition

To determine the frequency of an NMR signal correctly, it must be digitised at the appropriate rate. The *Sampling* or *Nyquist Theorem* tells us that to characterise a regular, oscillating signal correctly, it must be defined by at least two data points per wavelength. In other words, to characterise a signal of frequency F Hz, we must sample at a rate of at least $2F$; this is also known as the *Nyquist condition*. In NMR parlance, the highest recognised frequency is termed the *spectral width*, SW (Fig. 3.10). The time interval between sampled data points is referred to as the *dwell time*, DW, as given by:

$$\mathrm{DW} = \frac{1}{2\mathrm{SW}} \tag{3.3}$$

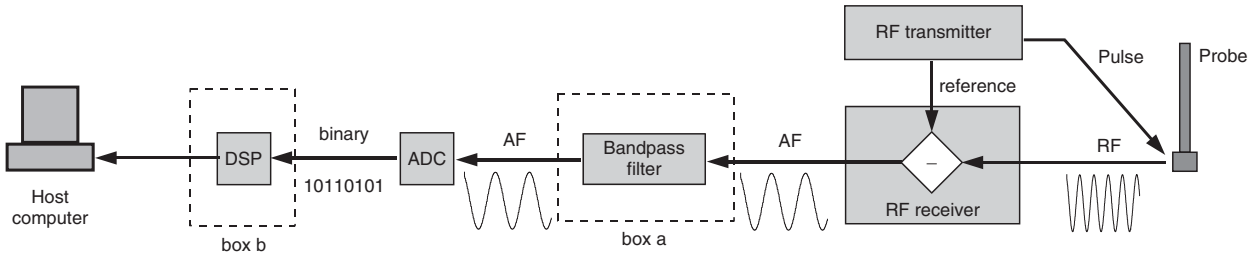


Figure 3.9. Schematic illustration of NMR data collection. Pulsed rf excitation stimulates the NMR response in the probe head, which is then amplified and detected in the receiver. The receiver reference frequency (that of the pulse rf) is subtracted to leave only audio frequencies (AF) that are digitised by the analogue-to-digital converter (ADC) and subsequently stored within the computer. The functions of the boxed sections are described later in the text (DSP: digital signal processor).

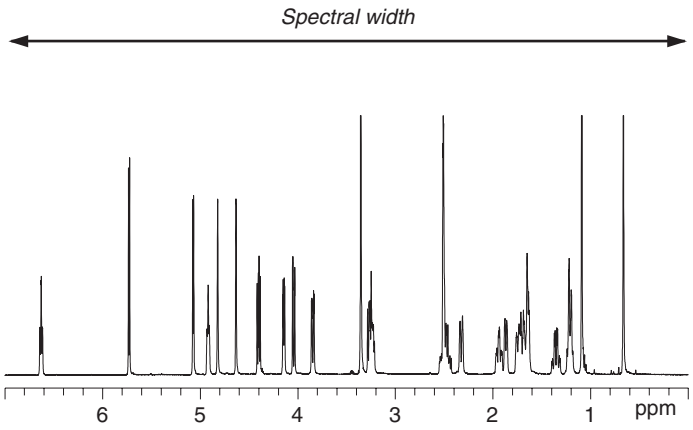


Figure 3.10. The spectral width defines the size of the observed frequency window. Only within this window are the line frequencies correctly characterised.

Signals with frequencies less than or equal to the spectral width will be characterised correctly as they will be sampled at two or more points per wavelength, whereas those with higher frequencies will be incorrectly determined and in fact will appear in the spectrum at frequencies that are lower than their true values. To understand why this occurs, consider the sampling process for three frequencies, two within and the third outside the spectral width (Fig. 3.11). The sampled data points of the highest frequency clearly match those of the lower of the three frequencies, and this signal will therefore appear with an incorrect frequency within the spectral window. Resonances that appear at incorrect frequencies, because in reality they exist outside the spectral width, are said to be *aliased* or *folded* back into the spectrum. Their location when corrupted in this manner is dependent on which of the two commonly employed *quadrature detection* schemes are in use, as explained in Section 3.2.4. On modern NMR spectrometers that employ digital signal detection (see Section 3.2.6), digital filters are able to completely eliminate signals that fall outside the defined spectral width and folded or aliased signals are therefore not observed. This filtering applies only for the directly sampled (observed) dimension of an NMR experiment, so aliasing of peaks may still be observed in the indirectly detected dimensions of a 2D experiment (see Section 5.3.1).

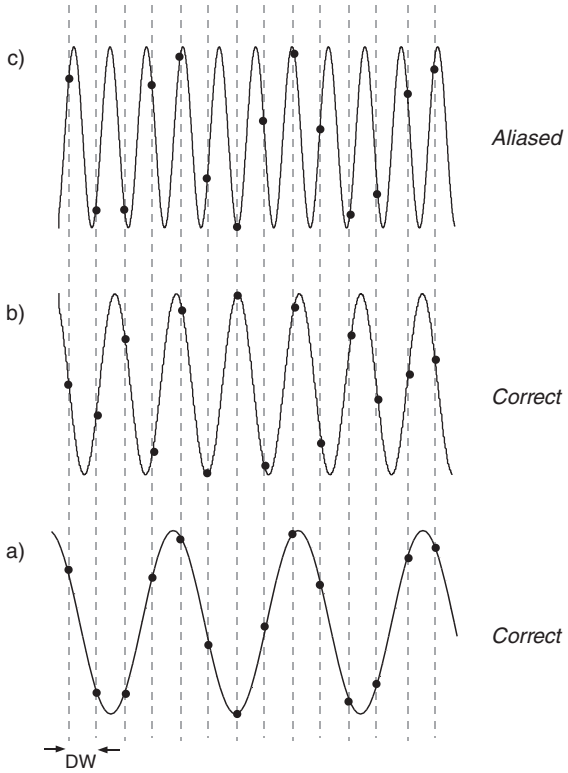


Figure 3.11. The Nyquist condition. To correctly characterise the frequency of a NMR signal, it must be sampled at least twice per wavelength, it is then said to fall within the spectral width. The sampling of signals (a) and (b) meet this criterion. Signals with frequencies too high to meet this condition are aliased back within the spectral width and so appear with the wrong frequency. The sampling pattern of signal (c) matches that of (a) and hence it is incorrectly recorded as having the lower frequency.

Filtering noise

One particularly insidious effect of aliasing or folding is that not only will NMR resonances fold into the spectral window but noise will as well. This can seriously compromise sensitivity, since noise can extend over an essentially infinite frequency range (so-called white noise) all of which, potentially, can fold into the spectrum and swamp the NMR resonances. It is therefore essential to prevent this by filtering out all signals above a certain frequency threshold with an audio *bandpass filter* after detection but prior to digitisation (Fig. 3.9, box a). The effect of the noise filter is demonstrated in Fig. 3.12, and the gain in signal-to-noise ratio in the spectrum employing this is clearly evident in Fig. 3.13. The cut-off bandwidth of the filter is variable, to permit the use of different spectral widths, and is usually automatically set by the spectrometer software to be a little greater than the spectral window. This is necessary because analogue filters are not perfect and do not provide an ideal sharp frequency cut-off, but instead tend to fall away steadily as a function of frequency, as is apparent in Fig. 3.12. A practical consequence is that these filters can attenuate the intensity of signals falling at the extreme ends of the spectrum, so may interfere with accurate signal intensity measurement and it is therefore wise to ensure that resonances do not fall at the edges. Analogue filters can also introduce a variety of distortions to the spectrum, and the use of digital filtration methods having superior characteristics is now a standard feature on modern NMR spectrometers. Digital signal processing is considered in Section 3.2.6.

Acquisition times and digital resolution

The total sampling period of the FID, known as the *acquisition time*, is ultimately dictated by the frequency resolution required in the final spectrum. From the Heisenberg's Uncertainty Principle, the resolution of two lines separated by $\Delta\nu$ Hz requires data collection for at least $1/\Delta\nu$ s. If one samples the FID for a time that is too short, then the frequency differences cannot be resolved and fine structure is lost (the minimum linewidth that can be resolved is given approximately by 0.6/AQ). Likewise, if the signal decays rapidly then one is unable to sample it for a long period of time and again can resolve no fine structure. In

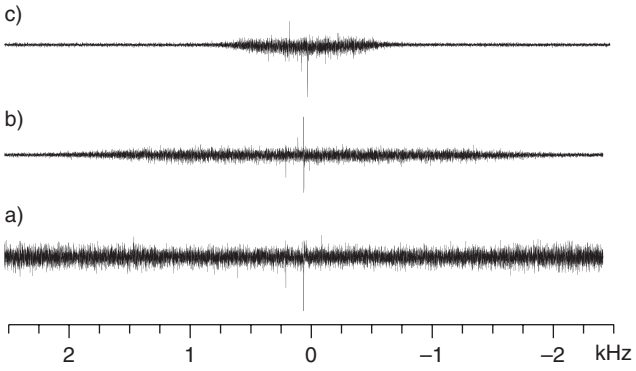


Figure 3.12. Eliminating noise with analogue filters. Spectra were recorded over a spectral width of 5 kHz with no sample in the probe. The analogue filter window was set to (a) 6 kHz, (b) 3 kHz and (c) 1 kHz and the associated noise attenuation is apparent, although the cut-off point is ill defined.

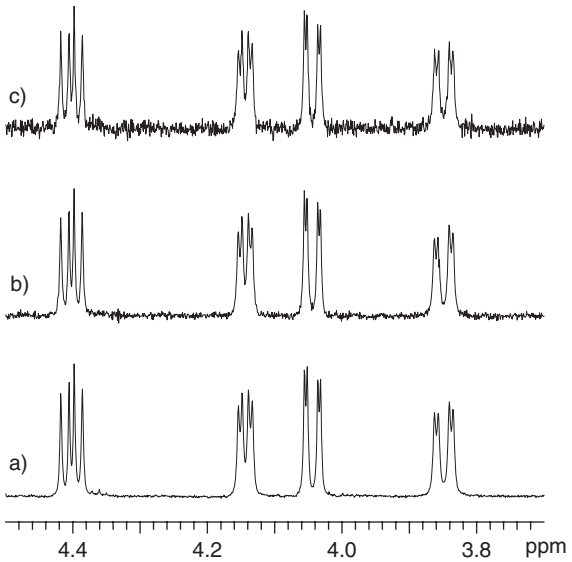


Figure 3.13. The effect of analogue filters on signal-to-noise. Spectrum (a) was recorded with the correct filter width (1.25 times the spectral width), and (b) and (c) with it increased 10- and 100-fold respectively.

this case, the rapid decay implies a large linewidth arising from natural (transverse) relaxation processes, or from poor field homogeneity, and one will be unable to recover resolution by extending the acquisition time. Thus, when selecting the appropriate acquisition time one needs to consider the likely frequency differences that need to be resolved and the relaxation times of the observed spins.

The acquisition time is defined by the digitisation rate, which is dictated by the spectral width and defines the sampling dwell time DW, and on how many data points are sampled in total. If the FID contains TD time-domain data points then:

$$\mathrm{AQ} = \mathrm{DW} \cdot \mathrm{TD} = \frac{\mathrm{TD}}{2 \cdot \mathrm{SW}} \tag{3.4}$$

When dealing with the final spectrum, we are concerned with frequency and not time, and what one really needs is some measure of how well the resonances in the spectrum are digitised. The figure of interest is the frequency between adjacent data points in the spectrum, the *digital resolution*, DR (quoted in hertz per data point or simply Hz/point). It should be stressed that this is not what spectroscopists speak of when they refer to the 'resolution' achieved on a spectrometer as this relates to the homogeneity of the magnetic field. Digital resolution relates only to the frequency window to data point ratio, which is small for a well-digitised spectrum but large when poorly digitised.

Following the FT, two data sets are generated representing the 'real' and 'imaginary' spectra (Section 2.3), so the real part with which one usually deals contains half the data points of the original FID (in the absence of further manipulation), and its data size, SI, is therefore TD/2. DR is then:

$$\mathrm{DR} = \frac{\text{total frequency window}}{\text{total number of data points}} = \frac{\mathrm{SW}}{\mathrm{SI}} = \frac{2 \cdot \mathrm{SW}}{\mathrm{TD}} = \frac{1}{\mathrm{AQ}} \tag{3.5}$$

Thus, digital resolution is simply the reciprocal of the acquisition time, so to collect a well-digitised spectrum one must sample the data for a long period of time; clearly this is the same argument presented above.

The effect of inappropriate digital resolution is demonstrated in Fig. 3.14. Clearly with a high value, that is short AQ, the fine structure cannot be resolved. Only with the extended acquisition times can the genuine spectrum be recognised. For proton spectroscopy, one needs to resolve frequency differences of somewhat less than 1 Hz to be able to recognise small couplings, so acquisition times of around 2–4 s are routinely used, corresponding to DRs of around 0.5 to 0.25 Hz/point. This then limits the accuracy with which frequency measurements can be made, including shifts and couplings, (even though peak listings tend to quote resonance frequencies to many decimal places).

The situation is somewhat different for the study of nuclei other than protons however, since they often exhibit rather few couplings (especially in the presence of proton decoupling), and because one is usually more concerned with optimising sensitivity, so does not wish to sample the FID for extended periods and thus sample more noise. Since one does not need to define lineshapes with high accuracy, lower digital resolution suffices to resolve chemical shift differences, and 1–2 Hz/point (0.5–1 s AQ) is adequate in ^{13}C NMR. Exceptions occur in the case of nuclei with high natural abundance that may show homonuclear couplings such as ^{31}P, or when the spectrum is being used to estimate the relative ratios of compounds in solution. This is true, for example, in the use of 'Mosher's acid' derivatives for the determination of enantiomeric excess from ^{19}F NMR spectra, for which the resonances must be well digitised to represent the true intensity of each species. In such cases, the parameters that may be used for the 'routine' observation of ^{19}F are unlikely to be suitable for quantitative measurements. Similarly, the use of limited digital resolution in routine carbon-13 spectra is one reason why signal intensities do not provide a reliable indication of relative concentrations. In contrast, the spectra of rapidly relaxing quadrupolar nuclei contain broad resonances that require only low digital resolution and hence short acquisitions times.

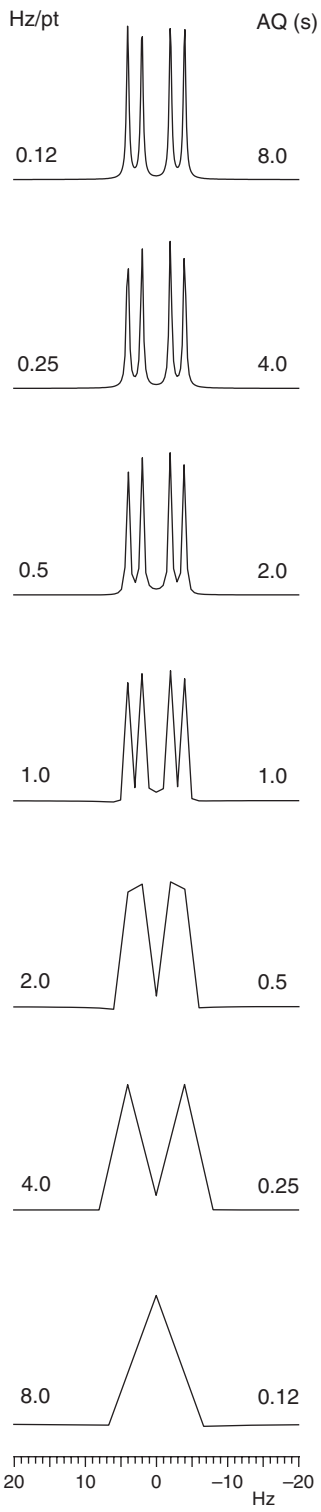


Figure 3.14. The influence of data acquisition times on the ability to resolve fine structure. Longer acquisition times correspond to higher digitisation levels (smaller Hz/pt), which here enable characterisation of the coupling structure within the double-doublet ($J = 6$ and 2 Hz).

Zero-filling and truncation artefacts

Whilst the NMR response decays throughout the FID, the noise component remains essentially constant and will eventually dominate the tail of the FID. At this point, there is little advantage in continuing acquisition since this only adds noise to the final spectrum. Provided the FID has fallen to zero when acquisition stops, one can artificially improve the digital resolution by appending zeros to the end of the FID. This process is known as *zero-filling*, and it interpolates these added data points in the frequency domain and so enhances the definition of resonance lineshapes (Fig. 3.15).

It has been shown [2] that by doubling the number of data points in the time domain by appending zeros (a single 'zero-fill'), it is possible to improve the frequency resolution in the spectrum. The reason for this gain stems from the fact that information in the FID is split

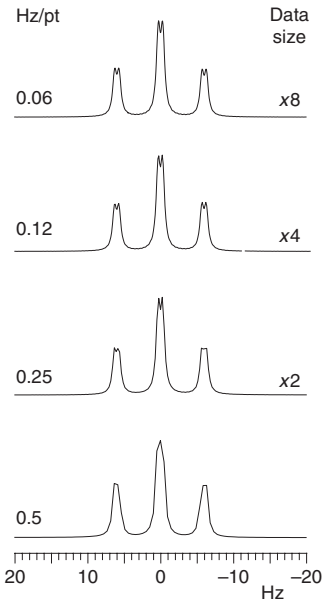


Figure 3.15. Zero-filling can be used to enhance fine structure and improve lineshape definition.

into two parts (real and imaginary) after the transformation, so one effectively looses information when considering only the real (absorption) part. Doubling the data size regains this lost information and is therefore a useful tool in the routine analysis of proton spectra. (One could of course simply double the acquired data points, but this leads to reduced sensitivity and requires more spectrometer time.) Further zero-filling simply increases the digital resolution of the spectrum by interpolating data points, so no new information can be gained and the improvement is purely cosmetic. However, because this leads to a better definition of each line, it can still be extremely useful when analysing multiplet fine structure in detail or when measuring accurate resonance intensities; in the determination of enantiomeric excess via chiral solvating reagents for example. In 2D experiments, acquisition times are necessarily kept rather short, and zero-filling is routinely applied to improve the appearance of the final spectrum. A more sophisticated approach to extending the time-domain signal known as *linear prediction* is described below.

If the FID has not decayed to zero at the end of the acquisition time, the data set is said to be *truncated* and this leads to distortions in the spectrum after zero-filling and FT. The distortions arise from the FT of the sudden step within the FID, the result of which is described by the function $(\sin x)/x$, also known as sinc x. Fig. 3.16 shows that this produces undesirable ringing patterns that are symmetrical about the base of the resonance, often referred to as 'sinc wiggles'. To avoid this problem, it is essential to ensure the acquisition time is sufficiently long, to force the FID to decay smoothly to zero with a suitable shaping function (Section 3.2.7) or to artificially extend the FID by LP.

In proton spectroscopy, acquisition times are sufficiently long that truncation artefacts are rarely seen in routine spectra, although they may be apparent around the resonances of small molecules with long relaxation times, solvent lines for example. As stated above, acquisition times for other nuclei are typically kept short and FIDs are usually truncated. This, however, is rarely a problem as it is routine practice to apply a window function to enhance sensitivity of such spectra, which itself also forces the FID to zero so eliminating the truncation effects. Try processing a standard ^{13}C spectrum directly with an FT but without the use of a window function; for some resonances, you are likely to observe negative responses in the spectra that appear to be phasing errors. In fact these distortions arise purely from the low digital resolution used, that is, the short acquisition times, and the corresponding truncation of the FID [3] (Fig. 3.17a). Use of a simple decaying function (Section 3.2.7) removes these effects (Fig. 3.17b) and when used in this way is often referred to as *apodisation*, literally meaning 'removing the feet'. Similar considerations apply to truncated 2D data sets where shaping functions play an essential role.

Linear Prediction

The method of linear prediction (LP) can play many roles in processing of NMR data, [4, 5] from the rectification of corrupted or distorted data, through to the complete generation of frequency-domain data from an FID; an alternative to the FT. Here we consider its most popular usage, known as forward linear prediction, which extends a truncated FID. Rather than simply appending zeros, this method, as the name suggests, predicts the values

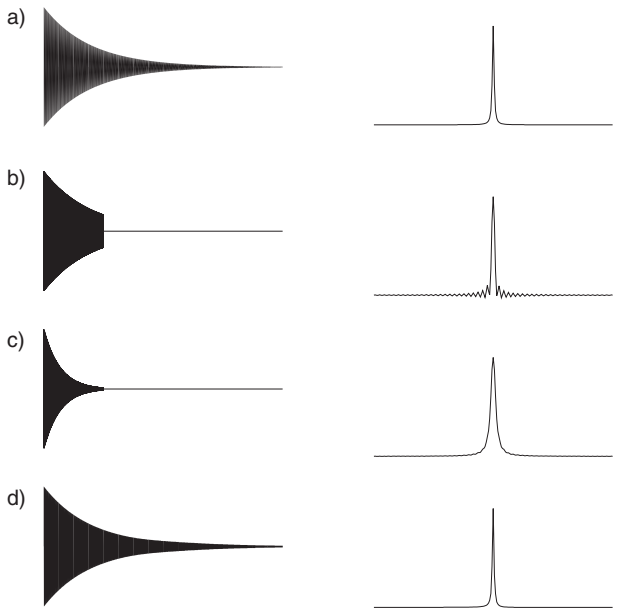


Figure 3.16. Processing a truncated FID. (a) A complete FID and the corresponding resonance, (b) a truncated FID that has been extended by zero-filling produces sinc wiggles in the spectrum, (c) apodisation of the FID in (b) together with zero filling and (d) linear prediction of the FID in (b).

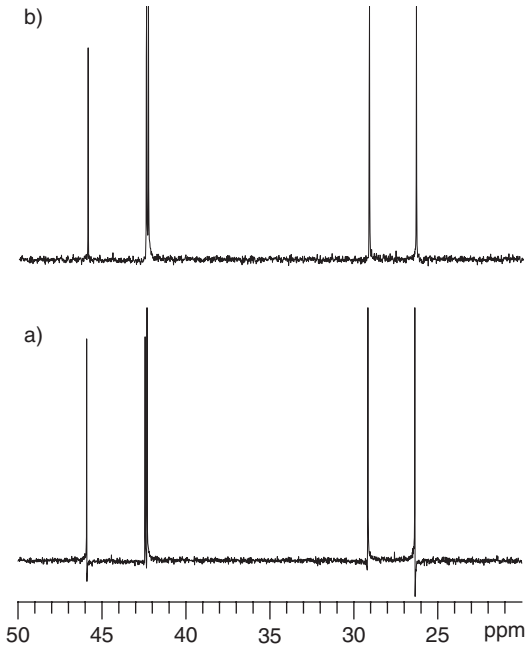


Figure 3.17. Carbon spectra often display distortions when transformed directly (a) which appear to be phase errors but which actually arise from a short acquisition time. Applying a line broadening apodisation function prior to the transform removes these distortions (b, 1 Hz line broadening).

of the missing data points by using the information content of the previous points and so genuinely extends the FID (Fig. 3.16d).

In a time sequence of data points, the value of a single point, d_n, can be estimated from a linear combination of the immediately preceding values:

$$d_n = a_1 d_{n-1} + a_2 d_{n-2} + a_3 d_{n-3} + a_4 d_{n-4} \ldots \tag{3.6}$$

where a_1, a_2, ... and so on represent the LP coefficients. The number of coefficients (referred to as the *order* of the prediction) corresponds to the number of data points used to predict the next value in the series. Provided the coefficients can be determined from the known data, it is then possible to extrapolate beyond the acquired data points. Repetition of this process incorporating the newly predicted points ultimately leads to the extended FID.

This process is clearly superior to zero-filling and produces a much better approximation to the true data than does simply appending zeros. It improves resolution, avoids the need for strong apodisation functions and greatly attenuates truncation errors. Naturally, the method has its limitations, the most severe being the requirement for high signal-to-noise in the FID for accurate estimation of the LP coefficients. Successful execution also requires that the number of points used for the prediction is very much greater than the number of lines that comprise the FID. This may be a problem for 1D data sets with many component signals, and LP of such data is less widely used. The method is far more valuable for the extension of truncated data sets in the indirectly detected dimensions of a 2D or 3D experiment (you may wish to return to these discussions when you are familiar with the 2D approach, see Chapter 5). Here, individual interferograms contain rather few lines (look at a column from a 2D data set) and are thus better suited to prediction, providing sensitivity is adequate. The use of LP in the routine collection of 2D spectra of organic molecules has been subject to detailed investigation [6]. Typically, a twofold to fourfold prediction is used, so for example, 128 t_1 data points may be readily extended to 256 or 512. Apodisation of the data is still likely to be required, although not as severely as for untreated data, and zero-filling may also be applied to further improve digitisation. Figure 3.18 clearly demonstrates the improved resolution attainable through the use of LP in the indirect dimension of

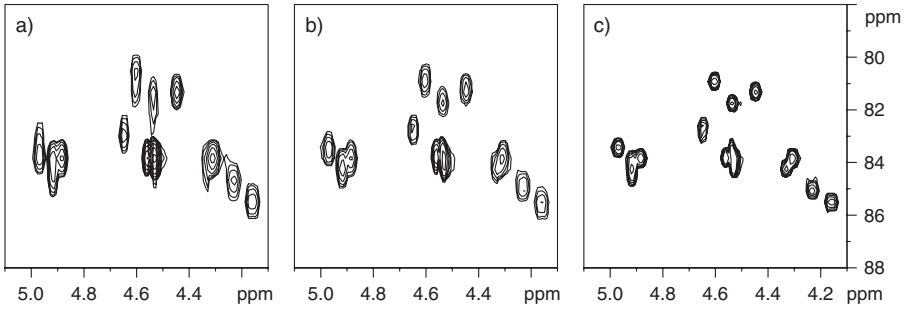


Figure 3.18. Improved peak resolution in a two-dimensional heteronuclear proton-carbon correlation experiment (Chapter 6) through linear prediction (LP). The same raw data was used in each spectrum, with the F_1 (carbon) dimension processed with (a) no data extension, (b) one zero-fill and (c) LP in place of zero-filling.

a 2D heteronuclear correlation experiment. The same principles can be used in *backward* LP. Data points at the start of an FID are sometimes corrupted by ringing in receiver circuitry when detection starts resulting in baseline distortion of the spectrum. Replacing these points with uncorrupted predicted points eliminates the distortion; an example of this is found in Section 4.5.

3.2.4. Quadrature detection

It was described above how during the classical NMR signal detection process a reference frequency equal to that of the excitation pulse is subtracted from the NMR signal to produce an audio frequency signal that is digitised and later subject to FT. The problem with analysing the data produced by this *single-channel detection* is that the FT is intrinsically unable to distinguish frequencies above the reference from those below it, that is, it cannot differentiate positive from negative frequencies. This results in a magnetisation vector moving at $+\nu$ Hz in the rotating frame producing two resonances in the spectrum at $+\nu$ and $-\nu$ Hz after FT (Fig. 3.19). The inevitable confusing overlap in a spectrum acquired with the reference positioned in the centre of the spectral width can be avoided by placing the reference at one edge of the spectrum to ensure all rotating-frame frequencies have the same sign. Whilst this will solve potential resonance overlap problems, it introduces a number of other undesirable factors. Firstly, although there will be no mirror image NMR resonances remaining within the spectral window, noise will still be mirrored about the reference frequency and added to that already present. This leads to a decrease in signal-to-noise by a factor of $\sqrt{2}$, or about 1.4 (not by a factor of 2 because noise is random and does not add coherently). Furthermore, this produces the greatest possible frequency separation between the pulse and the highest frequency resonance, so enhancing undesirable off-resonance effects. The favoured position for the reference is therefore in the centre of the spectrum, and two-channel (*quadrature*) detection is then required to distinguish sign.

To help visualise why single-channel detection cannot discriminate positive and negative frequencies, and how the quadrature method can, consider again a single magnetisation vector in the rotating frame. The use of the single-channel detector equates to being able to observe the precessing magnetisation along only one axis, say the y axis. Figure 3.20 shows that the resultant signal along this axis for a vector moving at $+\nu$ Hz is identical to that moving at $-\nu$ Hz, both giving rise to a cosinusoidal signal, so the two are indistinguishable (Fig. 3.20a). If, however, one were able to make use of two (phase-sensitive) detection channels whose reference signals differ in phase by 90° (hence the term *quadrature*), one would then be able to observe magnetisation simultaneously along both the x and the y axes such that one channel monitors the cosine signal and the other the sine (Fig. 3.20a and b).

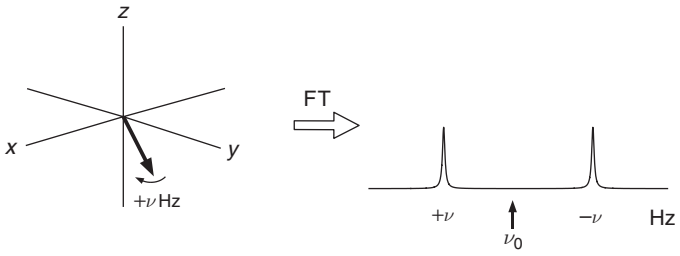


Figure 3.19. A single-channel detection scheme is unable to differentiate positive and negative frequencies in the rotating frame. This results in a mirror image spectrum being superimposed on the true one if the transmitter is placed in the centre of the spectrum.

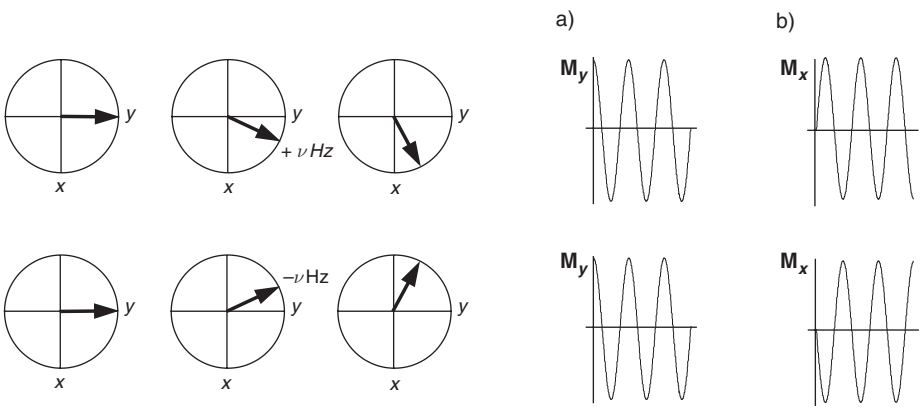


Figure 3.20. A two-channel quadrature detection system monitors magnetisation on two orthogonal axes, providing both (a) cosine- and (b) sine-modulated data, which ultimately allow the sense of precession to be determined (see text).

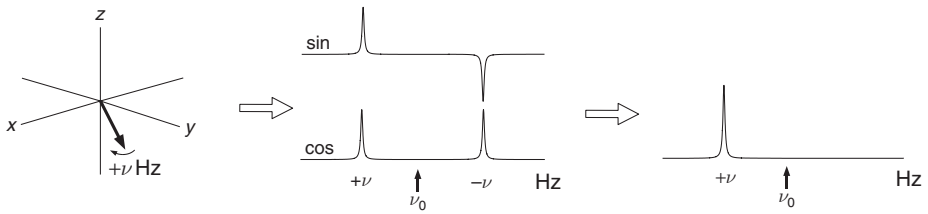


Figure 3.21. A two-channel detection scheme is able to differentiate positive and negative frequencies in the rotating frame, allowing the transmitter to be placed in the centre of the spectrum without the appearance of mirror image signals.

With the additional information provided by the sinusoidal response of the second channel, the sense of rotation can be determined, and vectors moving at $\pm\nu$ Hz can be distinguished (Fig. 3.21). Technically, the FT is then *complex*, with the x and y components being handled separately as the real and the imaginary inputs to the transform, following which the positive and negative frequencies are correctly determined. In the case of the single channel, the data are used as input to a *real* FT.

Simultaneous and sequential sampling

Quadrature detection is universally employed in all modern spectrometers. However, there exist a number of experimental schemes for implementing this, and which of these you are likely to use will be dictated by the spectrometer hardware and perhaps by the age of the instrument (on some modern instruments, the operator can choose between these methods). The differing approaches may be divided into the classical *analogue* scheme described here for which there are two widely used methods, and the more recent *digital quadrature detection* (DQD) scheme that is considered later.

In both analogue implementations, the incoming rf signal from the probe and preamplifier is split into two signals and each fed to separate phase-sensitive detectors whose reference frequencies are identical but differ in phase by 90° (Fig. 3.22). The resulting audio signals are then sampled, digitised and stored for subsequent analysis, and it is in the execution of these that the two methods differ. The first relies on *simultaneous* sampling of the two channels (i.e. the channels are sampled at precisely the same point in time), and the data points for each are stored in separate memory regions (Fig. 3.23a). The two sets of data so generated represent the cosine and sine components required for the sign discrimination and are used as the real and the imaginary inputs to a complex FT routine. As this enables positioning of the transmitter frequency in the centre of the spectrum, a frequency range of only $\pm$SW/2 need be digitised. According to the Nyquist criterion, the sampling rate now becomes equal to SW, that is, $(2 \times \text{SW}/2)$ or half of that required for single-channel detection, so the equations given in the previous section are modified slightly for *simultaneous quadrature detection* to give:

$$\text{DW} = \frac{1}{\text{SW}} \quad \text{and} \quad \text{AQ} = \text{DW} \cdot \frac{\text{TD}}{2} = \frac{\text{TD}}{2 \cdot \text{SW}} \tag{3.7}$$

also

$$\text{DR} = \frac{\text{SW}}{\text{SI}} = \frac{2 \cdot \text{SW}}{\text{TD}} \quad \text{and so again} \quad \text{DR} = \frac{1}{\text{AQ}} \tag{3.8}$$

Notice that in the calculation of the acquisition time, we need only consider *half* the total time-domain data points since now two points are sampled at the same time.

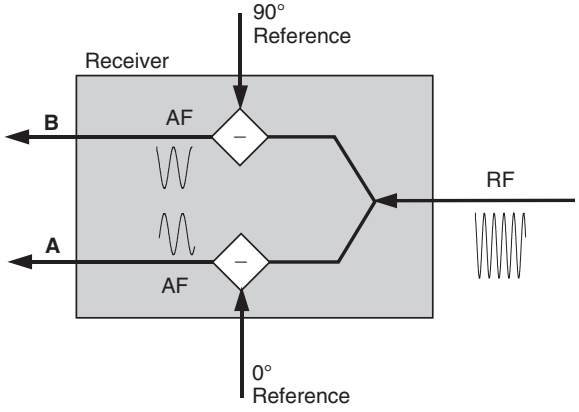


Figure 3.22. Schematic illustration of the experimental implementation of analogue quadrature detection in the NMR receiver. The incoming rf signal is split into two, and the reference signal, differing in phase by 90° in the two channels, subtracted. Channels A and B therefore provide the required sine and cosine components.

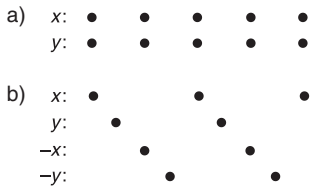


Figure 3.23. Data sampling schemes for the two common quadrature detection methods. (a) Simultaneous sampling: the two quadrature channels (representing *x*- and *y*-magnetisation) are sampled at the same point in time. (b) Sequential sampling: the two channels are sampled alternately at twice the rate of method (a), and the phase inverted for alternate pairs of data points (see text).

The second method for quadrature detection actually strives to eliminate the problems of positive and negative frequencies by effectively mimicking a single-channel detection scheme with the transmitter at the edge of the spectrum, so is sometimes referred to as 'pseudo-quadrature detection'. Although it allows the positioning of the transmitter frequency at the centre of the spectrum, it employs an ingenious sampling scheme to make it appear as if the reference frequency sat at one edge of the spectrum, so that only frequencies of one sign are ever characterised. In this method, data points are sampled *sequentially* at a rate suitable for single-channel detection (i.e. 2·SW or twice that in the simultaneous method), but for each data point sampled, the reference phase is incremented by 90° (Fig. 3.23b). The effect on a magnetisation vector is that it appears to have advanced in the rotating frame by 90° more between each sampling period than it actually has, so the signal appears to be precessing faster than it really is. As digitisation occurs at a rate of 2·SW and each 90° phase shift corresponds to an advance of 1/4 of a cycle, the frequency appears to have increased by 2·SW/4 or SW/2. Since the transmitter is positioned at the centre of the spectrum, the genuine frequency range runs from +SW/2 to −SW/2, meaning the artificial increase of +SW/2 moves the frequency window to +SW to 0. Thus, there are no longer negative frequencies to distinguish. Experimentally, the required 90° phase shifts are achieved by alternating between the two phase-sensitive detectors to give the 0° and 90° shifts, and simply inverting the signals from these channels to provide the 180° and 270° shifts respectively (Fig. 3.23b) The sampled data are handled by a single memory region as *real* numbers that are used as the input to a *real* FT. In this *sequential quadrature detection* scheme, the sampling considerations, given in Section 3.2.3, apply as they would for single-channel detection. In Chapter 5, these ideas are extended to quadrature detection in 2D spectroscopy.

For either scheme, the total number of data points digitised, acquisition times and spectral widths are identical, so the resulting spectra are largely equivalent. The most obvious difference is in the appearance of aliased signals, that is, those that violate the Nyquist condition, as described below. Experimentally, flatter baselines are observed for the simultaneous method as a result of the *symmetrical* sampling of the initial data points in the FID, and this is the recommended protocol for an analogue detection scheme.

Aliased signals

In Section 3.2.3 it was shown that a resonance falling outside the spectral window (because it violates the Nyquist condition) will still be detected but will appear at an incorrect frequency and is said to be *aliased* or *folded* back into the spectrum (if digital signal filters are not employed to eliminate this). This can be confusing if one is unable to tell whether the resonance exhibits the correct chemical shift or not. The precise location of the aliased signal in the spectrum depends on the quadrature detection scheme in use and on how far outside the window it truly resonates. With the simultaneous (complex FT) scheme, signals appear to be 'wrapped around' the spectral window and appear at the *opposite* end of the spectrum (Fig. 3.24b), whereas with the sequential (real FT) scheme, signals are 'folded back' at the *same* end of the spectrum (Fig. 3.24c). If you are interested to know why this difference occurs, see reference [7].

Fortunately, in proton spectroscopy it is generally possible to detect the presence of an aliased peak because of its esoteric phase that remains distorted when all others are correct. In heteronuclear spectra that display only a single resonance, such phase characteristics cannot provide this information as there is no other signal with which to make the comparison. In such cases, it is necessary to widen the spectral window and record any movement of the peak. A signal that is not aliased will be unchanged and will appear with the same chemical shift. In contrast, an aliased resonance will move towards the edge of the spectrum and will appear at a different shift. The situation is even worse if digital filters are employed as any resonance that falls outside the defined spectral window is simply not observed. If you happen to be searching for a resonance of unknown chemical shift, it would be wise to turn off digital filtration to increase the chance of detecting the peak.

Quadrature images

In addition to the corruption of spectra due to signal aliasing, there may also exist unwelcome artefacts that arise from imperfections in the analogue quadrature detection schemes. In these, one requires two signals to be digitised for subsequent storage in two separate memory blocks that differ in phase by 90° but which are otherwise identical. However, experimentally it is rather difficult to ensure the phase difference is *exactly* 90° and that the signal amplitudes are identical in each channel. Such channel imbalance in phase and amplitude leads to spurious images of resonances mirrored about the transmitter

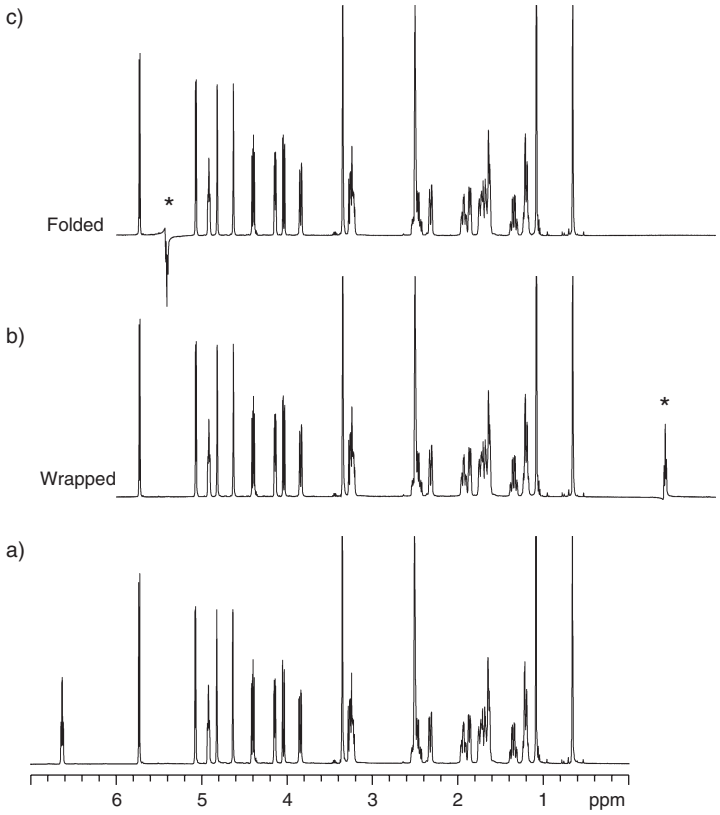


Figure 3.24. Aliasing of resonances. Spectrum (a) displays all resonances at their correct shifts, whilst (b) and (c) result from the spectral window being positioned incorrectly. Spectrum (b) shows how the resonance wraps back into the spectrum at the far end when simultaneous sampling is employed, whereas in (c) it folds back in at the near end with sequential sampling. Typically, the phase of the aliased resonance(s) is also corrupted.

frequency known as *quadrature images* or colloquially *quad images* (Fig. 3.25). Their origin can be realised if one imagines the signal amplitude in one channel to fall to zero. The scheme is then one of single-channel detection, so again positive and negative frequencies cannot be distinguished and thus mirror images appear about the reference (see Fig. 3.19 again). In a correctly balanced receiver, these images are usually less than 1% on a single scan, but even these can be significant if you wish to detect weak resonances in the presence of very strong ones, so some means of compensating the imbalance is required. The traditional approach to this is through *phase cycling* of the rf pulses and the receiver to cancel these artefacts, and the elimination of these images in this way serves as an introduction to the concept of phase cycling and is described in the section that follows. On modern instruments, these images are avoided through the use of digital quadrature detection in place of the analogue schemes described above.

Digital quadrature detection

In the traditional analogue detection scheme illustrated in Fig. 3.22, the two quadrature signals required for frequency discrimination were generated by mixing with analogue reference signals of (ideally) 90° phase difference prior to independently digitising the data from the two channels. In the digital quadrature detection (DQD) scheme, only a single channel is digitised, representing, say, the 0° component. This digital data may then be

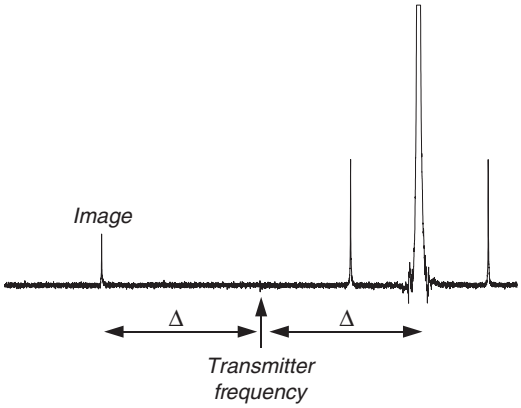


Figure 3.25. Quadrature images are unwanted mirror image artefacts that arise from spectrometer imperfections. Here an image of $CHCl_3$ can be seen at about one-half the height of the carbon satellites in a single-scan spectrum. Phase cycling is typically employed to suppress these, although digital quadrature detection can eliminate them completely.

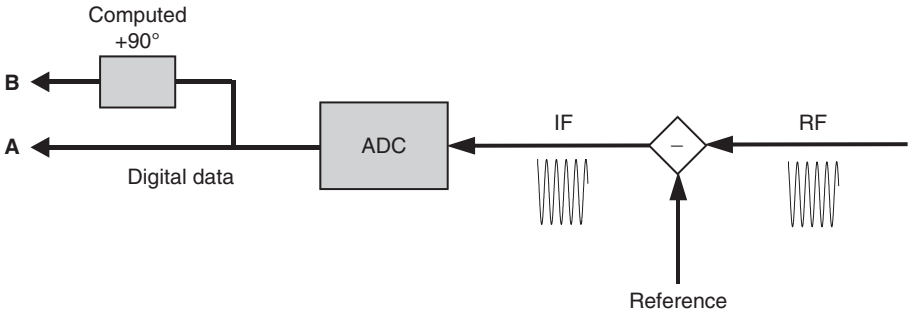


Figure 3.26. Schematic illustration of digital quadrature detection. A single channel is digitised as a suitable intermediate frequency input (*if*) and the second channel is generated numerically with an identical amplitude and a precise 90° phase shift.

copied and manipulated numerically to generate a second data set with a mathematically defined 90° phase difference from the originally detected data. This synthesized second channel data has the same amplitude and a precisely defined phase difference, so the imbalance in the quadrature channels that gives rise to quad images no longer occurs in the DQD scheme, and these images are thus eliminated. Since direct digitisation of the rf is not possible (current ADCs do not run fast enough for this) down-mixing of the NMR signal to a suitably low *intermediate frequency* (*if*) must occur prior to digitisation, which nowadays take place with input frequencies of anything from kilohertz up to tens of megahertz, this made possible by developments coming from telecommunications (Fig. 3.26). With digital schemes, it also becomes possible to offset the reference frequency used at the detection stage from that used for the transmitter pulse, and this then eliminates any "zero-frequency" glitch or spike that may be seen at the centre of spectra when analogue receivers are used that produce a residual DC offset of the detected FID.

3.2.5. Phase cycling

As mentioned above, the removal of quadrature images can be achieved through the process of *phase cycling*, an experimental method that lies at the very heart of almost every NMR experiment. This process involves repeating a pulse sequence with identical timings and pulse tip angles but with judicious changes to the *phases* of the rf pulse(s) and to the routing of the data to the computer memory blocks (often loosely referred to as the receiver phase cycle). The aim in all this is for the desired signals to add coherently with time averaging whereas all other signals, whether from unwanted NMR transitions or from instrumental imperfections, cancel at the end of the cycle and do not appear in the resulting spectrum. We shall see the importance of phase cycles throughout the remainder of the book, particularly in the sections on multipulse 1D and 2D NMR, but as an introduction we remain with the idea of analogue quadrature detection and the cancellation of quadrature images.

Recall that the quadrature scheme results in two signals differing in phase by 90° that are stored in two separate memory blocks (here designated 1 and 2) and that phase and amplitude imbalance between these gives rise to the images. One solution to the problem is to ensure that data from each channel, A and B, contributes equally to the two memory blocks 1 and 2. This can be achieved by performing two experiments, the first with a $90°_x$ pulse the second with $90°_y$ and adding the data. To keep the cosine and sine components in separate memory regions for use in the FT routine, appropriate data routing must also be used (Fig. 3.27a and b). This is generally handled internally by the spectrometer, being defined as the 'receiver phase' from the operator's point of view, and taking values of 0, 90, 180 or 270° or x, y, $-x$ and $-y$. Note, however, that the phase of the receiver *reference* rf does not alter on sequential scans, only the data routing is changed. For any pulse NMR sequence, it is necessary to define the rf phases and the receiver phase to ensure retention of the desired signals and cancellation of artefacts.

By use of the two-step phase cycle, it is therefore possible to compensate for the effects of imbalance in the two receiver channels. It is also possible to remove extraneous signals that may occur, such as from DC offsets in the receiver, by simultaneously inverting the phase of the rf pulse and the receiver; thus, as shown in Fig. 3.27, a steps to c and likewise b becomes d. The NMR signals will follow the phase of the pulse and so will add in the memory whereas offsets or spurious signals will be independent of this and so will cancel. This gives us a second possible two-step phase cycle which, when combined with the first, produces an overall four-step cycle known as CYCLOPS [8] (CYCLically Ordered PhaSe cycle, Table 3.1). This is the standard phase cycle used for one-pulse acquisitions on all spectrometers and is often nested within the phase cycles of 2D experiments again with the aim of removing receiver artefacts.

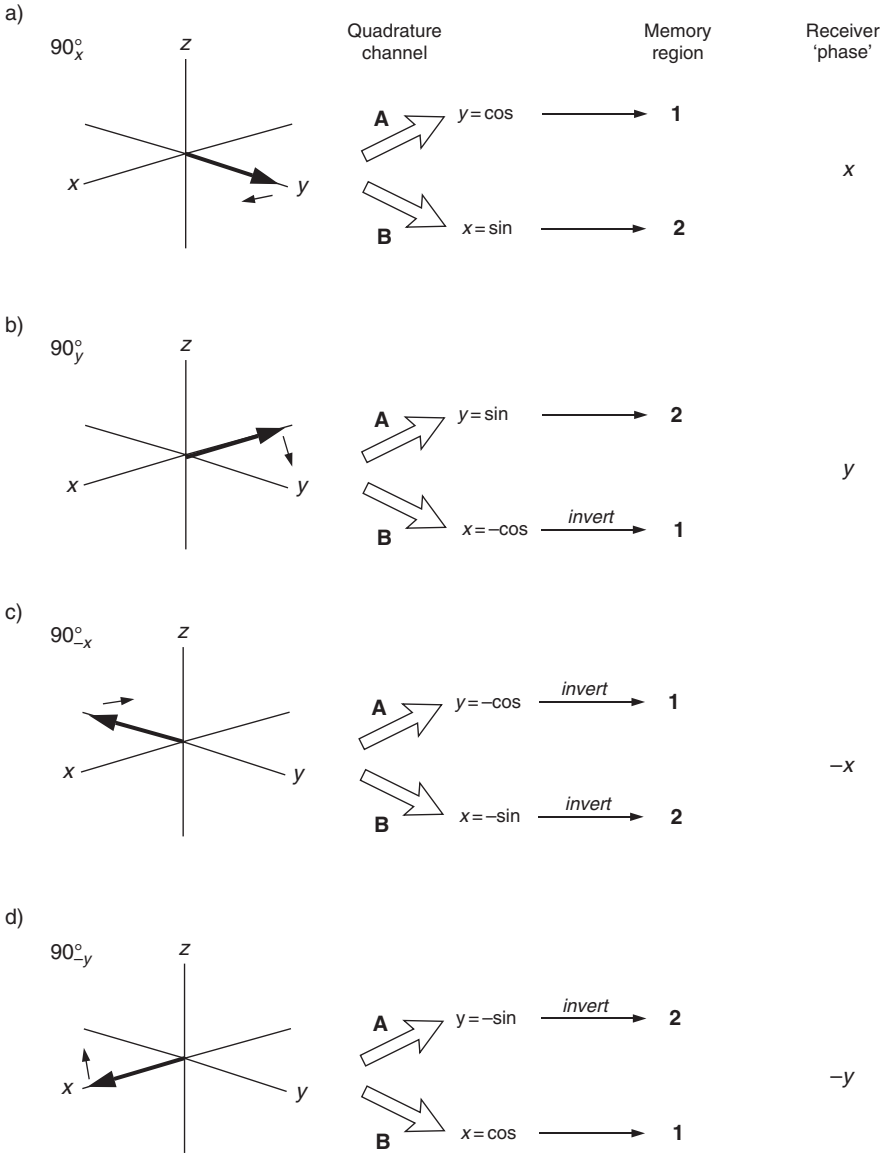


Figure 3.27. Phase cycling. The CYCLOPS scheme cancels unwanted artefacts whilst retaining the desired NMR signals. This four-step phase cycle is explained in the text.

Table 3.1. The four-step CYCLOPS illustrated in Fig. 3.27. This shorthand notation is conventionally used to describe all pulse sequence phase cycles

Scan number	Pulse phase	Receiver phase
1	x	x
2	y	y
3	$-x$	$-x$
4	$-y$	$-y$

3.2.6. Dynamic range and signal averaging

When sampling the FID, the analogue-to-digital converter (ADC or digitiser) limits the frequency range one is able to characterise (i.e. the spectral width) according to how fast it can digitise the incoming signal. In addition to limiting the *frequencies*, ADC performance also limits the *amplitudes* of signals that can be measured. The digitisation process converts the electrical NMR signal into a binary number proportional to the magnitude of the signal. This digital value is defined as a series of computer bits, the number of which describes the *ADC resolution*. Typical digitiser resolutions on modern spectrometers operate with 14 or 16 bits. The 16-bit digitiser is able to represent values in the range $\pm 32{,}767$, that is, 2^{15}-1 with one bit reserved to represent the sign of the signal. The ratio between the largest and smallest detectable value (the most and least significant bits), 32,767:1, is the *dynamic*

range of the digitiser. If we assume the receiver amplification (or *gain*) is set such that the largest signal in the FID on each scan fills the digitiser, then the smallest signal that can be recorded has the value 1. Any signal whose amplitude is less than this will not trigger the ADC; the available dynamic range is insufficient. However, noise will also contribute to the detected signal and this may be sufficiently intense to trigger the least significant bit of the digitiser. In this case, the small NMR signal will be recorded as it rides on the noise, and signal averaging therefore leads to summation of this weak signal, meaning even those whose amplitude is below that of the noise may still be detected. However, the digitiser may still limit the detection of smaller signals in the presence of very large ones when the signal-to-noise ratio is high. Figure 3.28 illustrates how a reduction in the available dynamic range limits the observation of smaller signals when thermal noise in the spectrum is low. This situation is most commonly encountered in proton studies, particularly of protonated aqueous solutions where the water resonance may be many thousand times that of the solute. Such intensity differences impede solute signal detection, so many procedures have been developed to selectively reduce the intensity of the H_2O resonance and ease the dynamic range requirements; some of these *solvent suppression* schemes are described in Chapter 10.

If any signal is so large that it cannot fit into the greatest possible value the ADC can record, the signal intensity will not be measured correctly and this results in severe distortion of the spectrum (Fig. 3.29a). The effect can be recognised in the FID as 'clipping' of the most intense part of the decay (Fig. 3.29b), which results from setting the receiver gain or amplification too high; this must be set appropriately for each sample studied, either by manual adjustment or more conveniently via the spectrometer's automated routines.

Signal averaging

The repeated acquisition and summation of a FID leads to an overall increase in the signal-to-noise ratio as the NMR signals add coherently over the total number of scans, NS, whereas the noise, being random, adds according to $\sqrt{NS}$. Thus, the signal-to-noise ratio increases according to $\sqrt{NS}$. In other words, to double the signal-to-noise ratio it is necessary to acquire *four times* as many scans (Fig. 3.30). It is widely believed that continued averaging leads to continuous improvement in the signal-to-noise ratio, although

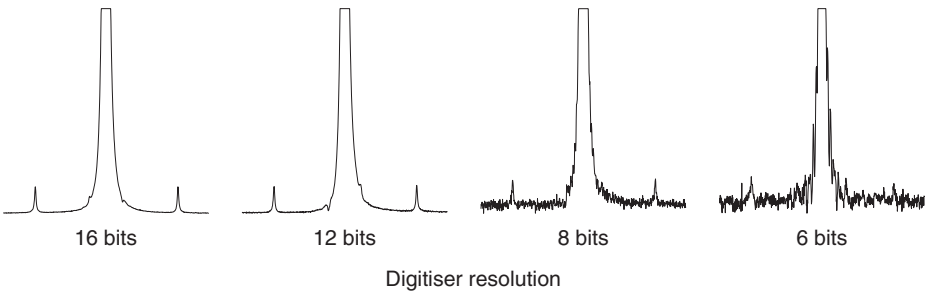


Figure 3.28. Dynamic range and the detection of small signals in the presence of large ones. As the digitiser resolution and hence its dynamic range are reduced, the carbon-13 satellites of the parent proton resonance become masked by noise until they are barely discernible with only 6-bit resolution (all other acquisition parameters were identical for each spectrum). The increased noise is *digitisation* or *quantisation noise* (see text below).

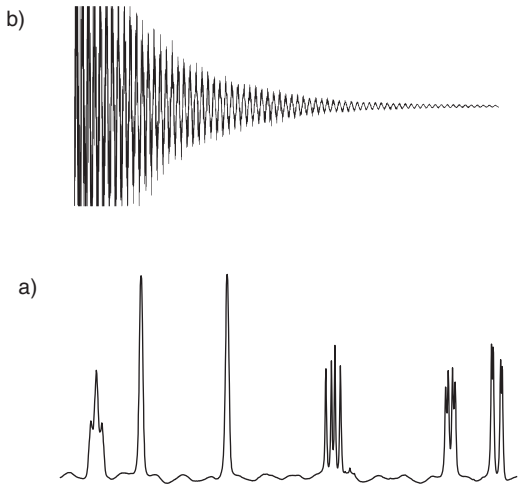


Figure 3.29. Receiver or digitiser overload distorts the spectrum baseline (a). This can also be recognised as a "clipping" of the FID (b).

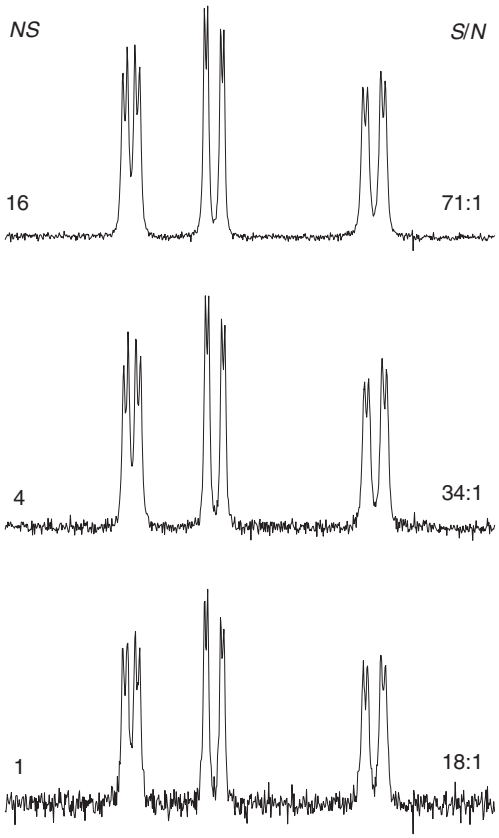


Figure 3.30. Signal averaging produces a net increase in the spectrum signal-to-noise ratio (*S/N*). This improves as the *square root* of the number of acquired scans (NS) because noise, being random, adds up *more slowly* than the NMR signal.

this is only true up to a point. Each time a scan is repeated, a binary number for each sampled data point is added to the appropriate computer memory location. As more scans are collected, the total in each location will increase as the signal 'adds up'. This process can only be repeated if the cumulative total fits into the computer word size; if it becomes too large, it cannot be recorded, potentially leading to corruption of the data (although spectrometers generally handle this problem internally by simply dividing the data and reducing the ADC resolution whenever memory overflow is imminent, so preserving the data and allowing the acquisition to continue). The point at which the computer word length can no longer accommodate the signal will be dependent on the number of bits used by the ADC (its resolution) and the number of bits in the computer word. For example, if both the ADC and the word length comprise 16 bits, then any signal that fills the ADC will also fill the memory, so no further scans may be collected. Thus, the word length must be greater than the ADC resolution to allow signal averaging, and the larger the word length or the smaller the ADC resolution, the more scans can be collected. Typical word lengths on modern instruments are 32 bits, so with a 16-bit ADC it is possible to collect around 66,000 scans (2^{32-16}) if the largest signal fills the digitiser (in fact, more scans than this could be collected since the NMR signal also contains noise that does not add coherently, meaning the memory will not fill as rapidly as this simple estimation suggests). Since signal averaging is most necessary when noise levels are relatively high and because longer word lengths are used on modern instruments, the problem of overflow is now rarely encountered.

Oversampling and digital filtering

As per the standard usage on the latest generation spectrometers, it has long been realised that the effective dynamic range can, in certain circumstances, be improved by *oversampling* the FID [9]. In particular, this applies when the *digitisation noise* (or *quantisation noise*) is significant and limits the attainable signal-to-noise level of the smallest signals in the spectrum. This noise arises from the 'rounding error' inherent in the digitisation process, in which the analogue NMR signal is characterised in discrete one-bit steps that may not accurately represent the true signal intensity (Fig. 3.31). Although this is not true noise, but arises from systematic errors in the sampling process, it does introduce noise into the NMR spectrum. Figure 3.28 has already shown how larger errors in the digitisation process, caused by fewer ADC bits, increases the noise level. In realistic situations, this noise becomes dominant when the receiver amplification or gain is set to a relatively low level such that the thermal (analogue) noise arising from the probe and

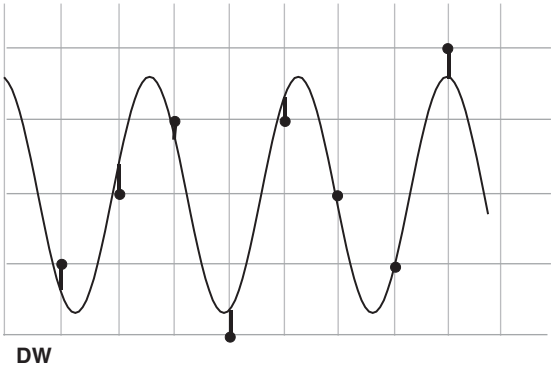


Figure 3.31. Digitisation errors. Discrete digital sampling of an analogue waveform introduces errors in amplitude measurements, indicated by the vertical bars. These errors ultimately contribute to additional noise in the spectrum.

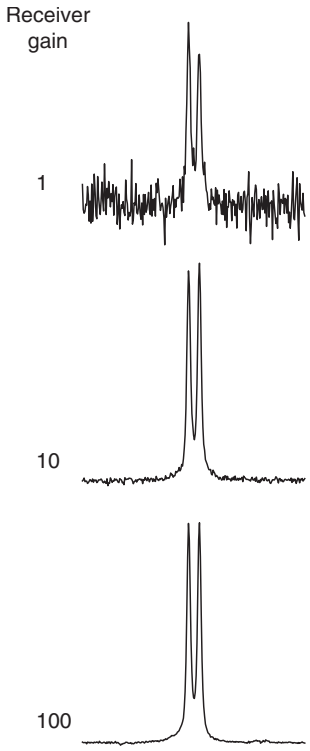


Figure 3.32. Digitisation noise. At high receiver gain settings, the noise in this proton spectrum is vanishingly small, meaning system thermal noise is low. As the gain is reduced, the amplitude of the NMR signal (and thermal noise) is reduced and digitisation noise becomes significant relative to the NMR signal.

amplifiers is negligible (Fig. 3.32). This is most likely to arise in proton spectroscopy when attempting to observe small signals in the presence of far larger ones and is especially significant with the introduction of low noise cryogenic probes. High gain settings are therefore favoured to overcome digitisation noise.

Oversampling, as the name suggests, involves digitising the data at a much faster rate than is required by the Nyquist condition or, equivalently, acquiring the data with a much greater spectral width than would normally be needed. The digitisation noise may then be viewed as being distributed over a far greater frequency range, such that in the region of interest this noise has reduced intensity so leading to a sensitivity improvement (Fig. 3.33). If the rate of sampling has been increased by an oversampling factor of N_{os}, then the *digitisation* noise is reduced by $\sqrt{N_{os}}$. Thus, sampling at four times the Nyquist frequency would theoretically produce a twofold reduction in noise, which in turn equates to an effective gain in ADC resolution of 1 bit. Likewise, oversampling by a factor of 16 corresponds to a fourfold noise reduction (a resolution gain of 2 bits) and a factor of 64 gives an 8-fold reduction or an extra 3 bits. The degree of oversampling that can be achieved is limited by the digitisation speed of the ADC, and typical values of N_{os} for proton observation are 16–32, meaning spectral widths become a few hundred kilohertz and the ADC resolution may be increased by 2–3 bits in favourable circumstances (when thermal noise can be considered insignificant relative to digitisation noise).

To maintain the desired digital resolution in the spectral region of interest when oversampling, data sets would have to be enlarged according to the oversampling factor also, which would demand greater storage capacity and slower data processing. To overcome these limitations, modern spectrometers generally combine oversampling with digital signal processing methods. Since one is really only interested in a relatively small part of our oversampled data set ($1/N_{os}$ of it), the FID is reduced after digitisation to the conventional number of data points by the process of *decimation* (literally 'removing one-tenth of') prior to storage. This, in effect, takes a running average of the oversampled data points, leaving one point for every N_{os} sampled, at intervals defined by the Nyquist condition for the desired spectral window (the peculiar distortions seen at the beginning of a digitally processed FID arise from this decimation process). The resulting FID then has the same number of data points as it would have if it had been sampled normally. These digital signal processing steps are generally performed with a dedicated processor after the ADC (Fig. 3.9, box b) and are typically left invisible to the user (short of setting a few software flags perhaps) although they can also be achieved by separate post-processing of the

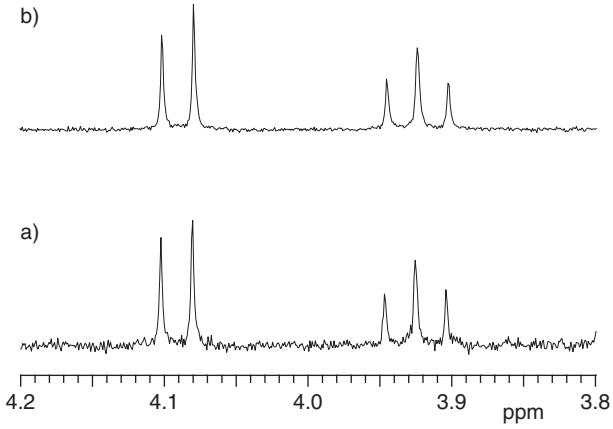


Figure 3.33. Enhanced sensitivity can be realised in favourable cases by oversampling the data and hence reducing digitisation noise. Spectrum (a) shows part of a conventional proton spectrum sampled according to the Nyquist criterion. Oversampling the data by a factor of 24 as in (b) provides a sensitivity gain (all other conditions as for (a)).

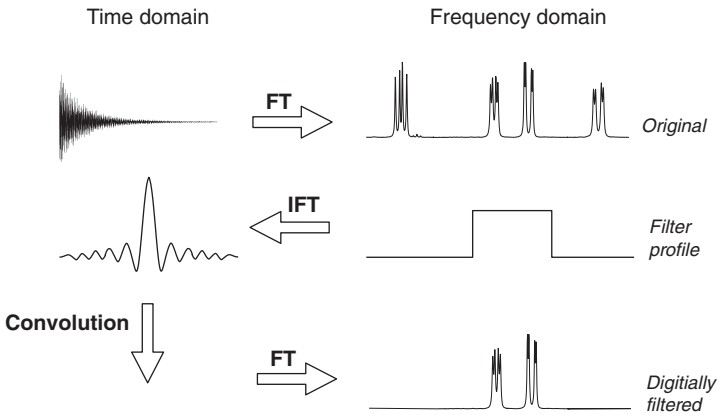


Figure 3.34. The use of a digital filter to observe a selected region of a spectrum. The desired frequency window profile is subject to an inverse Fourier Transformation and the resulting time-domain function convoluted with the raw FID. Transformation of the modified data produces a spectrum containing only a subset of all resonances as defined by the digital filter.

original data or even included in the FT routine itself [10]. The use of fast dedicated processors means the calculations can readily be achieved as the data are acquired thus not limiting data collection.

One further advantage of digital processing of the FID is the ability to mathematically define frequency filters that have a far steeper and more complete cut-off than can be achieved by analogue filtration alone, meaning signal aliasing can be eliminated. Spectral widths may then be set to encompass only a sub-section of the whole spectrum, allowing *selective detection* of NMR resonances [11]. The ability to reduce spectral windows in this way without complications from signal aliasing has considerable benefit when acquiring 2D NMR data in particular. The principle behind the filtration is as follows (Fig. 3.34). The digital time-domain filter function is given by the inverse FT of the desired frequency-domain window. This function may then be convoluted with the digitised FID such that, following FT, only those signals that fell within the originally defined window remain in the spectrum. One would usually like this window to be rectangular in shape although in practice such a sharp cut-off profile cannot be achieved without introducing distortions. Various alternative functions have been used which approach this ideal, generally with the property that the more coefficients used in the function, the steeper is their frequency cut-off. Since these operate only after the ADC, the analogue filter (Fig. 3.9, box a) is still required to reject broadband noise from outside the oversampled spectral width, but its frequency cut-off is now far removed from the normal spectral window and its performance less critical.

3.2.7. Window functions

There exist various ways in which the acquired data can be manipulated prior to FT in an attempt to enhance its appearance or information content. Most commonly one would either like to improve the signal-to-noise ratio of the spectrum to help reveal resonances above the noise level or to improve the resolution of the spectrum to reveal hidden fine structure. Many mathematical weighting functions, known as *window functions*, have been proposed to achieve the desired result, but a relatively small number have come into widespread use, some of which are illustrated in Fig. 3.35. The general philosophy behind all these functions is the same. In any FID, the NMR signal intensity declines throughout the acquisition time whereas the noise amplitude remains constant, meaning the relative noise level is greater in the latter part. Decreasing the tail of the FID will therefore help reduce the noise amplitude in the transformed spectrum, so enhancing *sensitivity*. Conversely, to increasing the middle and latter part of the FID equates to retarding the decay of the signal, thus narrowing the resonance and so enhancing *resolution*.

One word of warning here before proceeding. The application of window functions has also been referred to in the past as 'digital processing' or 'digital filtering' of the data. However, this should not be confused with the digital signal processing terminology introduced in previous sections and in widespread use nowadays. The adoption of the terms 'window' or 'weighting' function is therefore recommended in the context of sensitivity or resolution enhancement.

Sensitivity enhancement

As suggested above, the noise amplitude in a spectrum can be attenuated by de-emphasising the latter part of the FID, and the most common procedure for achieving this is to multiply the raw data by a decaying exponential function (Fig. 3.35a). This process is therefore also referred to as *exponential multiplication*. Because this forces the tail of the

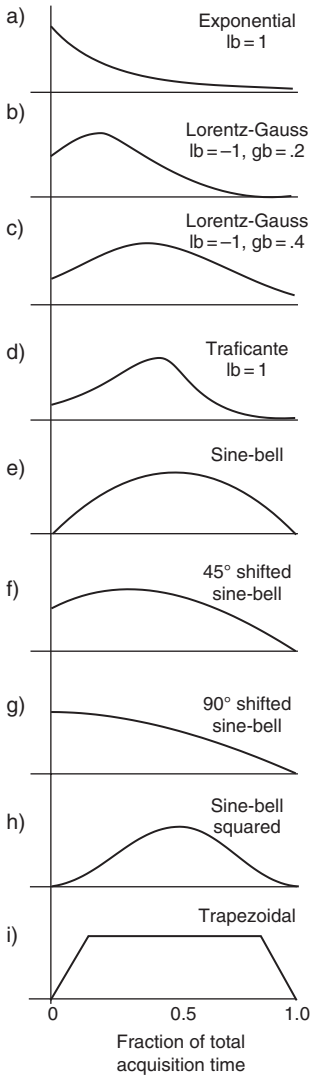


Figure 3.35. Some commonly employed window functions. These are used to modify the acquired FID to enhance sensitivity and/or resolution (lb = line broadening parameter, gb = Gaussian broadening parameter, i.e. the fraction of the acquisition time when the function has its maximum value; see text).

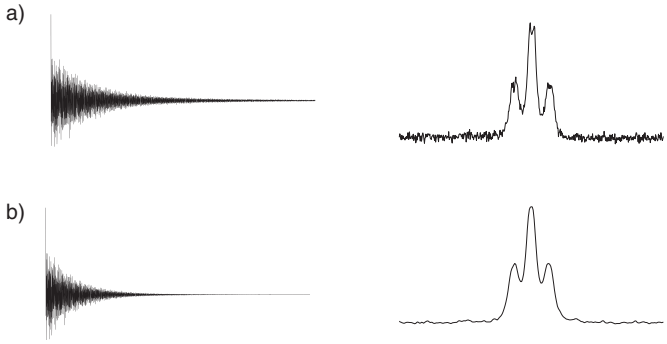


Figure 3.36. Exponential multiplication of the FID can be used to reduce noise in the spectrum. (a) Raw FID and spectrum following Fourier transformation. (b) Results after exponential processing with lb = 1 Hz.

FID towards zero, it is also suitable for the apodisation of truncated data sets. However, it also increases the apparent decay rate of the NMR signal, causing lines to broaden, meaning one compromises resolution for the gain in sensitivity (Fig. 3.36). Using too strong a function, that is one that causes the signal to decay too rapidly, can actually lead to a decrease in signal-to-noise ratio in the resulting spectrum because the broadening of the lines causes a reduction of their peak heights. Spectrometer software usually allows one to define the amount of line broadening directly in hertz (here parameter lb), so exponential multiplication is rather straightforward to use and allows the chemist to experiment with different degrees of broadening to attain a suitable result. The optimum balance between reducing noise and excessive line broadening is reached when the decay of the window function matches the natural decay of the NMR signal, in which case it is known as the *matched filter* [12], which results in a doubling of the resonance linewidth. For the exponential function to truly match the FID, the decay of the NMR signal must also be an exponential (meaning the NMR resonances have a Lorentzian lineshape), which on a correctly adjusted spectrometer is usually assumed to be the case (at least for spin-½ nuclei). Despite providing the maximum gain in sensitivity, the matched condition is often not well suited to routine use in proton spectroscopy as the resulting line broadening and loss of resolution may preclude the separation of closely space lines, so less line broadening than this may be more appropriate. Furthermore, different resonances in the spectrum often display different unweighted linewidths, so the matched condition cannot be met for all resonances simultaneously. For routine proton work, it turns out to be convenient to broaden the line by an amount equal to the digital resolution in the spectrum, as this leads to some sensitivity enhancement but to a minimal increase in linewidth. For heteronuclear spectra, resolution is usually a lesser concern and line broadening of a few hertz is commonly employed (comparable to or slightly greater than the digital resolution). This can be considered essential to attenuate the distortions about the base of resonances arising, in part, from truncation of the data set (Section 3.2.3 and Fig. 3.17). For spectra that display resonances that are tens or even hundreds of hertz wide (most notably those of quadrupolar nuclei), the amount of line broadening must be increased accordingly to achieve any appreciable sensitivity gain.

Resolution enhancement

One might suppose that to improve the resolution in spectra, we should apply a function that enhances the latter part of the FID, so increasing the decay time. The problem with this approach is that by doing this one also increases the noise amplitude in the tail of the FID and emphasises any truncation of the signal that may be present, so increasing the potential for undesirable 'truncation wiggles'. A better approach is to apply a function that initially counteracts the early decay but then forces the tail of the FID to zero to provide the necessary apodisation. The most popular function for achieving this has been the *Lorentz–Gauss transformation* [13] (Fig. 3.35 b and c), sometimes loosely referred to as *Gaussian multiplication* (although this strictly refers to yet another mathematical weighting function) and also known as double-exponential multiplication. This transforms the usual Lorentzian lineshape into a Gaussian lineshape, which has a somewhat narrower profile, especially around the base (Fig. 3.37), and it is this feature that allows the resolution of closely spaced lines.

The shape of the function is altered by two variable parameters that define the degree of line narrowing in the resulting spectra and the point during the acquisition time at which the function reaches its maximum value. These are usually presented to the operator as negative line broadening in hertz (here parameter lb) and as a fraction of the total acquisition time (here parameter gb), respectively. The choice of suitable values for these usually comes down to a case of trial-and-error, and different optimum values may be found for different groups of peaks within the same spectrum. Modern NMR

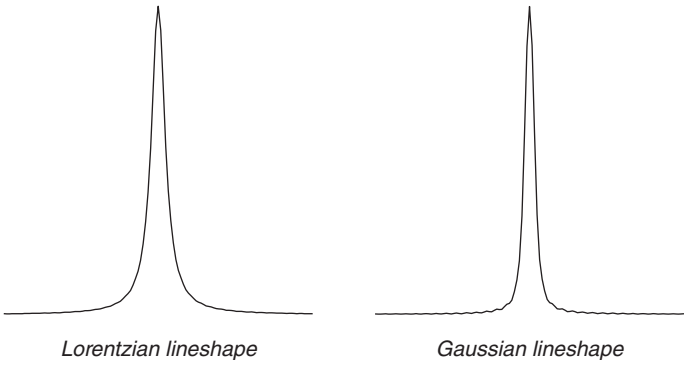


Figure 3.37. A comparison of the Lorentzian and Gaussian lineshapes.

processing packages usually allow an interactive variation of the parameters during which one can observe both the window function itself and the resulting spectrum, so an optimum can rapidly be determined. A more negative line broadening will produce narrower lines whilst positioning the maximum further along the decay will enhance this effect. This will also lead to greater distortion of the resonances about the base and to degradation of sensitivity as the early signals are sacrificed relative to the noise (Fig. 3.38). In any case, some reduction in sensitivity will invariably result, and it is necessary to have reasonable signal-to-noise for acceptable results. It has been suggested [14] that the optimum Gaussian resolution enhancement for routine use on spectra with narrow lines aims for a reduction in linewidth by a factor of 0.66 for which the function maximum should occur at a time of $1/(\Delta\nu_{1/2})$ s, that is, the inverse of the resonance linewidth in the absence of window functions. The appropriate figures for setting in the processing software can then be arrived at by simple arithmetic. For example, a starting linewidth of 1 Hz will require a line narrowing of 0.66 Hz and the maximum to occur after 1 s. With a typical proton acquisition time of around 3 s, the function should therefore reach a maximum at 1/3 of the total acquisition. Gaussian multiplication can be used to good effect when combined with zero-filling, thus ensuring the digital resolution is sufficiently fine to define any newly resolved fine structure.

In more recent years, new window functions have been introduced [15, 16] that are similar to the Lorentz–Gauss window but which aim to improve resolution without a discernible reduction in sensitivity. These so called *TRAF* functions (Fig. 3.35d) generally aim to enhance the middle part of the FID but also use matched filtration of the later part to attenuate noise. They are appearing in NMR software packages and provide an alternative to the established functions. Another function, more commonly used in the processing of 2D data sets, is the *sine-bell* window, which has no adjustable parameters. This comprises one half of a sine wave, starting at zero at the beginning of the FID, reaching a maximum half way through the acquisition and falling back to zero by the end of the decay (Fig. 3.35e). The function always has zero intensity at the end, so eliminates the truncation often encountered in 2D data sets. This tends to be a rather

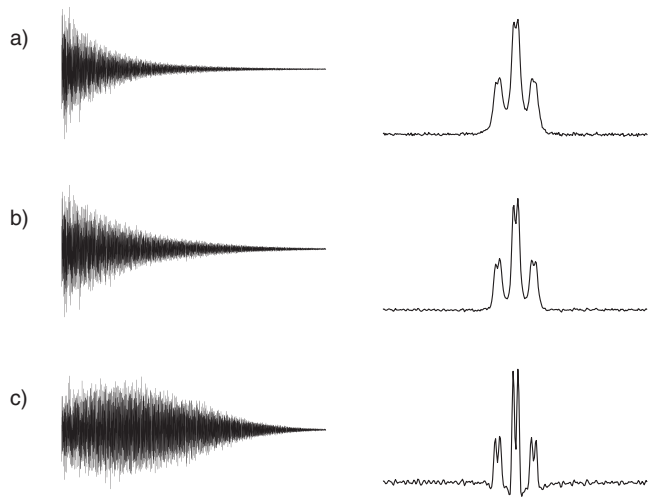


Figure 3.38. The Lorentz–Gauss transformation ("Gaussian multiplication") can be used to improve resolution. (a) Raw FID and spectrum following Fourier transformation and results after the Lorentz–Gauss transformation with (b) lb = −1 Hz, gb = 0.2 and (c) lb = −3 Hz and gb = 0.2.

severe resolution enhancement function that can introduce undesirable lineshape distortions and also produces severe degradation in sensitivity because the early part of the NMR signal is heavily attenuated, and hence this is rarely used in 1D spectra. A variation on this is the *phase-shifted sine-bell* for which the point at which the maximum occurs is a variable (Fig. 3.35f and g) and can be moved towards the start of the FID. Again, this is commonly applied in 2D processing and has the advantage that the user has the opportunity to balance the gain in resolution against the lineshape distortion and sensitivity degradation. A final variant is the (optionally shifted) *squared sine-bell* (Fig. 3.35h), which has similar properties to the sine-bell but tails more gently at the edges, which can invoke subtle differences in 2D spectra. Likewise, the *trapezoidal* function (Fig. 3.35i) is sometimes used in the processing of 2D data.

3.2.8. Phase correction

It has already been mentioned in Section 2.2 that the phase of a spectrum needs correcting following FT because the receiver reference phase does not exactly match the initial phase of the magnetisation vectors. This error is constant for all vectors, and since it is independent of resonance frequencies, it is referred to as the 'zero-order' phase correction (Fig. 3.39). Practical limitations also impose the need for a frequency-dependent or 'first-order' phase correction. Consider events immediately after the pulse is applied to the sample. A short period of time, the pre-acquisition delay, DE, is required for the spectrometer electronics to recover from the effect of the pulse before an undistorted FID can be collected. This delay is typically tens of microseconds, during which magnetisation vectors will evolve a little according to their chemical shifts so that at the point digitisation begins they no longer have the same phase (Fig. 3.40). Clearly those resonances with the greatest shifts require the largest corrections. If the DE is small relative to the frequency offsets, the phase errors have an approximately linear offset dependency and can be removed. If the delay becomes large, the correction cannot be made without introducing a rolling spectrum baseline. The appearance of these zero- and first-order phase errors is illustrated in the spectra of Fig. 3.41.

Typically, both forms of error occur in a spectrum directly after the FT. The procedure for phase correction is essentially the same on all spectrometers. The zero-order correction is used to adjust the phase of one signal in the spectrum to pure absorption mode, as judged 'by eye', and the first-order correction is then used to adjust the phase of a signal far away from the first in a similar manner. Ideally, the two chosen resonances should be as far apart in the spectrum as possible to maximise the frequency-dependent effect. Experimentally, this process of phase correction involves mixing of the real and imaginary parts of the spectra produced by the FT process such that the final displayed 'real' spectrum is in pure absorption mode whereas the usually unseen 'imaginary' spectrum is pure dispersion.

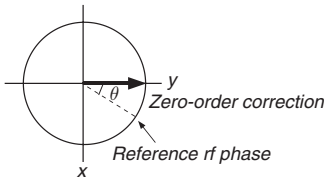
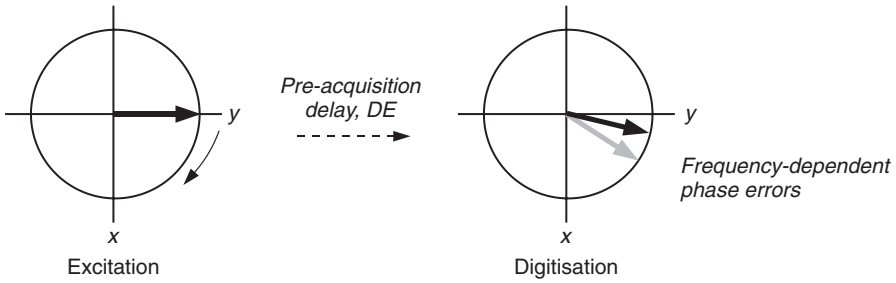


Figure 3.39. Zero-order (frequency-independent) phase errors arise when the phase of the detected NMR signals does not match the phase of the receiver reference rf. All resonances in the spectrum are affected to the same extent.



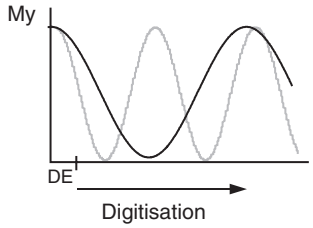


Figure 3.40. First-order (frequency-dependent) phase errors arise from a dephasing of magnetisation vectors during the pre-acquisition delay that follows the excitation pulse. When data collection begins, vectors with different frequencies have developed a significant phase difference, which varies across the spectrum.

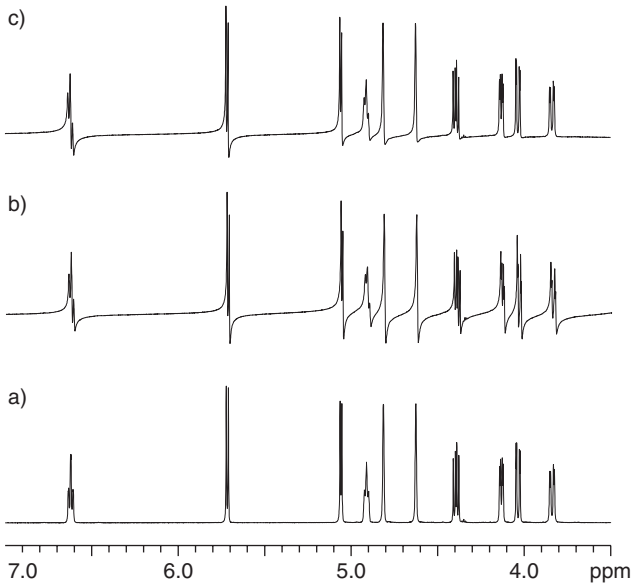


Figure 3.41. Zero- and first-order phase errors in a ^{1}H spectrum. (a) The correctly phased spectrum, (b) the spectrum in the presence of frequency-independent (zero-order) phase errors and (c) the spectrum in the presence of frequency-dependent (first-order) phase errors.

3.3. PREPARING THE SAMPLE

This section describes some of the most important aspects of sample preparation for NMR analysis, a topic that is all too frequently given insufficient consideration in the research laboratory. Even for routine applications of NMR, it is wise for the chemist to adopt a systematic strategy for sample preparation that will give consistently good results, so saving instrument time and eliminating frustration.

3.3.1. Selecting the solvent

Since all modern NMR spectrometers rely on a deuterium lock system to maintain field stability, a deuterated solvent is invariably required for NMR, and these are now available from many commercial suppliers. In a busy organic chemistry laboratory, the volume of NMR solvents used can be surprisingly large, and the cost of the solvent may play some part in its selection. Generally, chloroform and water are the most widely used solvents and are amongst the cheapest, although dimethylsulfoxide is becoming increasingly favoured, notably in the industrial sector, because of its useful solubilising properties. Acetone, methanol and dichloromethane are also in widespread routine use.

The selection of a suitable solvent for the analyte is based on a number of criteria. The most obvious requirement is that the analyte be soluble in it at the concentration required for the study. This will be dependent on a diverse range of factors including the fundamental sensitivity of the nucleus, the overall sensitivity of the instrument, the nature of the experiment and so on. If experiments are likely to be performed at very low temperatures, then it is also important to know that the solute remains in solution when the temperature is reduced so that a precipitate does not form in the NMR tube and degrade resolution. For experiments at temperatures other than ambient probe temperatures (which may be a few degrees greater than the ambient room temperature), the melting and boiling points of the solvents must also be considered. Table 3.2 summarises some important properties of the most commonly encountered solvents. For work at very high temperatures, dimethylsulphoxide or toluene is generally the solvents of choice, whilst for very low temperatures, dichloromethane, methanol or tetrahydrofuran are most appropriate. Even for experiments performed at ambient temperatures, solvent melting points may be limiting; on a cold day, dimethylsulfoxide solutions can freeze in the laboratory. The viscosity of the solvent will also influence the resolution that can be obtained with the best performance provided by the least viscous solvents, such as acetone (which, for this reason, is used as the solvent for spectrometer resolution tests). The sharpest possible lock signal is also beneficial in experiments based on difference spectroscopy, the most important of these being the NOE difference experiment (Chapter 8).

For proton and carbon spectroscopy, the chemical shifts of the solvent resonances need to be anticipated as these may occur in a particularly unfortunate place and interfere with resonances of interest. In proton spectroscopy, the observed solvent resonance arises from the residual *protonated* species (NMR solvents are typically supplied with deuteration

Table 3.2. Properties of the common deuterated solvents. Proton shifts, δH, and carbon shifts, δC, are quoted relative to tetramethylsilane (TMS) (proton shifts are those of the residual partially *protonated* solvent). The proton shifts of residual HDO/H_2O vary depending on solution conditions

Solvent	δ_H/ppm	$\delta_{(HDO)}$/ppm	δ_C/ppm	Melting point/°C	Boiling point/°C
Acetic acid-d_4	11.65, 2.04	11.5	179.0, 20.0	16	116
Acetone-d_6	2.05	2.0	206.7, 29.9	−94	57
Acetonitrile-d_3	1.94	2.1	118.7, 1.4	−45	82
Benzene-d_6	7.16	0.4	128.4	5	80
Chloroform-d_1	7.27	1.5	77.2	−64	62
Deuterium oxide-d_2	4.80	4.8	–	4	101
Dichloromethane-d_2	5.32	1.5	54.0	−95	40
N,N-dimethyl formamide-d_7	8.03, 2.92, 2.75	3.5	163.2, 34.9, 29.8	−61	153
Dimethylsulfoxide-d_6	2.50	3.3	39.5	18	189
Methanol-d_4	4.87, 3.31	4.9	49.2	−98	65
Pyridine-d_5	8.74, 7.58, 7.22	5.0	150.4, 135.9, 123.9	−42	114
Tetrahydrofuran-d_8	3.58, 1.73	2.4	67.6, 25.4	−109	66
Toluene-d_8	7.09, 7.00, 6.98, 2.09	0.4	137.9, 129.2, 128.3, 125.5, 20.4	−95	111
Trifluoroacetic acid-d_1	11.30	11.5	164.2, 116.6	−15	75
Trifluoroethanol-d_3	5.02, 3.88	5.0	126.3, 61.5	−44	75

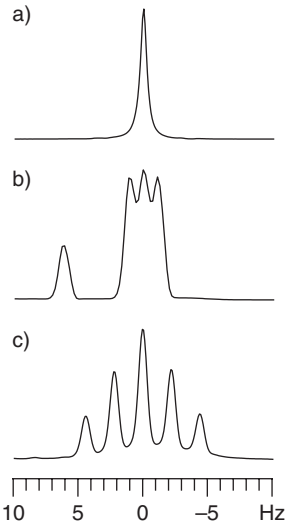


Figure 3.42. Residual protonated resonances of deuterated solvents (a) $CHCl_3$ in $CDCl_3$, (b) $CHDCl_2$ in CD_2Cl_2 and (c) CHD_2COCD_3 in $(CD_3)_2CO$. The multiplicity seen in (b) and (c) arises from two-bond (geminal) couplings to spin-1 deuterium producing a 1:1:1 triplet and a 1:2:3:2:1 quintet respectively. The left-hand singlet in (b) is residual CH_2Cl_2 in the solvent, the shift difference arises from the H–D isotope shift of 6 Hz.

levels in excess of 99.5%). For routine studies, where a few milligrams of material may be available, the lower specification solvents should suffice for which the residual protonated resonance is often comparable in magnitude to that of the solute. Solvents with higher levels of deuteration are beneficial for proton spectroscopy when sample quantities are of the order of tens of micrograms or less and are likely to be used in conjunction with microsample techniques (Section 3.3.3). The solvent proton resonance, with the exception of chloroform and water, will comprise a multiplet from coupling to spin-1 deuterium. For example, dichloromethane displays the triplet of $CDHCl_2$ whilst acetone, dimethylsulphoxide or methanol show a quintuplet from CD_2H (Fig. 3.42). In carbon spectroscopy, the dominant resonance is often that of the deuterated solvent that again will be a multiplet owing to coupling with deuterium. For other common nuclei, interference from the solvent is rarely a consideration, the one notable exception to this being deuterium itself. Unless very large sample quantities are being used (many tens of milligrams) and one can be sure the solvent resonance will not overlap those of the solute, it is usually necessary to record ^2H spectra in *protonated* solvent since the huge signals of the deuterated solvent are likely to swamp those of interest. This then precludes the use of the lock system for maintaining field stability. The lack of a suitable lock reference requires a small quantity of the deuterated solvent to be added to the solution to provide an internal chemical shift reference (see below).

Aside from the solvent resonance itself, the other significant interference seen in proton spectra is water, which is present to varying degrees in all solvents and is also often associated with solutes. The water resonance is often rather broad, and its chemical shift can vary according to solution conditions. All solvents are hygroscopic to some degree (including deuterated water) and should be exposed to the atmosphere as little as possible to prevent them from becoming wet. Very hygroscopic solvents such as dimethylsulphoxide, methanol and water are best kept under a dry inert gas atmosphere. Solvents can also be obtained in smaller glass ampoules that are particularly suitable for those that are used only infrequently. Molecular sieves may also be used to keep solvents dry, although some care is required in filtering the prepared solution before analysis to remove fine particulates that may be present. Alternatively, the NMR solution may be passed through activated alumina to remove water, as part of the filtration process.

The solvent may also result in the loss of the resonances of exchangeable protons since these will become replaced by deuterons. In particular, this is likely to occur with deuterated water and methanol. To avoid deuterium replacement altogether, one must consider using H_2O containing 10–20% D_2O (to provide a lock signal) or d_3-methanol (CD_3OH). In these cases, a suitable solvent suppression scheme (Chapter 10) is required to attenuate the large solvent resonance (which may still lead to the loss of the exchangeable protons of the solute by the process known as saturation transfer). Studies in 90% H_2O dominate the NMR of biological macromolecules because solvent exchangeable protons, such as the backbone amide NH protons in peptides and proteins or the imido protons of DNA and RNA base pairs, often play a key role in the structure determination of these molecules [17, 18].

The nature of the solvent can also have a significant influence on the appearance of the spectrum, and substantial changes can sometimes be observed on changing solvents. Whether these changes are beneficial is usually difficult to predict, and a degree of trial

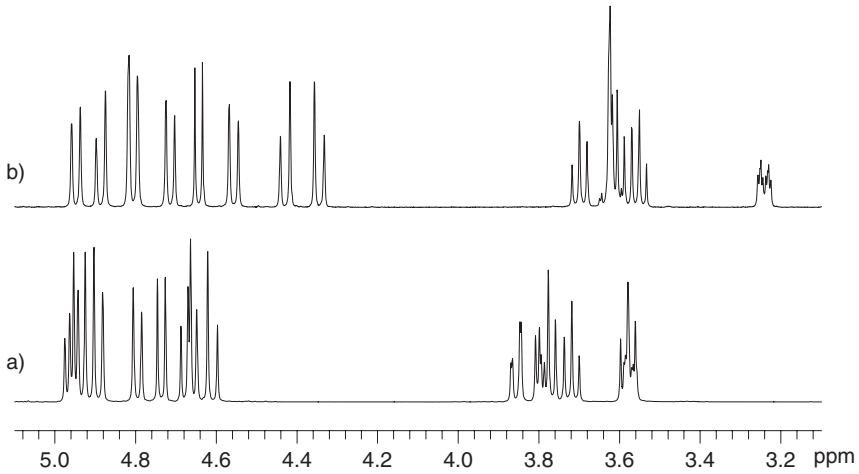
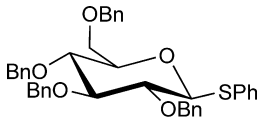


Figure 3.43. Changes in solvent can be used to improve resonance dispersion. The proton spectrum of the sugar **3.1** is shown in (a) $CDCl_3$ and (b) C_6D_6. Notice the appearance of the resonance at 3.2 ppm in (b) that was hidden at 3.6 ppm by another resonance in (a).



3.1

and error is required. A useful switch of solvent is from a non-aromatic solvent into an aromatic one, for example, chloroform to benzene (Fig. 3.43), making use of the magnetic anisotropy exhibited by the latter. In cases of particular difficulty, where the change from one solvent to another simply produces a different but equally unsuitable spectrum, then titration of the second solvent into the first may provide a suitable compromise between the two extremes. Such changes may prove useful in 1D experiments that use selective excitation of a specific resonance, by revealing a target proton when previously it was hidden, as in the example in Fig. 3.43. The selective removal of a resonance by the addition of another solvent can also prove useful in spectrum interpretation. Adding a drop of D_2O to an organic solution in the NMR tube, mixing thoroughly and leaving the mix to settle removes (or attenuates) acidic exchangeable protons. Acidic protons that are protected from the solvent, such as those in hydrogen bonds, may not fully exchange immediately, but may require many hours to disappear. This can provide a useful probe of H-bonding interactions.

3.3.2. Reference compounds

When preparing a sample it is common practice to add a suitable compound to act as an *internal* chemical shift reference in the spectrum, and the selection of this must be suitable for the analyte and the solvent. In proton and carbon NMR, the reference used in organic solvents is tetramethylsilane (TMS, 0.0 ppm), which has a number of favourable properties; it has a sharp 12-proton singlet resonance that falls conveniently to one end of the spectrum, and it is volatile so can be readily removed and it is chemically inert. In a few cases, this material may be unsuitable, such as in the study of silanes or cyclopropanes. For routine work, it is often not necessary to add any internal reference as the residual lines of the solvent itself can serve this purpose (Table 3.2). For aqueous solutions, the water soluble equivalent of TMS is partially deuterated sodium 3-(trimethylsilyl)propionate-d_4 (TSP-d_4), which is also referenced to 0.0 ppm. A volatile alternative is 1,4-dioxane (^{1}H 3.75 ppm, ^{13}C 67.5 ppm), which can be removed by lyophilisation. The standard reference materials for some other common nuclides are summarised in Table 3.3. Often these are not added into the solution being studied but are held in an outer, concentric jacket or within a separate axial capillary inside the solution, in which case the reference material may also be in a different solvent to that of the sample.

An alternative to adding additional reference materials is to use a so-called external reference. Here the spectrum of a separate reference substance is acquired before and/or after the sample of interest and the spectrum reference value carried over. Identical field settings should be used for both, which, on some older instruments, requires the same lock solvent, or an additional correction to the spectrum reference frequency must be used to compensate any differences. This restriction does not arise on instruments that use shifting of the lock transmitter frequency to establish the lock condition.

The referencing of more 'exotic' nuclei is generally less clear-cut than for those in common use, and in many cases, it is impractical to add reference materials to precious samples, and it is sometimes even difficult to identify what substance is the 'accepted' reference standard. In such cases, the IUPAC Ξ-scale can be used, which does not require use of a specific reference material. Instead this scheme defines the *reference frequency* for the reference material of each nuclide at a field strength at which the proton signal of TMS resonates at exactly 100.000 MHz. The reference frequencies are scaled appropriately for

Table 3.3. Spectrum reference materials (0 ppm) and reference frequencies (Ξ values) for selected nuclides

Nuclide	Primary reference	Ξ value/MHz	Alternative reference
^{1}H	Me_4Si	100.000,000	TSP-d_4 (aq)
^{2}H	Me_4Si	15.351	Trace deuterated solvent
^{6}Li	LiCl (aq)	14.717	
^{7}Li	LiCl (aq)	38.866	
^{10}B	$BF_3O(Et)_2$	10.746	H_2BO_3 (aq)
^{11}B	$BF_3O(Et)_2$	32.089	H_2BO_3 (aq)
^{13}C	Me_4Si	25.145,004	1,4-dioxan @ 67.5 ppm, TSP-d_4 (aq)
^{14}N	CH_3NO_2	7.224	
^{15}N	CH_3NO_2	10.136,767	NH_4NO_3 (aq), NH_3 (liq)[a]
^{17}O	H_2O	13.557	
^{19}F	$CFCl_3$	94.094,003	
^{29}Si	Me_4Si	19.867,184	
^{31}P	H_3PO_4 (85%)	40.480,747	
^{119}Sn	Me_4Sn	37.290,665	
^{195}Pt	$K_2[Pt(CN)_6]$	21.414,376	
^{207}Pb	Me_4Pb	20.920,597	

[a] ^{15}N shifts of biomolecular materials are more often quoted relative to external liquid ammonia.

the magnetic field in use and this then defines the absolute frequency at 0.0 ppm for the nuclide in question. The 'master' reference sample for these measurements is 1% TMS in $CDCl_3$, which is used to determine the absolute frequency of TMS for the instrument in use. The Ξ values for selected nuclei are also summarised in Table 3.3, whilst more extensive tables are available [19, 20].

3.3.3. Tubes and sample volumes

Newcomers to the world of practical NMR often find the cost of NMR tubes surprisingly high. The prices reflect the need to produce tubes that conform to strict requirements of straightness (camber) and concentricity, as deviation from these can produce undesirable artefacts in spectra, usually in the form of 'spinning-sidebands'. Generally, the higher quality (and more expensive) tubes are required on higher field instruments, and this is particularly so for proton work. Some experimenting may be required to decide on a suitable balance between cost and performance for your particular instruments.

The diameter of the tube is generally determined by the dimensions of the probe to be used, or more precisely the probe's rf coils. Tube diameters employed in chemical laboratories (Fig. 3.44) include 1, 1.7, 2.5 or 3 mm (for use with the so-called microprobes), 5 mm (the most widely used) or 10 mm (typically used for the observation of low sensitivity nuclei where the solubility of the material to be studied is limiting or for dedicated polymer studies). Eight millimetre tubes have found use in biological macromolecular work, where solubility and aggregation considerations preclude the use of concentrated samples. The benefit of the smaller diameter tubes and probes lies in their greater 'mass sensitivity' as will be explained in Section 3.4.2, which describes available probe technologies. As sensitivity is usually of prime importance when performing any NMR experiment, it is

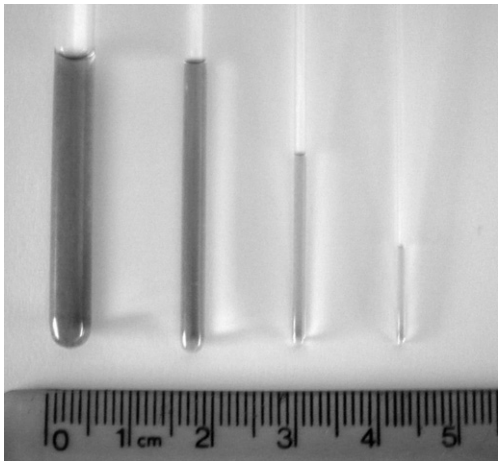


Figure 3.44. NMR tubes of various diameters and their optimum solution volumes (from left to right): 5 mm (500 µL), 3 mm (150 µL), 1.7 mm (30 µL) and 1 mm (5 µL).

Table 3.4. Sample tube diameters and their associated working sample volumes and heights

Tube dimensions (outer diameter)	Solute volume/μL	Solute column height/mm
10	2500	40
8	1500	40
5	500	40
3	150	40
1.7	30	20
1	5	10

prudent to use the correct sample volume for the probe/tube configuration and not to dilute the sample any more than is necessary. An important factor here is the length of the detection coil within the probe. The sample volume should be sufficient to leave a little solution above and below the coil since magnetic susceptibility differences at the solution/air and solution/glass interfaces lead to local distortions of the magnetic field. For a standard 5 mm tube, the coil length is typically 16 mm and the required volume will be around 500–600 μL whereas for the narrower coils and tubes this may reduce to 100–150 μL for the 3 mm tube, 25–30 μL for the 1.7 mm tube or as little as 5–8 μL for the 1 mm microtube (Table 3.4). For 10 mm tubes, it may be necessary to use a 'vortex-suppresser' to prevent whirlpool formation in the tube if the sample is spun. These are usually polytetrafluoroethylene plugs designed to fit tightly within the NMR tube, which are pushed down to sit on the surface of the sample and thus hold the solution in place. Some care is required when using these plugs in variable temperature work due to the thermal expansion or contraction of the plug, and at low temperatures, the plug may even fall into the solution so its use may be inappropriate. Filling the smaller 1 and 1.7 mm tubes requires the use of fine needles with a manual syringe or robotic system, or alternatively can be readily filled by placing the solution in the top tube funnel and centrifuging this down.

For the handling of smaller sample quantities when microprobes are not available, the microtubes of small diameter that are designed for use with microprobes may be employed to concentrate the available material within the detection coil. Thus, using a 3 mm tube in a 5 mm probe has been shown to provide a modest signal-to-noise improvement for conventional probes whereas this approach can provide greater gains when using cryogenically cooled probe heads; see Section 3.4.2. Alternatively, microcells can be obtained that fit within standard 5 mm NMR tubes, but which require only tens of microlitres of solution again allowing greater sample concentration. When using such cells, it is advisable to fill the volume around the outside of the cell with solvent, as this minimises susceptibility differences and improves resolution. Additionally, special tubes and plugs ('Shigemi tubes') are available that allow all the sample volumes to be held within the sensitive region of the rf coil (Fig. 3.45). These use glasses that have a magnetic susceptibility matched to that of the solvent and so allow shorter sample heights without the introduction of lineshape defects [21]. Typically, tubes are available individually matched to water and to a small selection of organic solvents and are available in 5 and 3 mm sizes. When using these it is imperative that no air bubbles become trapped between the solvent and the upper plug to avoid lineshape perturbations, and some practise in their use is highly recommended

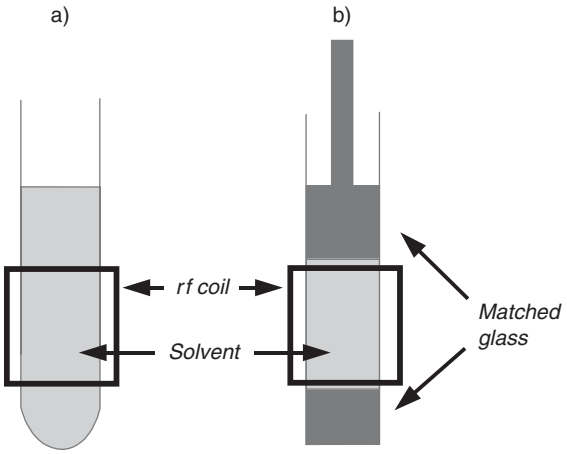


Figure 3.45. Schematic comparison of (a) a conventional NMR tube and (b) a Shigemi matched tube and plug in which the magnetic susceptibility of the glass is matched to that of the solvent. This allows a smaller sample volume and so concentrates more of the analyte within the detection coil.

before embarking on the preparation of your most precious sample. More recently, specialised tubes have been introduced specifically for use with cryogenically cooled probes that have rectangular cross-sections, for example, 3×6 mm (Section 3.4.2). These have been shown to provide enhanced sensitivity when observing samples of high ionic strength (buffered or salty samples) by reducing pick-up of electrical noise arising from the sample solution and so are likely to be of greatest applicability for the study of biomacromolecules or biofluids.

NMR tubes should be kept clean, dry, free from dust and must also be free from scratches on the glass as this distorts the cylinder in which the sample is confined. New tubes may not always clean when delivered, although this is often assumed to be the case. Organic lubricants used in their manufacture may remain on the glass and become all too apparent when the tube is first used. For routine washing of tubes, rinsing with a suitable organic solvent, such as acetone, or with distilled water a few times is usually sufficient. More stubborn soiling is best tackled by chemical rather than mechanical means and either soaking in detergent or strong mineral acids is recommended. The use of chromic acid must be avoided as this can leave traces of paramagnetic chromium(VI), which will degrade resolution. Tubes should not be dried by subjecting them to high temperatures for extended periods of time as this tends to distort them. A better approach is to keep the tubes under vacuum, at slightly elevated temperatures if possible, or to blow filtered nitrogen into the tubes. If oven drying is used, the tubes must be laid on a flat tray and heated for only 30 min, not placed in a beaker or tube rack as this allows them to bend. If the removal of all traces of protiated water from the tube is important, then it is necessary to rinse the tube with D_2O prior to drying to ensure the exchange of water adsorbed onto the surface of the glass. Unused tubes should be stored with their caps on to prevent dust from entering. Finally, it is also important to avoid contamination of the *outside* of the NMR tube (which includes fingerprint oils) as this leads to the transfer of the contaminants into the probe head. The accumulation of these contributes to degradation of the instrument's performance over time.

3.3.4. Filtering and degassing

For any NMR experiment, it is necessary to achieve a uniform magnetic field throughout the whole of the sample to obtain a high-resolution spectrum, and it is easy to imagine a range of circumstances that detract from this condition. One of the most detrimental is the presence of particulates in the sample since these can distort of the local magnetic field, so it is of utmost importance to remove these from an NMR sample prior to analysis. Even very slight changes in the field within the sample can produce a noticeable reduction in resolution. This should not come as a surprise if one recalls that one often wishes to resolve lines whose resonant frequencies are many hundreds of megahertz but which are separated by less than 1 Hz. This corresponds to a resolution of 1 part in 10^9, roughly equivalent to measuring the distance from the Earth to the Sun to an accuracy of 1 km! Commercial sintered filters can be obtained but require constant washing between samples. A convenient and cheap alternative is to pass the solution through a small cotton wool plug in a Pasteur pipette directly into the NMR tube (it is good practise to rinse the cotton wool with a little of the NMR solvent before this). If cotton wool is not suitable for the sample, a glass wool plug may be used although this does not provide such a fine mesh. If metal ions are likely to be present in solution (paramagnetic ions being the most unwelcome), either from preparation of the sample or from decomposition of the sample itself, then passing the solution through a small column of a metal chelating resin should remove these.

In situations where very high resolution is demanded or if relaxation studies are to be performed, then it may also be necessary to remove all traces of oxygen from the solution. This may include the measurement of nuclear Overhauser enhancements, although for routine NOE measurements, it is usually not necessary to go to the trouble of degassing the sample (see Chapter 8). The need to remove oxygen arises because O_2 is paramagnetic and its presence provides an efficient relaxation pathway that leads to line broadening. Essentially, there are two approaches to degassing a sample. The first involves bubbling an inert gas through the solution to displace oxygen. This is usually oxygen-free nitrogen or argon, which need be passed through the solution in the NMR tube for about 1 minute for organic solvents and double this for aqueous solutions. Great care is required when attempting this, and a fine capillary must be used to introduce the gas *slowly*. It is all too easy to blow the sample clean out of the tube in an instant by introducing gas too quickly, so it is probably wise to get some feel for this before attempting the procedure on your most precious sample. Note that with volatile solvents a significant loss in volume is likely to occur and TMS, if used, may well also be lost.

The second and more thorough approach is the 'freeze-pump-thaw' technique. The NMR tube containing the solution is frozen with liquid nitrogen or dry ice and placed under vacuum on a suitable vacuum line. Commercial tube manufacturers produce a variety of specialised tubes and adapters for this purpose, but a simple alternative is to connect a standard tube to a vacuum line via a needle through a rubber septum cap on the tube. The tube is then isolated from the vacuum by means of a stopcock and allowed to thaw, during which the dissolved gases leave the solution. The procedure is then repeated typically at least twice more, this usually being sufficient to fully degas the sample. When using an ordinary vacuum pump, it may be necessary to place a liquid nitrogen trap between the pump and the sample to avoid the possibility of vacuum oils condensing in the sample tube. Furthermore, when freezing aqueous samples, it is easy to crack the tube if this is carried out too fast; holding the sample tube just above the freezing medium whilst tilting it is usually sufficient to avoid such disasters. An alternative approach for aqueous samples containing involatile solutes is to carefully place the sample directly under vacuum without freezing, for example, in a vacuum jar or schlenk, and allowing dissolved gas to bubble out of solution.

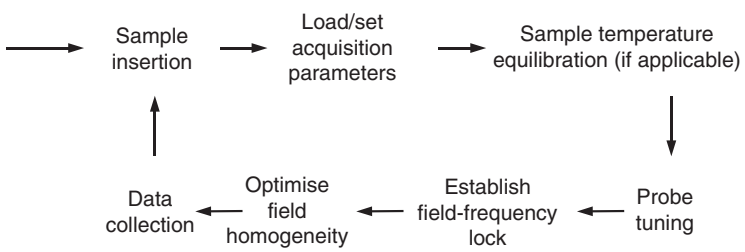
Following the degassing procedure, the tube should be sealed. If the sample is likely to be subject to short-term analysis, say over a few hours, then a standard tight-fitting NMR cap wrapped with a *small* amount of paraffin wax film or, better still, use of a rubber septum, is usually sufficient since diffusion into solution of gases in the tube will be rather slow. For longer duration studies, specialised adapters or screw-top tubes can be purchased or standard tubes can be flame-sealed, for which NMR tubes with restrictions towards the top are available to make the whole process very much easier. When flame sealing, it is highly advisable to cool or freeze the solution and then seal whilst pulling a gentle vacuum. Failure to cool the solution or evacuate the tube may cause the pressure within it to become dangerously high as the sample warms. The seal should be symmetric, otherwise the sample will spin poorly, so practice on discarded tubes is likely to pay dividends.

3.4. PREPARING THE SPECTROMETER

Whenever a new sample is placed within the NMR spectrometer, the instrument must be optimised for this. The precise nature of the adjustments required and the amount of time spent making these will depend on the sample, the spectrometer and the nature of the experiment, but in all cases, the aim will be to achieve optimum resolution and sensitivity and to ensure system reproducibility. The details of the approaches required to make these adjustments depend on the design of spectrometer in question, so no attempt to describe such detail is made here. They all share the same general procedures however, which are summarised in Scheme 3.1 and described below.

3.4.1. The probe

If you are a chemist making use of the NMR facilities available to you, then choosing the appropriate probe for the study in question may not be relevant as you will probably be forced to use that which is available. If, however, you are involved with instrument purchasing or upgrading or you are fortunate enough to have available a variety of probes for a given instrument, then it is important to be able to make the appropriate selection. Since it is the probe that must receive the very weak NMR signals, it is perhaps the most critical part of the NMR spectrometer, and its particular design and construction will influence not only the types of experiments it is able to perform but also its overall performance. There has been considerable progress over the years in the design of probes, resulting in ever-increasing performance and a greater array of available probe technologies, and these will be reviewed in more detail in Section 3.4.2 below.



Scheme 3.1. The typical procedure followed in preparing a spectrometer for data collection. All these steps can be automated, although probe tuning demands additional equipment or suitably equipped auto-tuneable probe heads.

Important considerations when selecting a probe are the frequency range(s) it may be tuned to, since this will dictate the nuclei one is able to observe, and how the rf coil configuration has been optimised. The simplest design is a probe containing a single coil, which is designed for the observation of only one nucleus. In fact, it would be 'doubly-tuned' to enable the simultaneous observation of deuterium for the field-frequency regulation via the lock system, although the presence of a deuterium channel is usually implicit when discussing probe configurations. However, many modern NMR experiments require pulses to be applied to two (or more) different nuclides, of which one is most often proton, for which two rf coils are necessary. In this case, two possible configurations are in widespread use depending on whether one wishes to observe the proton or another nuclide, referred to as the X-nucleus. The traditional two-coil design is optimised for the observation of the X-nucleus, with the X-coil as the inner of the two allowing it to sit closer to the sample, so offering the best possible sensitivity for X-observation; it is said to have the greatest 'filling factor'. This configuration can be described by the shorthand notation X{^{1}H}. Nowadays, multipulse experiments tend to utilise the higher sensitivity offered by proton observation wherever possible and benefit from probes in which the proton coil sits closest to the sample with the X-coil now the outer most; ^{1}H{X}. It is this design of probe that is widely referred to as having the *inverse* configuration because of this switch in geometry. In either case, the X-coil circuitry can be designed to operate at only a single frequency or can be tuneable over a wide frequency range, such probes being known as *broadband observe* or *broadband inverse* probes. An alternative popular configuration in organic chemistry is the *quad-nucleus* probe that allows observation of the four most commonly encountered nuclei ^{1}H, ^{13}C, ^{19}F, and ^{31}P. In studies of biological macromolecules, and less frequently in organic chemistry, *triple-resonance* probes are employed, allowing proton observation and pulsing or decoupling of two other nuclei; ^{1}H{X, Y}, the most common configurations being ^{1}H{^{13}C, ^{15}N} or ^{1}H{^{13}C, ^{31}P}. Alternatively, triple-resonance inverse probes that include one BB channel may offer greater flexibility in a chemical laboratory where many nuclei may be of interest with ^{1}H{^{13}C, BB} or ^{1}H{^{31}P, BB} being most common.

A further feature offered is the addition of magnetic field gradient coils to the probe head. These surround the usual rf coils and are designed to destroy the static magnetic field homogeneity throughout the sample for short periods of time in a very reproducible manner; gradient-selected techniques are introduced in Chapter 5. Whilst offered as optional features, the inclusion of a single-gradient coil in a probe can now be considered as standard.

Flow probes

A fundamentally different approach to sample handling is used in the case of flow probes in which a flow cell sits within the rf coils, and solutions are pumped into the cell for analysis, dispensing with the need for NMR tubes altogether. Traditional designs held a fixed cell assembly meaning these probes could only be used in the flow mode, whereas newer versions of both conventional and cryogenic probes (Section 3.4.2) utilise an interchangeable flow cell that allows this to be placed within probes designed to accommodate conventional tubes, so providing greater flexibility. Flow probes are typically used either as the NMR detector at the output of hyphenated systems including HPLC for mixture analysis (as LC-NMR or LC-NMR-MS configurations), or in systems where direct flow injection is employed for automated sample handling often of very large sample numbers, a notable example being the screening analysis of biofluids. NMR analyses may then be performed in an *on-flow mode* wherein the sample moves continually through the cell, so limiting observation to 1D ^{1}H (or ^{19}F) spectra only due to the short sample residence time. This mode is typically used for front-line screening of mixtures and has limited applicability to full structural characterisation. Alternative static modes include the *stopped-flow mode* in which the analyte slug is held in the cell for a discrete period by stopping sample elution, so enabling the collection of longer running and 2D NMR spectra for structure identification, or *loop-collection mode* where the analyte slug is diverted to a storage loop during a chromatographic run and later parked within the flow cell for NMR analyses. Some of the main issues that arise with flow-based NMR when compared with tube-based NMR include the potential for cross-contamination between samples, requiring appropriate cleaning cycles between these, sample dilution during the transfer process and the fact that samples are not isolated in convenient storage vessels (tubes) should further analysis be required. Hyphenated methods for on-line separation also tend to be time consuming and technically demanding, and procedures employing off-line chromatography followed by tube-based NMR analysis still appear to be widely employed. Developments such as solid-phase extraction (SPE) for sample concentration combined with high sensitivity microscale or cryogenic probes also provide competitive routes to the analysis of complex mixtures

[22]. Indeed, the 30 µL volume of a 1.7 mm NMR tube matches the elution volume of SPE cartridges, and the most recent 1.7 mm cryogenic probe design (see below) was no doubt developed with this in mind.

LC-NMR coupling technology has been recently reviewed [23], and the interested reader is also referred to the detailed text edited by Albert [24]. The miniaturisation of flow probes to the capillary scale has been developed to enhance signal detection sensitivity, as discussed in Section 3.4.2, and more recently the hyphenation of gas chromatography with microflow cells (GC-NMR) has even been demonstrated [25].

3.4.2. Probe design and sensitivity

There now exists an ever-increasing array of probe designs and dimensions aimed at delivering improved and optimised detection sensitivity for the analytes of interest, and this section sets out some of these key developments and the rationale underlying them. Over the years, a number of factors have been employed that have a progressively improved probe performance and sensitivity in addition to novel probe designs. One such development has been the material used in the construction of the receiver coil itself. Since this coil sits in very close proximity to the sample, this material may distort the magnetic field within this, compromising homogeneity. Modern composite metals are designed so that they do not lead to distortions of the field (they are said to have *zero magnetic susceptibility*), so allowing better lineshapes to be obtained, which ultimately leads to improved signal-to-noise figures. Increasingly, coil materials are also matched to the properties of the cooling gas that flows over the NMR tube, typically either air or pure nitrogen meaning it may be advantageous to define this when ordering probe heads. A second factor lies in the coil dimensions, with modern coils on standard probes tending to be longer than used previously so that a greater sample volume sits within them. This demands a greater volume in which the magnetic field is uniform, so these changes in probe construction have largely followed improvements in room temperature shim systems (see below). Developments in electronics have also contributed to enhanced instrument sensitivity.

Detection sensitivity

In order to appreciate the developments in probe and instrument design, we shall first consider some key factors that contribute to the detection sensitivity of an NMR measurement and hence how this may be improved. As an illustration of the progressive enhancements in instrument sensitivity, Fig. 3.46 shows the specified signal-to-noise ratios of commercial ^{1}H observe probes at the time of new magnet launch (as judged by the 0.1% ethyl benzene 'sensitivity test', Section 3.6.2). Although the data reflect in part increasing performance with higher magnetic fields, the improvement in overall instrument and probe performance that paralleled these developments is illustrated by the comparison of 500 MHz data at the time of launch (1979) with that in 2007, which shows an approximately fivefold improvement in *S/N* (dashed line in Fig. 3.46). The current *S/N* specifications for various field strengths (represented by ^{1}H frequency) are also shown for the most recent

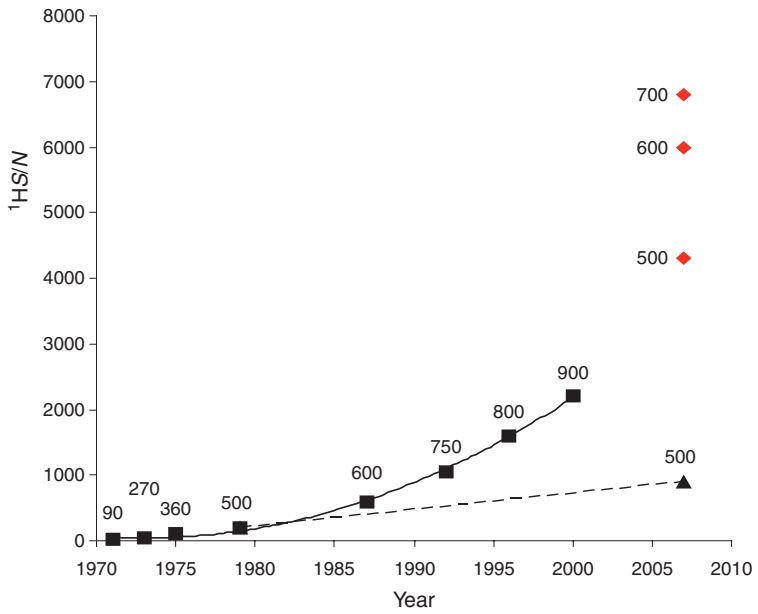


Figure 3.46. The specified ^{1}H signal-to-noise ratios for ^{1}H observe probes at the time of magnet launch (indicated by ^{1}H frequency) as a function of year (squares) and for the more recent cryogenically cooled probes (red diamonds). The current data for a conventional 500 MHz probe is shown by the triangle (data courtesy of Bruker Biospin in 2007).

(2007) cryogenically cooled probes (described below). These data demonstrate that a quarter of century of combined instrument development has led to a tenfold gain in field strength and ~400-fold enhancement of ^{1}H signal-to-noise. Nevertheless, sensitivity limitations can still present a barrier when compared against other spectrometric methods, and experimental technologies for enhancing this further continue to represent an active research area; see Section 10.8.2 for example.

The signal-to-noise of an NMR measurement depends on many factors, but improving this ultimately comes down to either boosting the signal intensity or reducing the background noise. A general expression for this parameter for a conventional NMR spectrometer operating at ambient temperature may be formulated thus:

$$\frac{S}{N} \propto NAT_s^{-1}B_0^{3/2}\gamma^{5/2}T_2^*(\mathrm{NS})^{1/2} \tag{3.9}$$

where N is the number of molecules in the observed sample volume, A is a term that represents the abundance of the nuclide, T_S is the temperature of the sample and surrounding rf coil, B_0 is the static magnetic field, γ represents the magnetogyric ratio of the nuclide, T_2^* is the effective transverse relaxation time and NS is the total number of accumulated scans. Many of these factors are dictated by the properties of the nuclide involved, including the natural abundance, the magnetogyric ratio and relaxation behaviour so are not dictated by instrument design and are considered further in Chapter 4 where pulse techniques that exploit these parameters are introduced. Suffice it to say that the high magnetogyric ratio, near 100% natural abundance and favourable relaxation properties of the proton, explains its popularity in high-resolution NMR spectroscopy.

The number of observable molecules N is, of course, dictated by the amount of material available but also by how much of this is held within the observe region of the detection coil, the so-called *active volume* (or *observe volume*) of the coil (Fig. 3.47). Any solution that sits outside this region will not induce a response in the coil and so does not contribute to the NMR signal but may be required to avoid lineshape distortions arising from magnetic susceptibility discontinuities too close to the coil. The susceptibility matched ('Shigemi') tubes introduced in Section 3.3.3 enhance sensitivity by concentrating more of the sample within the active volume and represent a cost-effective approach to improving signal-to-noise.

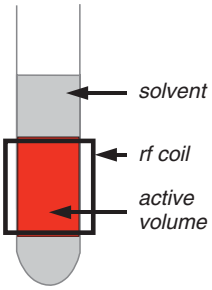


Figure 3.47. The active (observe) volume of an NMR coil is the region of the sample that sits within the coil and so contributes to the detected response.

From an instrumental point of view, sensitivity scales as $B_0^{3/2}$ [2], and an increase in field strength has traditionally been one of the drivers of enhanced sensitivity, although is becoming more difficult to justify on these grounds alone given the enormous costs associated with the very highest field magnets. Increasingly, advances in probe design are now employed to provide substantial sensitivity gains at more affordable prices, as illustrated below. In addition, improved electronics have reduced system noise, especially that associated with the NMR signal preamplifier stage and digital data handling in the form of oversampling reduces so-called quantisation noise in spectra (Section 3.2.6) adding further enhancement. These advances combined with the aforementioned refinements in probe technology have themselves lead to progressive increases in signal-to-noise figures for conventional 5 mm probes, as illustrated by the data for 500 MHz probes in Fig. 3.46.

Mass and concentration sensitivity

When discussing sensitivity in the context of real laboratory samples as opposed to instrument testing, it becomes useful to consider two definitions of this. The classic measurement of instrument sensitivity uses a fixed solution concentration (0.1% ethyl benzene in $CDCl_3$, equivalent to 14 mM) under standard conditions and so measures the *concentration sensitivity* of the system.

$$S_c = \frac{\text{signal-to-noise}}{\text{solution concentration}}$$

Concentration sensitivity may be improved by using greater solution volumes within the active volume of the coil. Thus, a wider or longer detection coil will hold more sample and show an enhanced concentration sensitivity, and it follows that the use of wider sample tubes and probes is advantageous when sample solubility is a limiting factor. The use of longer coils in 5 mm probes has contributed to enhanced signal-to-noise performance specifications for certain probe designs. However, such gains may not be realised for *real*

[2] A more thorough theoretical analysis shows signal-to-noise should scale as $B_0^{7/4}$ [26].

Table 3.5. Comparative ^{1}H mass sensitivities for various probe configurations discussed in the text (data for entries 1–6 courtesy of Bruker Biospin)

Probe diameter (inverse configuration)	Sample volume (µl)	Relative mass sensitivity
5 mm	500	1
3 mm	150	1.5
1.7 mm	30	2
1 mm	5	4
5 mm Cryogenic	500	4
1.7 mm Cryogenic	30	14
Capillary micro-flow	5	10

mass-limited laboratory samples. Such samples are typically not presented as fixed solution concentrations but are more commonly available as materials of fixed mass and can thus be prepared in any chosen sample volume, to the limit of sample solubility. In such cases, it is more useful to refer to the *mass sensitivity* of the probe or system, a measure of signal-to-noise as a function of sample mass or the number of moles of material.

$$S_{\mathrm{M}} = \frac{\text{signal-to-noise}}{\text{moles of material}}$$

A probe of higher mass sensitivity will provide a greater signal-to-noise ratio in a spectrum for a fixed amount of material, and this better reflects the intrinsic detection sensitivity of the probe design. When making comparisons between the performances of different probes, it is common practice to relate the results to those of a 'conventional' 5 mm probe operating at the same field strength, and this approach will be employed in the discussions that follow. The relative mass sensitivities of some commercially available probes are summarised in Table 3.5. Note that these are conservative figures that the manufacturer expects always to meet, and it is not uncommon to exceed these but they nevertheless provide us with an approximate guide to the sensitivity gains afforded by the differing probe configurations.

Microprobes

The mass sensitivity of an rf coil scales inversely with the diameter of the coil, d ($S/N \propto 1/d$), to first approximation, meaning that greater intrinsic sensitivity can be achieved by narrowing the coil. This is the basis for the development of the so-called *microprobes*. This terminology is used somewhat loosely and there exists no formal definition of a microscale probe, but it is generally taken to mean a probe geometry that is small when compared to a 'standard' probe. Nowadays, the standard remains a probe designed to accept a 5 mm diameter maximum NMR tube, colloquially referred to as a *5 mm probe*, so commercial microprobes would now include the 3.0, 1.7 and 1.0 mm probes.

The mass sensitivity benefits of a microprobe relative to a larger diameter probe can be exploited for a fixed mass of sample through the enhanced signal-to-noise achievable or through reduced data collection times for data sets of equivalent signal-to-noise ratio, or perhaps as a balance between these. This of course demands that the solute can be dissolved in the reduced sample volumes employed; see below. As a consequence, such probes enable data to be collected on reduced sample masses and so extend the operating range of NMR experiments for mass-limited samples. For example, a microprobe that has a fivefold gain in mass sensitivity relative to a larger probe will provide a fivefold improvement in signal-to-noise ratio, which may be traded for a 25-fold reduction in data collection time for a fixed sample mass. Alternatively, this would allow sample masses to be reduced by a factor of 5 for data of comparable quality to that from the larger probe collected in a similar timeframe. An additional benefit from the reduced solvent volumes of the microprobes is the reduction in background solvent resonances and potentially also those of solvent impurities that can become apparent when dealing with very small sample masses. The smaller sample volumes also mean solvent suppression becomes much easier. The microprobes also provide reduced pulse widths leading to more uniform excitation over wider bandwidths. Indeed, it is generally the case that the length of a 90° pulse width for a given pulse power provides a direct indication of the detection sensitivity of the coil with smaller pulse widths correlating with higher sensitivity, and this itself can be useful way of gauging a probe's performance. They may also demand lower pulse powers to prevent damage to the more delicate rf coils, so even greater caution is required with such probes.

The availability of commercial tube-based microprobes came about in 1992 with the introduction of the 3 mm inverse probe, although much earlier work had demonstrated the potential for microscale probes [27]. The 3 mm probe with a 150 µL sample volume demonstrated an approximately twofold gain in signal-to-noise in ^{1}H-detected heteronuclear correlation spectra relative to a 5 mm probe with the same sample mass in 600 µL [28]. Subsequent to this in 1999 [29], a 1.7 mm probe [30] utilising a 30 µL sample demonstrated a further approximately twofold sensitivity gain in 2D correlation experiments relative to the 3 mm probe, and most recently in 2002 [31], a 1 mm probe using only a 5 µL sample volume demonstrated a fivefold mass sensitivity gain in ^{1}H observation relative to a 5 mm probe. Owing to limitations in sample handling and the delicate nature of 1 mm tubes, it seems likely that this will represent the smallest size for tube-based NMR probes, although smaller dimensions are feasible for microflow probes as described below.

Figure 3.48 compares the performance capabilities for ^{1}H observation of 5 and 1 mm triple-resonance inverse probes. Data were collected under identical conditions (single scan, 90° pulse excitation) on the same 500 MHz spectrometer for a 50 µg sample of sucrose in D_2O. The relative signal-to-noise of the anomeric resonance (not shown) is 7.8:1 and 39.5:1 for the 5 and 1 mm probes respectively, indicating a fivefold sensitivity gain for the smaller diameter probe. The reduced solvent volume also attenuates the solvent background, and this can also be a significant factor in the selection of smaller diameter tubes for mass-limited samples. Despite the miniscule samples and hence the small volume over which good field homogeneity is required, experience suggests the 1 mm probe resolution tends to be a little poorer than that of the larger geometry probes, and the vendor probe specifications also support this observation. This most probably is due to the proportionately larger volume of glass present from the NMR tube relative to the solution within the coil and with the associated susceptibility discontinuities; susceptibility matched 1 mm NMR tubes may yield still greater performance but are not currently available.

One further approach to miniaturisation in the Varian NanoProbe^{TM} has been the use of small rotors of 40 µL total volume inside the specially constructed probe head that spins the sample at the 'magic angle' (54.7° from the axis of the B_0 field) at several kilohertz to eliminate lineshape distortions that would arise from sample discontinuities close to or within the rf coil (magic angle spinning is described in Section 10.7). This allows the whole sample volume to be retained within the coil's active volume, unlike tube-based approaches where up to 50% of the sample may sit outside of this volume, and so improves sensitivity. The use of MAS also makes the probe suitable for the analysis of heterogeneous samples such as those attached to resins from combinatorial synthesis [32], but it lacks the convenience of tube-based sample handling and appears to be less widely adopted for work with mass-limited samples.

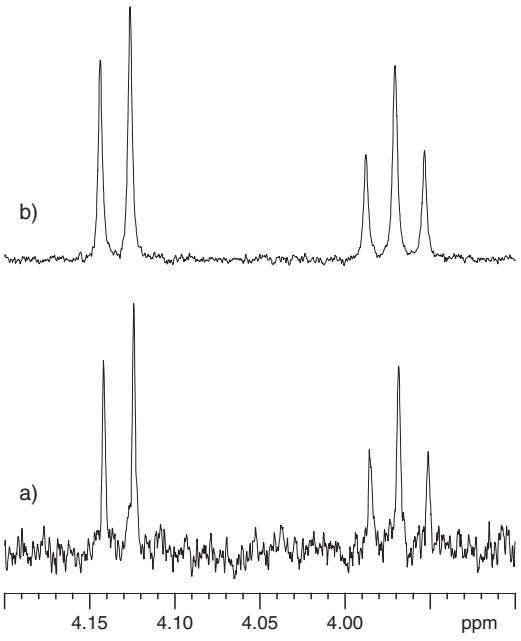


Figure 3.48. Partial ^{1}H spectra of 50 µg sucrose in D_2O recorded with probes of different dimensions, (a) 500 µL in 5 mm probe and (b) 5 µL in 1 mm probe. Spectrum (b) displays a fivefold gain in signal-to-noise relative to spectrum (a).

Microflow probes

An alternative approach to reducing sample volume requirements in probes is avoiding the use of sample tubes altogether and to use flow technology to place the sample within a fixed coil geometry. In the context of probe miniaturisation, this is achieved using capillary flow cells ('capillary NMR') in which the rf coil is wound directly around the capillary tube itself. Such probes have been developed, which have active volumes as little as 5 nL [33] although commercially available probes currently utilise active volumes of 2.5 µL that require working sample volumes of 5 µL and have a total *probe* volume of a mere 15 µL [34]. These probes employ rf coils that are surrounded by an inert perfluorinated fluid whose magnetic susceptibility matches that of the coil assembly, as needed to produce acceptable lineshapes. As there is no need to align the coil with the field axis for sample insertion, as is the case for tube-based designs, this allows the use of solenoid coils mounted perpendicular to the field, which give an intrinsic sensitivity advantage of approximately twofold over the traditional, vertically aligned saddle coil designs [35]. In accord with this, the data in Table 3.5 indicate the capillary probe to demonstrate a ~10-fold mass sensitivity gain over a conventional 5 mm probe (or approximately 5-fold if Shigemi tubes are employed [36]). Furthermore, it is possible to tune the single coil to multiple frequencies simultaneously, making it possible to use the coil for both ^{1}H and ^{13}C observation without any loss of filling factor.

The flow cell design naturally requires a different approach to sample handling [37] with sample insertion performed via manual syringe injection, robotic injection or via direct hyphenation of the probe with capillary liquid chromatography (CapLC). The narrow capillaries (<100 µm diameter) may be prone to blockage, and effective sample filtration and the routine use of in-line filters (2 µm) are essential. Arguably, the flow approach lacks the convenience of NMR tubes, which are effective storage vessels should one wish to undertake further analysis of a sample without further preparative steps. The practicalities of employing the flow probe in conjunction with CapLC has also been considered and contrasted with off-line sample separation coupled with tube-based analysis with a cryogenic probe [22].

Cryogenic probes [38, 39]

A fundamentally different approach to improving detection sensitivity is to reduce the background noise in spectra, and the most recent probe designs aim to do this by cooling both the rf detection coils and the preamplifier for optimum results. This is achieved through the use of a cold helium gas closed-cycle system, which cools the rf coils to typically 25 K and the preamplifier (which is housed within the body of the probe rather than being separate from this as with conventional probes) to around 70 K. This concept was proposed [26] and demonstrated experimentally [40, 41] some years before the first commercial systems became available (in 1999), a testament to the demanding technical challenges in the development of these systems, not least of which is the need to maintain the detection coil at around 25 K whilst the adjacent sample, a matter of some millimetres away, remains at ambient temperature. The result of cooling the rf coils leads to an approximately twofold gain in signal-to-noise whilst cooling of the preamplifier leads to a similar gain, resulting typically in a fourfold signal-to-noise gain relative to a conventional probe of similar dimensions[3]. The cooling of the rf coils is achieved through conduction from a cryo-cooled block on which the coils are mounted rather than by direct cooling of these by the flowing He gas so as to avoid disturbance of the coils.

The gains arising from cryogenic cooling of probes may be described more formally by considering the noise contributions arising within the hardware. This may be summarised

$$S/N \propto \frac{1}{\sqrt{T_{\mathrm{c}}\mathrm{R}_{\mathrm{c}} + T_{\mathrm{a}}(R_{\mathrm{c}} + R_{\mathrm{s}}) + T_{\mathrm{s}}R_{\mathrm{s}}}} \tag{3.10}$$

where T_{c} and R_{c} represent the temperature and resistance of the coil, T_{s} is the sample temperature, R_{s} is the resistance generated in the coil by the sample itself (the 'sample resistance') and T_{a} is the effective noise temperature of the preamplifier. By cooling both the coils and the preamplifier, the first two terms of the denominator in Eq. (3.10) are reduced and so contribute to the higher sensitivity of cryogenically cooled probes. The third term, representing noise generated by the sample, becomes important for ionic solutions [42], as discussed below, and is of greater significance for cryogenic probes than conventional probes because of the smaller contributions of the first two terms. The comparison of

[3] The dewar assembly required for the cryoprobes actually results in a reduced filling factor that places limitations on the sensitivity improvement achievable through probe cooling; for example, a 5 mm cryogenic probe typically has the coil dimensions of a traditional 8 mm probe.

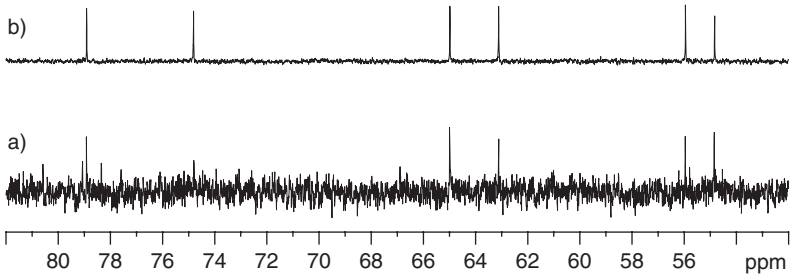


Figure 3.49. A comparison of the carbon-13 sensitivity performance of (a) a conventional room temperature probe and (b) a cryogenic probe operating at 500 MHz (^{1}H). Data were collected on the same sample under identical experimental conditions using a broadband (BB) observe probe tuned to carbon-13 and a cryogenic carbon-13 observe probe respectively.

cryogenic and conventional probes in Table 3.5 demonstrates sensitivity gains of (at least) 4, equating to a minimum 16-fold time-saving for a given sample mass or the ability to collect data on fourfold less material in a given data acquisition period. This advantage is demonstrated experimentally in Fig. 3.49 for a ^{13}C-optimised cryogenic probe for which there is observed a sixfold gain in signal-to-noise ratio relative to a room temperature probe operating under otherwise identical experimental conditions. Whilst the presence of some resonances is dubious in Fig. 3.49 trace (a), all are clearly revealed in trace (b). As the performance gain is not dependent on increased sample concentration, as is the case with microprobes, there is a simultaneous gain in both concentration and mass sensitivity with this technology. Further enhancements of mass sensitivity may thus be achieved by combining miniaturisation with cryogenic technology in the form of microcryogenic probes, as is illustrated by the figures for the 1.7 mm cryogenic probe.

It is also possible to manipulate the influence of sample noise to enhance sensitivity. Consider again the sample resistance, R_s, which may be represented in the form

$$R_s \propto \varpi_0^2 l r^4 \sigma \tag{3.11}$$

where ω_0 is the observation frequency, l and r are the sample length and radius respectively and σ represents a term for the dielectric properties of the solution. This suggests that the noise contribution from the solution may be reduced by employing tubes of smaller diameter, the most common approach being a 3 mm tube used in a 5 mm probe. Whilst this approach is valid in principle for both conventional and cryogenic probes, the advantages are more pronounced for the cooled probes because the noise contribution from the sample relative to that of the instrument is more significant, as described above. Despite the associated reduction in filling factor, the measure of how much of the rf coil volume is filled with sample, a *S/N* gain of ~1.5-fold has been demonstrated using this approach for a 3 mm *vs* a 5 mm tube containing identical sample masses when acquired in a 5 mm inverse cryogenic probe, [39] equating to a potential time-saving factor of ~2.2. Personal experience suggests that comparable results to the 3 mm tubes are obtained for identical sample masses in both 2 and 1 mm diameter tubes, suggesting further possible gains in noise reduction are offset by the losses in filling factor. Greater gains may be anticipated, however, by the use of 3 mm matched 'Shigemi' tubes. For instances where sample concentration cannot be increased, this method is clearly inappropriate. It has been demonstrated that the contribution from sample noise may also be diminished through the use of NMR tubes with rectangular instead of circular cross-sections with the longer transverse axis of the tube aligned along the rf magnetic field [43]. Such tubes are now commercially available with typical dimensions of 6 × 3 mm and demonstrate a ~20% gain in *S/N vs* a 5 mm tube. Such gains are most pronounced for samples of high ionic strength, which compromise the performance of cryogenic probes because of the associated increase in the R_s term in Eq. (3.10). The most common examples of this are the buffered aqueous solutions widely employed in biomolecular studies. Whilst higher salt concentrations generally lead to decreased performance, the nature of the buffer itself can also have a significant effect, and the behaviour of many common buffers has been assessed in detail in this context [42]. Such adverse effects are generally less pronounced for organic solvents, although losses may still arise for the more polar solvents, especially methanol and dimethylsulfoxide. Finally, note also that Eq. (3.11) predicts that the use of cryogenic probe technology becomes less efficient at higher field strengths and for higher γ nuclides due the to ω_0 dependence of R_s. Thus, despite their high profile for biomolecular studies, the mid-range frequencies (400–600 MHz) routinely employed for organic chemistry combined with the prevalence of organic solvents and the importance of ^{13}C NMR might suggest this to be the home territory of cryogenic probe technology.

Despite the obvious benefits of these high-sensitivity probes, there are also, inevitably, some drawbacks in their use. These include the high initial purchase costs for the complete

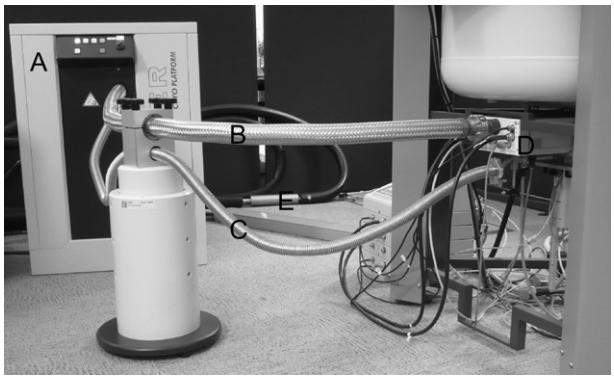


Figure 3.50. A cryogenic probe installation. The cryogenic bay (A) contains the cold expansion head that generates cold helium gas that passes along the insulated transfer line (B) to the probe. The vacuum line (C) maintains the cryogenic probe vacuum and the base of the probe (D) houses the cold preamplifier stages. At the rear can be seen the high pressure lines (E) that carry helium gas to/from the remote compressor.

system, which can be over five times those of a conventional probe, the running and maintenance costs and the need to accommodate the additional instrumentation to operate the cryogenics. A complete system comprises the cold probe, the cryogenic platform or bay that directly supplies the probe with the cooled helium gas and acts as the master controller for the complete assembly, a remote helium gas compressor unit and some means of dissipating the considerable quantities of heat generated by the closed-cycle cooling system (Fig. 3.50). The probe contains the indirectly cooled rf coils and also houses the cooled preamplifiers that are fed with the same cold gas flow. The more recent probe designs cool the preamplifiers associated with both the inner and the outer coils providing enhanced detection sensitivity for multiple nuclei. The internal dewared probe assembly demands continuous pumping of the probe vacuum from the platform and/or the use of internal ion pumps to maintain the cryogenic efficiency. The cryogenic bay generates the cold helium gas through a compression/expansion cycle, producing the characteristic 'chirping' associated with the cryogenic probes. The heat generated by this refrigeration process is removed either through the use of a chilled water supply or by external air-cooled assemblies as used in air-conditioning systems. Selecting the most appropriate method can be a critical aspect of site planning these installations. The water-cooled systems may provide the convenience of local access, but the instrument is then beholden to the demands on the building-wide installation, whilst the external compressor systems require suitable routing either out of the building or into internal space with sufficient cooling capacity. The use of these probes also imposes different operational protocols, chiefly because of the time required to cool and warm the probes that are typically $\sim$4 and $\sim$2 h respectively. This, combined with the more complex nature of the probe and its installation, has led to most cryogenic probe assemblies being considered as permanent fixtures within an instrument and probe removal being reserved for maintenance procedures only.

In summary, it is apparent that there now exist a range of approaches to improving the sensitivity of NMR measurements using modern probe and tube technologies as an alternative to seeking higher field strengths. These often represent more economical and more practical routes to extending the range of samples that can be studied through probe miniaturisation and cryogenic cooling, through the appropriate use of microtubes and susceptibility-matched tubes, or indeed by a combination of these methods. Clearly, the adoption of appropriate probe technologies offers the chemist exciting new opportunities for the investigation of mass-limited samples.

3.4.3. Tuning the probe

The NMR probe is a rather specialised (and expensive) piece of instrumentation whose primary purpose is to hold the transmit and receive coils as close as possible to the sample to enable the detection of the weak NMR signals. For the coils to be able to transmit the rf pulses to the sample and to pick up the NMR signals efficiently, the electrical properties of the coil circuit should be optimised for each sample. The adjustments are made via variable capacitors that sit in the probe head a short distance from the coil(s) and comprise the tuning circuitry. There are two aspects to this optimisation procedure known as *tuning* and *matching*, although the whole process is more usually referred to as 'tuning the probe'. The first of these, as the name implies, tunes the coil to the rf of the relevant nucleus and is analogous to the tuning of a radio receiver to the desired radio station. A poorly tuned probe will lead to a severe degradation in sensitivity, just as a radio broadcast becomes swamped with hissing noise. The second aspect aims to equalise (or match) the impedance (the total effective resistance to alternating current) of the coil/sample combination with that of the transmitter and receiver so that the maximum possible rf energy can pass from the

transmitter into the sample and subsequently from the sample into the receiver. As electrical properties differ between samples, the optimum tune and match conditions will also vary and so require checking for each new sample. These differences can be largely attributed to differences between solvents with the most significant changes occurring between low polarity organic solvents and ionic aqueous solutions.

Probe tuning is necessary for a number of reasons. Other than the fundamental requirement for maximising sensitivity, it ensures pulse widths can be kept short, which in turn reduces off-resonance effects and minimises the power required for broadband decoupling. A properly tuned probe is also required if previously calibrated pulse widths are to be reproducible, an essential feature for the successful execution of multipulse experiments.

Tuning and matching

The process of probe tuning involves applying rf to the probe, monitoring the response from it by some suitable means and making adjustments to the capacitors in the head of the probe (via long rods that pass through its base) to achieve the desired response. In the case of broadband probes, which may be tuneable over very wide frequency ranges, the capacitors may even need to be physically exchanged, either by removing them from the probe altogether or by means of a switching mechanism held within the probe. Various procedures exist for monitoring the response of the probe, but the most useful and the one supplied widely on modern instruments uses a frequency sweep back and forth over a narrow region (of typically a few megahertz) about the target frequency during which the probe response is compared to that of an internal 50 Ω reference load (all NMR spectrometers are built to this standard impedance). The display then provides a simultaneous measure of the tune and match errors (Fig. 3.51a), allowing one to make interactive adjustments to the tuning capacitors to achieve the desired result (Fig. 3.51b); the response reaches a minimum when the probe is matched to the 50 Ω load. With this form of display, it is possible to see the direction in which adjustments should be made to arrive at the correct tune and match condition. It is important to note that the tune and match controls are not mutually exclusive, and adjustments made to one are likely to alter the other, so a cyclic process of tune, match, tune and so on will be required to reach the optimum.

This process is clearly amenable to automation under computer control, and auto-tune configurations are now widely available through two distinct approaches. The first uses remote motorised units that connect to the probe tuning rods via flexible cables that drive this process and has the advantage of being compatible with standard probes. The second, more sophisticated, approach employs motors built into the base of the probe itself that perform the optimisation but has the limitation of being incompatible with older probe designs. The fact that the drive units operate within an intense magnetic field also poses rather severe technical challenges that have been overcome by the use of rare earth metals

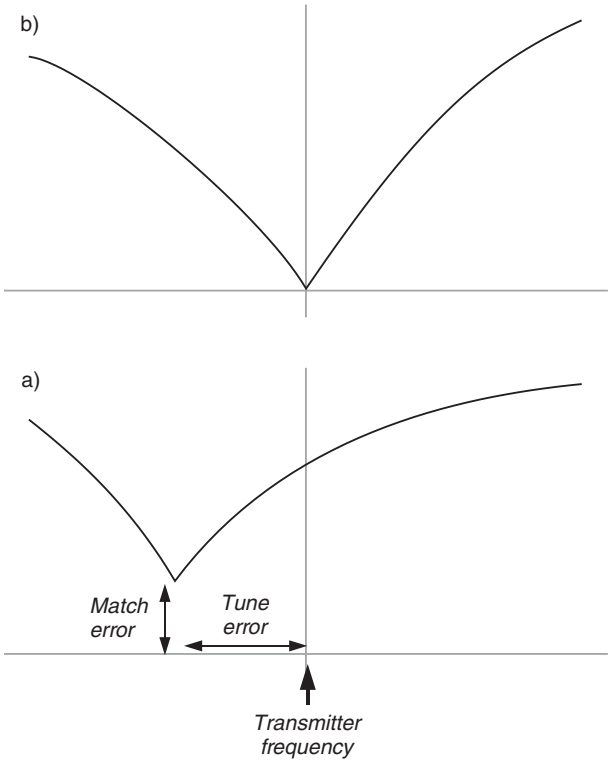


Figure 3.51. Tuning and matching a NMR probe head. The dark line represents the probe response seen for (a) a mis-tuned and (b) a correctly tuned probe head.

in the motorised units. Which ever approach is adopted the benefits of auto-tune probes mean that optimum tuning performance can be obtained under fully automated operation, a wide range of nuclei likewise become accessible without operator intervention (such as with broadband auto-tuneable probes), and for manually operated instruments, this provides the ease and convenience of not having to reach under magnets to tune, a sometimes awkward and potentially hazardous process.

The method for probe tuning on older spectrometers that are unable to produce the frequency sweep display is to place a *directional coupler* between the transmitter/receiver and the probe and to apply rf as a series of very rapid pulses. The directional coupler provides some form of display, usually a simple meter, which represents the total power being reflected back from the probe. The aim is to *minimise* this response by the tuning and matching process so that the maximum power is able to enter the sample. Unfortunately with this process, unlike the method described above, there is no display showing errors in tune and match separately, and there is no indication of the direction in which changes need be made, one simply has an indication of the overall response of the system. This method is clearly the inferior of the two, but may be the only option available.

When more than one nuclear frequency is of interest, the correct approach is to make adjustments at the *lowest* frequency first and work up to the highest. Tuning will also be influenced by sample and probe temperature and must be checked whenever changes are made. If large temperature changes are required, it is wise to quickly recheck the tuning every 10 or 20° so that one never becomes too far from the optimum; this is especially important if using the directional coupler method. Where the spectrometer is used in an open access environment, where interaction with the spectrometer is kept to a minimum or where the instrument runs automatically, probe tuning for each sample may not be not viable (unless an auto-tuneable probe is installed), in which case it is appropriate to tune the probe for the most frequently used solvent, and accept some degradation in performance for the others. Different solvents may require different pulse width calibrations under these conditions.

Finally in this section, we consider one particular situation in which it is beneficial to deliberately detune a probe. When performing studies in *protonated* water, the linewidth of the solvent resonance is broadened significantly in a well-tuned probe, because of the phenomenon of *radiation damping* [44]. This is where the FID of the solvent decays at an accelerated rate because the relatively high current generated in the coil by the intense NMR signal itself produces a secondary rf field that drives the water magnetisation back to the $+z$ axis at a faster rate than would be expected from natural relaxation processes alone. The rapid decay of the FID in turn results in a broadened water resonance. This only occurs for very intense resonances and has greatest effect when the sample couples efficiently with the coil, that is, when the probe is well tuned. Detuning the probe a little provides a sharper (and weaker) water resonance whose lineshape gives a better indication of the field homogeneity. Retuning the probe is essential for subsequent NMR observations employing solvent suppression schemes.

3.4.4. The field-frequency lock

Despite the impressive field stability provided by superconducting magnets, they still have a tendency to drift significantly over a period of hours, causing NMR resonances to drift in frequency leading to a loss of resolution. To overcome this problem, some measure of this drift is required so that corrections may be applied. On all modern spectrometers, the measurement is provided by monitoring the frequency of the deuterium resonance of the solvent. The deuterium signal is collected by a dedicated ^{2}H observe spectrometer within the instrument that operates in parallel with the principle channels, referred to as the *lock channel* or simply the *lock*.

The lock system

The lock channel regulates the field by monitoring the *dispersion* mode deuterium resonance rather than the absorption mode signal that is usually considered in NMR and aims to maintain the centre of this resonance at a constant frequency (Fig. 3.52). A drift in the magnetic field alters the resonance frequency and therefore produces an error signal that has both magnitude and sign (unlike the absorption mode resonance that has only magnitude). This then controls a feedback system that adjusts the field setting. The dispersion signal also has the advantage of having a rather steep profile, providing the greatest sensitivity to change. Monitoring the deuterium resonance also provides a measure of the magnetic field homogeneity within the sample since only a homogeneous field produces a sharp, intense resonance. In this manner, the lock signal may be used as a guide when optimising (shimming) the magnetic field (Section 3.4.5) where the absorption mode signal

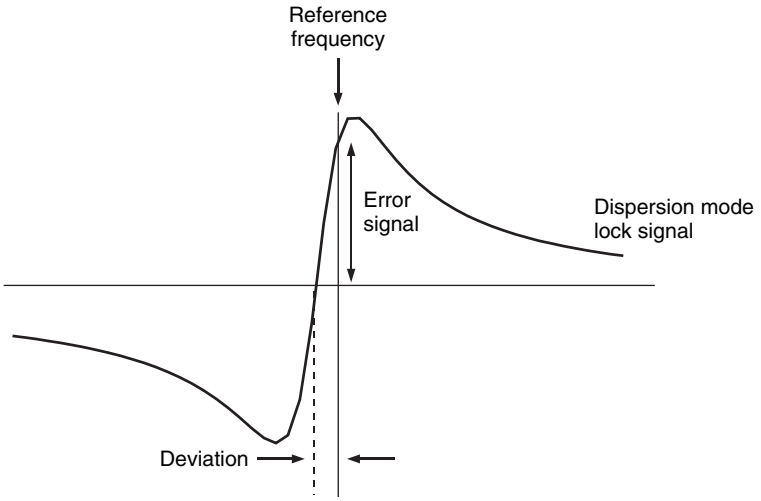


Figure 3.52. The spectrometer lock system monitors the dispersion mode signal of the solvent deuterium resonance. A shift of the resonance frequency due to drift in the static field generates an error signal that indicates the magnitude and direction of the drift, enabling a feedback system to compensate this.

is presented to the operator as the lock display. Since the best field-frequency regulation is provided by the strongest and sharpest lock resonance, the more highly deuterated non-viscous solvents provide the best regulation. Water is especially poor since it tends to have a rather broad resonance because of exchange processes that are very dependent on solvent temperature and pH.

Optimising the lock

The first procedure for locking is to establish the resonance condition for the deuterium signal, which involves altering either the field or the frequency of the lock transmitter. Of these two options, the latter is preferred since it avoids the need for changing transmitter frequencies and is now standard on modern instruments. Beyond this, there are three fundamental probe-dependent parameters that need to be considered for optimal lock performance. The first of these is the lock transmitter *power* used to excite the deuterium resonance. This needs to be set to the highest usable level to maximise the ^{2}H signal-to-noise ratio but must not be set so high that it leads to lock *saturation*. This is the condition in which more energy is applied to the deuterium spins than can be dissipated through spin relaxation processes and is evidenced by the lock level drifting up and down erratically. The saturation level can be determined by making small changes to the lock power and observing the effect on the lock signal. If the power is increased, the lock signal should also increase, but if saturated it will then drift back down. Conversely, if the power is decreased, the lock level will drop but will then tend to creep up a little. The correct approach is to determine the point of saturation and then to be sure you are operating below this before proceeding. The second feature is the lock *gain*. This is the amplification that is applied to the detected lock signal and is generally less critical than the power although it should not be so high as to introduce excessive lock noise. The final feature is the lock *phase*. We have already seen that the field-frequency regulation relies on monitoring the pure dispersion mode lineshape, and this is present only when the resonance has been phased correctly. This is the case when the *observed* lock signal produces the maximum intensity as the lock phase is adjusted since only then does this have the required pure absorption lineshape.

Acquiring data unlocked

In cases where deuterium is not available for locking, for example, if one wishes to observe or decouple deuterium itself, then it may be necessary to acquire data 'unlocked' (although recent hardware developments allow the decoupling of deuterium during acquisition by appropriate blanking of the lock channel). Running unlocked is perfectly feasible over limited time periods, which depend on the drift rate of the magnet. For experiments that must run for very many hours for reasons of sensitivity, then a useful approach is to acquire a series of FIDs each being collected over a shorter time period, say 1 h. Drift within each data set will then be essentially negligible whereas any drift occurring over the whole duration of the experiment can be corrected by adding *frequency-domain spectra* for which any frequency drift has been compensated manually by internal spectrum referencing. The addition then provides the required enhancement of signal-to-noise without the deleterious effects of field drift.

3.4.5. Optimising the field homogeneity: shimming

NMR experiments require a uniform magnetic field over the whole of the NMR sample volume that sits within the detection coil. Deviation from this ideal introduces various lineshape distortions, compromising both sensitivity and resolution. Thus, each time a sample is introduced into the magnet it is necessary to 'fine-tune' the magnetic field and a few minutes spent achieving good resolution and lineshape is time well spent. For anyone actually using an NMR spectrometer, competence in the basic level of field optimisation is essential, but even if you only need to *interpret* NMR spectra, perhaps because someone else has acquired the data or if the whole process is performed through automation, then some understanding of the most common defects arising from remaining field inhomogeneities can be invaluable.

The shim system

Maintaining a stable magnetic field that is uniform to 1 part in 10^9 over the active volume within modern NMR probes (typically 0.1–1.5 cm^3) is extremely demanding. This amazing feat is achieved through three levels of field optimisation. The first lies in the careful construction of the superconducting solenoid magnet itself, although the field homogeneity produced by these is rather crude when judged by NMR criteria. This basic field is then modified at two levels by sets of 'shim' coils. These coils carry electrical currents that generate small magnetic fields of their own, which are employed to cancel remaining field gradients within the sample. In fact, shims are small wedges of metal used in engineering to make parts fit together and were originally used in the construction of iron magnets to modify the position of the poles to adjust the field. Still in the present day where superconducting magnets dominate, this name permeates NMR, as does the term 'shimming', referring to the process of field homogeneity optimisation.[4] The superconducting shim coils sit within the magnet cryostat and remove gross impurities in the magnet's field. The currents are set when the magnet is first installed and do not usually require altering beyond this. The room temperature shims are set in a former that houses the NMR probe itself, the whole assembly being placed within the bore of the magnet such that the probe coil sits at the exact centre of the static field. These shims (of which there are typically around 20 to 40 on a modern instrument) remove any remaining field gradients by adjusting the currents through them, although in practice only a small fraction of the total number need be altered on a regular basis (see below).

The static field in vertical bore superconducting magnets also sits vertically and this defines, by convention, the z axis. Shims that affect the field along this axis are referred to as axial or Z shims, whereas those that act in the horizontal plane are known as radial or X/Y shims (Table 3.6). When acquiring high-resolution spectra, it is traditional practice to spin the sample (at about 10–20 Hz) about the vertical axis. This has the effect of averaging field inhomogeneities in the X–Y plane, so improving resolution. The averaging means that adjustments to shims containing an X or Y term must be made when the sample is static; hence, these shims are also commonly referred to as the 'non-spinning shims'. Modern shim sets are capable of delivering non-spinning lineshapes that almost match those when spinning, and it is becoming increasingly common not to spin samples. For multidimensional studies, this is certainly the case, since sample spinning can introduce modulation effects to the acquired data, leading to unwanted artefacts particularly in the form of so-called t_1 noise (see Chapter 5).

Shimming

In order to achieve optimum field homogeneity, high quality samples are essential. The depth of a sample also has a considerable bearing on the amount of Z shimming required, which can be kept to a minimum by using solutions of similar depth each time. Most spectrometers possess software that is capable of carrying out the shimming process automatically, and clearly this is essential if an automatic sample changer is used. However, such systems are not infallible and can produce spectacularly bad results in some instances. Here, reproducible sample depths are vitally important for the auto-shimming procedures to be successful and to reach an optimum rapidly. It is also crucial for the whole of the sample to be at thermal equilibrium so that convection currents do not exist. For aqueous solutions away from ambient, this may demand 10–20 min equilibration in the probe.

To provide an indication of progress when shimming, one requires a suitable indicator of field homogeneity. Essentially, there are three schemes that are in widespread use, all of which have their various advantages and disadvantages; (1) the lock level, (2) the shape of the FID and (3) the shape of the NMR resonance. The ultimate measure of homogeneity is the NMR resonance itself, since defects apparent in the spectrum can often be related

[4] Note that some (older) texts may refer to this process as "tuning", which is now exclusively reserved for processes involving rf; for example, one may *shim* a magnet but will *tune* a probe.

Table 3.6. Shim gradients found on high-resolution spectrometers. Lower field instruments (<500 MHz) may utilise only 20 or so gradients (such as those in the left panel), whilst higher field spectrometers may employ in excess of 30. Shims up to third order are those most likely to need periodic optimisation as part of long-term spectrometer maintenance for which the most significant interacting shims are listed. Those shown in square brackets interact less strongly with the listed gradient whilst those that interact with Z^0 (the main field) may cause momentary disruption of the lock signal when adjusted

Shim gradient	Gradient order	Principle interacting shim gradients	Shim gradient	Gradient order
Z^0 (the main field)	0	–	XYZ^2	4
Z^1 (Z)	1	–	$(X^2-Y^2)Z^2$	4
Z^2	2	Z	X^3Z	4
Z^3	3	Z, [Z^2]	Y^3Z	4
Z^4	4	Z^2, Z^0,[Z, Z^3]	XZ^4	5
Z^5	5	Z^3, Z, [Z^2, Z^4]	YZ^4	5
Z^6	6	Z^4, Z^2, Z^0, [Z, Z^3, Z^5]	XYZ^3	5
X	1	Y, [Z]	$(X^2-Y^2)Z^3$	5
Y	1	X, [Z]	XZ^5	6
XZ	2	X, [Z]	YZ^5	6
YZ	2	Y, [Z]	XYZ^4	6
XY	2	X, Y	$(X^2-Y^2)Z^4$	6
X^2-Y^2	2	XY, [X, Y]	XYZ^5	7
XZ^2	3	XZ, [X, Z]	$(X^2-Y^2)Z^5$	7
YZ^2	3	YZ, [Y, Z]		
XYZ	3	XY, [X, Y, Z]		
$(X^2-Y^2)Z$	3	X^2-Y^2, [X, Y, Z]		
X^3	3	X		
Y^3	3	Y		
XZ^3	4			
YZ^3	4			

directly to deficiencies in specific shim currents, as described below. Most often field homogeneity is monitored by the height of the deuterium lock resonance that one aims to maximise. Whilst conceptually this is a simple task, in reality it is complicated by the fact that most shims interact with others. In other words, having made changes to one it will then be necessary to re-optimise those with which it interacts. Fortunately, shims do not influence all others, but can be sub-divided into smaller groups that are dealt with sequentially during the shimming process. A detailed account of the shimming procedure has been described [45] and the fundamental physics behind field-gradient shims has also been presented [46], but here we shall be concerned more with addressing the lineshape defects that are commonly encountered in the daily operation of an NMR spectrometer.

When shimming, it is not always sufficient to take the simplest possible approach and maximise the lock level by adjusting each shim in turn, as this is may lead to a 'false maximum', in which the lock level appears optimum yet lineshape distortions remain. Instead, shims must be adjusted interactively. As an example of the procedure that should be adopted, the process for adjusting the Z and Z^2 shims (as is most often required) should be:

1. Adjust the Z shim to maximise the lock level, and note the new level.
2. Alter Z^2 so that there is a noticeable change in the lock level, which may be up or down, and remember the direction in which Z^2 has been altered.
3. Readjust Z for maximum lock level.
4. Check whether the lock level is greater than the starting level. If it is, repeat the whole procedure, adjusting Z^2 in the same sense, until no further gain can be made. If the resulting level is lower, the procedure should be repeated but Z^2 altered in the opposite sense.

If the magnetic field happens to be close to the optimum for the sample when it is initially placed in the magnet, then simply maximising the lock level with each shim directly will achieve the optimum since you will be close to this already. Here again, a reproducible sample depth makes life very much easier.

Shimming is performed by concentrating on one interacting group at a time, always starting and finishing with the lowest order shim of the group. Principle interactions for selected shims are summarised in Table 3.6. Whenever it is necessary to make changes to a high order shim, it will be necessary to readjust all the low order shims within the same

interacting group, using a similar cyclic approach to that described for the adjustment of Z and Z^2 above. Generally, the order of optimisation to be followed will be:

1. Optimise Z and Z^2 interactively, as described above. If this is the first pass through Z and Z^2, then adjust the lock phase for maximum lock level.
2. Optimise Z^3: make a known change, then repeat step 1. If the result is better than previously, repeat this procedure, if not alter Z^3 in the opposite sense and repeat step 1.
3. Optimise Z^4 interactively with Z^3, Z^2 and Z.
4. Stop the sample spinning (if applicable) and adjust Z to give the maximum response (this is likely to have changed a little as the position of the sample relative to the field will change). Adjust X and Y in turn to give the maximum response.
5. Optimise X and XZ interactively. Adjust Z to give the maximum response.
6. Optimise Y and YZ interactively. Adjust Z to give the maximum response.
7. Optimise XY interactively with X and Y.
8. Optimise X^2-Y^2 interactively with XY, X and Y.
9. Repeat step 1.

The higher the shim order, the greater the changes will be required, and when far from the optimum shim settings, large changes to the shim currents may have only a small effect on the lock level and the shim response will feel rather 'sluggish'. When close to the optimum, the response becomes very sensitive and small changes can have a dramatic effect. The above procedure should be sufficient for most circumstances and any field strength, unless the basic shims set has become grossly misset. If lineshape distortions remain, then it may be possible to identify the offending shim(s) from the nature of the distortion (see below), allowing the appropriate corrections to be applied.

Common lineshape defects

The NMR resonance lineshape gives the ultimate test of field homogeneity, and it is a useful skill to be able to recognise the common distortions that are caused by errors in shim settings (Fig. 3.53). Thus, the Z shims all influence the width of the NMR resonance, but in subtly different ways; impurities in the even order shims (Z^2, Z^4 and Z^6) will produce unsymmetrical distortions to the lines whereas those in the odd orders (Z, Z^3 and Z^5) will result in symmetrical broadening of the resonance. In any case, the general rule is that the higher the order of the shim, the lower down the resonance the distortions will be seen. Errors in Z^3 usually give rise to a broadening of the base of a resonance and, since a broad resonance corresponds to a rapid decay of the FID, such errors are sometimes seen as a sharp decay in the early part of the FID. Another commonly observed distortion is that of a shoulder on one side of a peak, arising from poorly optimised Z and Z^2 shims (this is often associated with reaching a 'false maximum' simply by maximising the lock level with each shim and is usually overcome by making a significant adjustment to Z^2, and following the procedure described above).

Errors in the low order X/Y shims give rise to the infamous 'spinning sidebands' (for a spinning sample). These are images of the main resonance displaced from it by multiples of the spinning frequency. Shims containing a single X or Y term produce 'first order sidebands' at the spinning frequency from the main line whereas XY and X^2-Y^2 give second order sidebands at double the spinning frequency. However, unless something has gone seriously amiss, you should not encounter more than the first order sidebands in every day NMR at most, and these should certainly be no greater than one or two percent of the main resonance. If there is any doubt as to the presence of sidebands, a simple test is to alter the spinning speed by say 5 Hz and re-acquire the data; only the sidebands will have moved. If the sample is not spinning, errors in the low order X/Y shims contribute to a general broadening of the resonance.

Shimming using the FID or spectrum

Although the lock level is used as the primary indicator of field homogeneity, it is not always the most accurate one. The lock level is dependent on only one parameter; the height of the deuterium resonance. This, whilst being rather sensitive to the width of the main part of the resonance, is less sensitive to changes in the broad base of the peak. The presence of such low level lumps can be readily observed in the spectrum (particularly in the case of protons), but for this to be of use when shimming, the spectrometer must be able to supply a 'real-time display' of a single-scan spectrum so that changes to the shim currents can be assessed rapidly. With modern host computers, the FT and phase correction of a spectrum can be performed very rapidly, allowing one to correct for lineshape distortions as one shims. Alternatively, the shape and the duration of the FID may be used as a more immediate indicator of homogeneity and can often be used simultaneously with the lock display. This approach works best when a singlet resonance dominates the spectrum (such

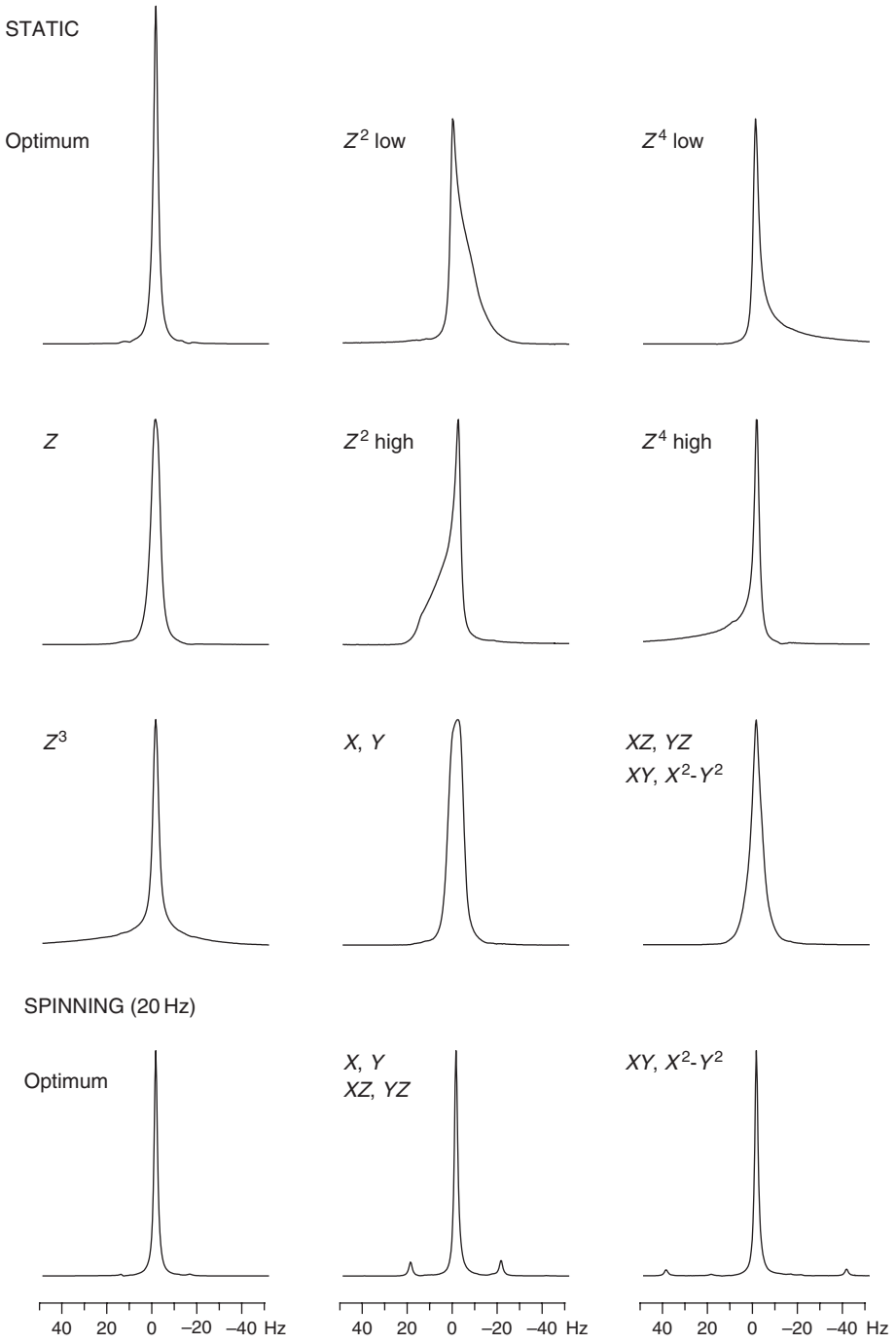


Figure 3.53. Lineshape defects that arise from inappropriate settings of various shims. These effects have been exaggerated for the purpose of illustration.

as for aqueous solutions) for which the shape of the FID should be a smooth exponential decay. Since with this method of shimming it is likely that changes to the shim currents will be made *during* the acquisition of the spectrum (which will certainly lead to a peculiar lineshape), it is essential that one assesses a later spectrum for which there have been no adjustments during acquisition to decide whether improvements have been made.

Gradient shimming

The most recent approach to field optimisation comes from the world of magnetic resonance imaging and makes use of field gradients to map B_0 inhomogeneity within a sample. This can then be cancelled by calculated changes to the shim settings [47, 48]. The results that can be attained by this approach are little short of astonishing when seen for the first time, especially for anyone who has had to endure the tedium of extensive manual shimming of a magnet, and this method now enjoys widespread popularity.

The discussions below assume some understanding of the action of PFGs, and the reader not familiar with these may wish to return to this section after they are introduced in Chapter 5. In any case, an appreciation of the capabilities of this method should be readily achieved from what follows. Here we shall consider the basis of the method with reference

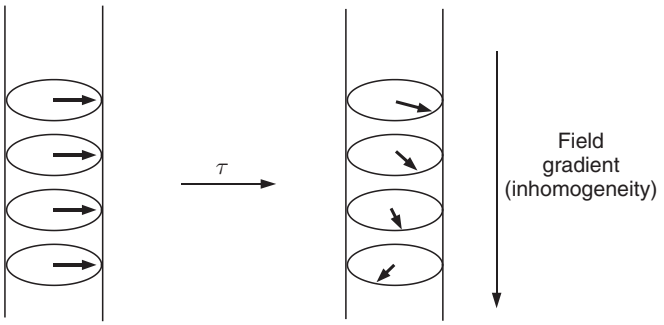


Figure 3.54. Errors in the local static field along the length of a sample are encoded as signal phases when magnetisation is allowed to precess in a inhomogeneous static field for a time τ. The resulting spatially dependent phase differences are used as the basis of gradient shimming.

to the optimisation of Z shims which requires z axis PFGs that are commonly found in modern probeheads (although the use of conventional shim assemblies to generate the appropriate gradients with so-called homospoil pulses has been demonstrated [49], which has the advantage of not requiring specialised gradient hardware). The underlying principle is that all spins throughout a sample contributing to a singlet resonance will posses the same precession (Larmor) frequency *only* if the static field is homogeneous throughout (of course this is what we aim for when shimming!). Any deviation from this condition will cause spins in physically different locations within the sample to precess at differing rates according to their *local* static field. If the excited spins are allowed to precess in the transverse plane for a fixed time period prior to detection, these differing rates simply correspond to different phases of their observed signals (Fig. 3.54). By detecting these signals in the presence of an applied field gradient, the *spatial distribution* of the spins becomes encoded as the *frequency distribution* in the spectrum allowing the inhomogeneity (encoded as phase differences) to be mapped along the length of the sample in the case of z axis gradients (or across the sample for x and y axis gradients).

A suitable scheme for recording this is the gradient echo of Fig. 3.55 in which spins are first dephased by a PFG and later rephased after a period of precession τ_1 to allow detection. The resulting spectrum is the one-dimensional spatial profile (or image) of the sample (Fig. 3.56). Recording a second echo with delay τ_2 and taking the difference yields the *phase map* in which only free precession during the period $(\tau_2-\tau_1)$ is encoded. The phase distribution in this profile therefore directly maps the inhomogeneity along the sample. The necessary corrections to shim currents to remove these inhomogeneities are calculated from a series of reference phase maps recorded with known offsets in each of the Z shims. Once these reference maps have been recorded for a given probe, they can be used for the gradient shimming of all subsequent samples, with the whole process operating automatically. A more recent approach to optimisation [48] seeks to provide the best possible resonance lineshape rather than aiming for minimal B_0 field errors by simulating

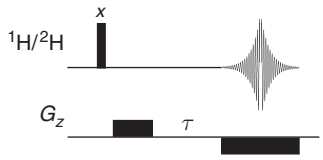


Figure 3.55. A gradient echo sequence suitable for z axis field gradient shimming. The pulsed field gradients provide the spatial encoding whilst the delay τ encodes the static field inhomogeneity as signal phase.

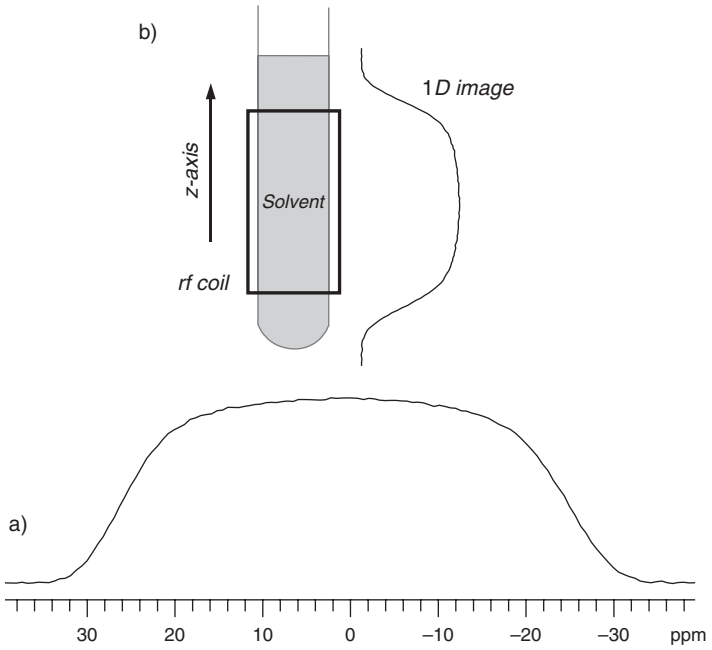


Figure 3.56. (a) The one-dimensional (1D) z axis deuterium image of dimethylsulfoxide (magnitude mode). (b) The frequency axis of this image encodes the spatial dimension along the length of the sample.

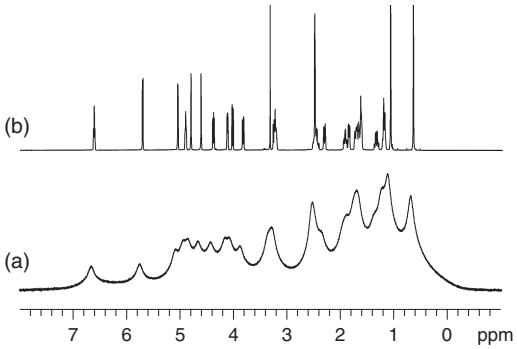


Figure 3.57. Automatic deuterium gradient shimming. Spectrum (a) was acquired with all z axis shims set to zero. After less than two minutes of gradient shimming of the Z–Z^5 shims (three iterations), spectrum (b) was obtained. The solvent was dimethylsulfoxide and one scan was acquired for each deuterium gradient echo collected via the probe lock coil. Automatic switching for deuterium observation was achieved with a home-built switching device.

peak shapes from the measured field maps and knowledge of the B_1 (rf) profile of the rf coil. This is better able to factor in field discontinuities at the ends of sample columns, for example, and is claimed to provide a more robust and reliable approach with higher quality results; experience suggest this is indeed case.

The primary experimental requirement for gradient shimming is a sample containing a dominant strong, singlet resonance. A good candidate for proton observation is 90% H_2O, but although this is ideal for biomolecular studies, it is clearly of little use for the majority of solvents used in organic spectroscopy. An alternative in this case is to observe the deuterium resonance of the solvent [50] (which in most cases is also a singlet) using the lock channel of the probe. The potential problem then is one of sensitivity and the need for appropriate hardware to allow deuterium observation on the lock coil without manually recabling the instrument each time a sample is shimmed. The necessary lock channel switching devices are commercially available, or, more recently, dedicated deuterium transmitters are built into lock systems to allow direct ^2H observation (and decoupling).

The remarkable power of gradient shimming is illustrated in Fig. 3.57. The lower spectrum was recorded with the Z–Z^5 shims all set to zero whilst the upper trace was the result of only three iterations of deuterium gradient shimming using the dimethylsulfoxide solvent resonance. The whole process took less than 2 min without operator intervention. Although a rather extreme example, the capabilities of this approach are clearly evident and explain the increasing use of this methodology, which now plays a valuable role in automated spectroscopy where irreproducible sample depths can lead to rather poor results with conventional simplex optimisation shim routines. The individual mapping of field errors within each and every sample overcomes these problems in a time-efficient manner.

3.4.6. Reference deconvolution

Despite efforts to achieve optimum shimming, either manually or through automated procedures, there are invariably instances when the field homogeneity in a sample does not meet expectations, leading to lineshape distortions in the resulting spectrum. Reference deconvolution is a post-processing strategy that aims to correct such distortions and is now a standard feature of some NMR processing software [51–54]. The process relies on the fact that shim errors will influence all resonances in the distorted spectrum in the same way. If the type and extent of lineshape error can be determined and a correction factor computed for a single resonance, this adjustment may be applied to all others to yield a spectrum free from the unwanted distortions. The principal limitation of the technique is in the selection of a suitable reference resonance with many common multiplet structures proving unsuitable since they give rise to null points in their FIDs; a singlet is thus most appropriate although multiplets have been employed [55]. The reference deconvolution procedure is summarised pictorially in Fig. 3.58. Having identified a suitable reference peak in the spectrum that is free from others, this is extracted (by zeroing all other data points) and its FID generated by inverse fourier transformation (iFT) of the reduced spectrum. An idealised lineshape is then chosen and the corresponding FID similarly computed. This shape will most likely have a Lorentzian profile, corresponding to an exponential decay of the FID, but others may also be selected including a Gaussian profile should resolution enhancement of the spectrum be desired. A comparison of the experimentally derived FID and the idealised FID is then made, a correction function computed and this applied to the original complete FID. This may then be transformed to yield the corrected spectrum.

The reference deconvolution procedure is not limited to addressing lineshape distortions in 1D spectra, having a number of potential applications in high-resolution NMR spectroscopy [53, 54] and has been shown to be effective in improving NOE difference spectra [52], for t_1 noise reduction in 2D spectra [56, 57] and in the production of high-resolution diffusion-ordered spectroscopy (DOSY) [58].

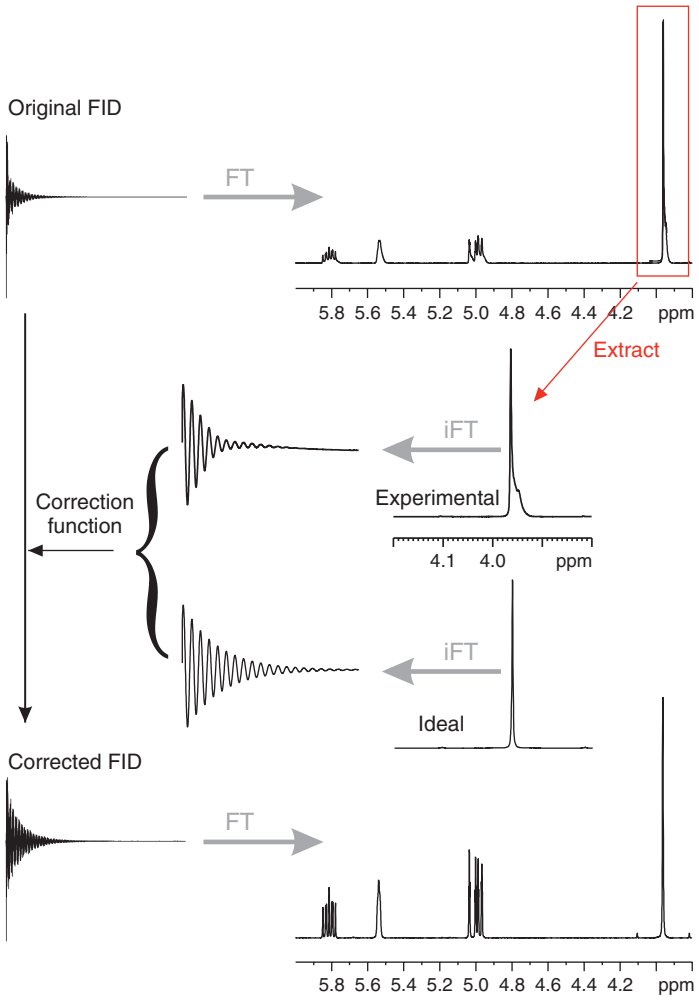


Figure 3.58. A schematic representation of the reference deconvolution procedure.

3.5. SPECTROMETER CALIBRATIONS

This section is primarily intended for those who need to set up experiments or those who have new hardware to install for which new calibrations are required. As with any analytical instrumentation, correct calibrations are required for optimal and reproducible instrument performance. All the experiments encountered in this book are critically dependent on the application of rf and gradient pulses of precise amplitude, shape and duration, and the calibrations described below are therefore fundamental to the correct execution of these sequences. Periodic checking of these calibrations, along with the performance tests described in the following section, also provides an indication of the overall health of the spectrometer.

3.5.1. Radiofrequency pulses

Modern multipulse NMR experiments are critically dependent on the application of rf pulses of known duration (the *pulse width*) that correspond to precise magnetisation tip angles, most frequently 90° and 180°. Pulse width calibrations are normally defined for the 90° pulse (PW_{90}), from which all other tip angles may be derived. Those with PW_{90} in the microsecond range generally excite over a rather wide bandwidth and are termed *non-selective* or *hard* pulses whilst those in the millisecond range are of lower power and are effective over a much smaller frequency window and are thus termed *selective* or *soft* pulses. A third class of pulse is the *frequency-swept* or *adiabatic* pulse, which can be effective over far wider frequency bandwidths than simple hard pulses and have become especially important with higher field spectrometers. The implementation of selective and adiabatic pulses and their calibration is considered separately in Chapter 10, whilst here we shall concentrate on the more widely used hard pulses and on weaker rf fields used for decoupling purposes.

The rf field strengths

The calibration of a pulse width is equivalent to determining the radiofrequency (B_1) field strength of the pulse, and, in fact, it is often more useful to think in terms of field strengths than pulse widths when setting up experiments. For example, excitation bandwidths and off-resonance effects are best considered with reference to field strengths, as are decoupling bandwidths. The relationship between pulse width and the rf field strength γB_1 is straightforward:

$$\gamma B_1 = \frac{1}{\mathrm{PW}_{360}} \equiv \frac{1}{4\mathrm{PW}_{90}}\,\mathrm{Hz} \tag{3.12}$$

Thus, a 90° pulse width of 10 μs corresponds to a field strength of 25 kHz, a typical value for pulse excitation on a modern spectrometer.

Although spectrometer transmitters are frequently used at full power when applying single pulses, there are many instances when lower transmitter powers are required, for example, the application of decoupling sequences or of selective pulses. These lower powers are derived by attenuating the transmitter output, and the units used for defining the level of attenuation are decibel or dB. This is, in fact, a measure of the ratio between two power levels, P_1 and P_2, as defined by:

$$\mathrm{dB} = 10\log_{10}\left(\frac{P_1}{P_2}\right) \tag{3.13}$$

When one speaks of attenuating the transmitter output by so many dBs, one must think in terms of a change in power relative to the original output. In fact, it is more convenient to think in terms of changes in output voltage rather than power, since the rf field strength and hence pulse widths (our values of interest) are proportional to voltage. Since power is proportional to the *square* of the voltage ($P = V^2/R$, where R is resistance), we may rewrite Eq. (3.13) as:

$$\mathrm{dB} = 20\log_{10}\left(\frac{V_1}{V_2}\right) \equiv 20\log_{10}\left(\frac{\mathrm{PW}_1}{\mathrm{PW}_2}\right) \tag{3.14}$$

where PW_1 and PW_2 are the pulse widths for the same net tip angle at the two attenuations. Thus, if one wished to double a pulse width, an additional 20 log 2 dB (6 dB) attenuation is required. An alternative expression has the form:

$$+n\mathrm{dB} \Rightarrow \mathrm{PW} \times 10^{\frac{n}{20}} \tag{3.15}$$

or in other words, addition of ndB attenuation increases the pulse width by a factor of $10^{n/20}$. Some example attenuation values are presented in Table 3.7.

An illuminating experiment to perform, if your spectrometer is able to alter the output attenuation internally, is to determine the pulse width over a wide range of attenuations. A plot of pulse widths *vs* $\log_{10}$(attenuation) should yield a single straight line over the full range. Discontinuities in the plot may arise when different power amplifiers come into use or when large attenuators are switched in place of many smaller ones. More recent spectrometers make use of linear amplifiers together with calibrated power correction look-up tables, providing pulse output powers that are linear over the whole attenuation range. This means pulse calibrations at different power settings become superfluous since all pulse widths can be calculated with Eq. (3.14) from a single, accurate high power calibration, so it is worthwhile checking if your spectrometer is equipped for this. This also makes the application of more elaborate pulse profiles whose performance can be critically dependent on accurate pulse calibrations (such as the shaped pulses described in Chapter 10) far more straightforward since their calibrations may also be derived in this way.

Observe pulses: high sensitivity

We begin with the most basic NMR pulse calibration, that for the observed nucleus. When the spectrum of the analyte can be obtained in a single scan, or only a few scans, it is straightforward to calibrate the pulse width simply by following the behaviour of the

Table 3.7. Pulse width *vs* attenuation. The additional attenuation of transmitter output levels increases the pulse width by the factors shown. Thus, adding 6 dB will double the pulse width whereas removing 6 dB will halve it

Attenuation (dB)	1	3	6	10	12	18	20	24
Pulse width factor	1.1	1.4	2.0	3.2	4.0	7.9	10	15.8

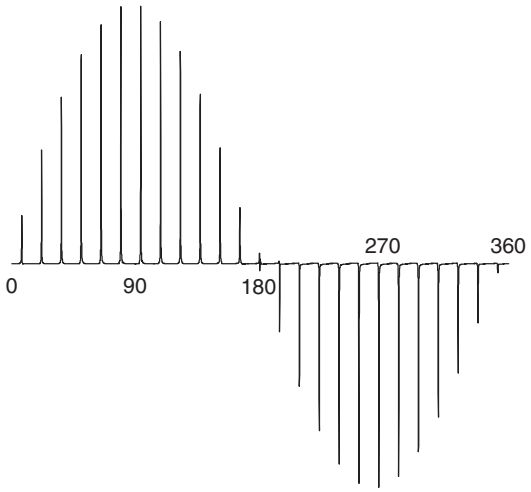


Figure 3.59. Pulse width calibration for the observe channel. A sequence of experiments is recorded with a progressively incremented excitation pulse. The maximum signal is produced by a 90° pulse and the first null with a 180° pulse. Either the 180° or the 360° condition can be used for the calibration (but be sure to know which of these you are observing).

magnetisation as the excitation pulse is increased, with nulls in the signal intensity occurring with a 180° or 360° pulse. To perform any pulse calibration, it is essential that the probe is first tuned for the sample so that the results are reproducible. The transmitter frequency should be placed on-resonance for the signal to be monitored, to eliminate potential inaccuracies arising from off-resonance effects, and a spectrum recorded with a pulse width less than 90°, say 2–3 µs. The resonance is used to define the phase correction for all subsequent experiments and is phased to produce the conventional positive absorption signal. The experiment is repeated with identical phase correction but with a progressively larger pulse width (Fig. 3.59). Passing through the 180° null then provides the PW_{180} calibration, whilst going beyond this yields an inverted resonance, so it becomes clear when you have gone too far. Detecting the null precisely can sometimes be tricky, as there is often some slight residual signal remaining, the exact appearance depending on the probe used, so in practice one aims for the minimum residual signal.

Between acquisitions there must be a delay sufficient for complete relaxation of the spins, and if signal averaging is used for each experiment, this must also be applied between each scan to obtain reliable calibrations. In proton spectroscopy, such delays are unlikely to be too much of a burden, but can be tediously long for slower relaxing spins. If you have some feel for what the pulse width is likely to be, perhaps from similar samples or a 'reference' calibration, then a better approach is to search for the 360° null. Since magnetisation remains close to the $+z$ axis, its recovery demands less time and the whole process can be performed more quickly. The process can be further simplified by monitoring the FID alone, which also has minimum amplitude with the 360° pulse. In any case, it is wise to check the null obtained is in fact the one you believe it to be, as all multiples of 180° give rise to nulls.

Observe pulses: single-scan nutation spectroscopy

A more recent approach to pulse calibration allows this to be performed in a single scan and provides a rapid and convenient method for ^{1}H pulse calibration in particular that is also well suited to use under automation [59]. The method has most significance for inverse and cryogenic probes where pulse calibrations may vary significantly between samples. The measurement derives from a single-scan nutation experiment in which an rf pulse is applied during each of the dwell periods of a FID acquisition between data point sampling in a manner similar to that used in the homonuclear decoupling experiments described in Section 4.2.2. No initial excitation pulse is required. Provided the rf pulses are applied on-resonance for a single peak, the associated magnetisation vector is driven by each of the pulses away from its equilibrium position along the longitudinal axis as the acquisition proceeds; the magnetisation is said to *nutate* about the applied rf field. The transverse magnetisation generated by this process describes an oscillatory behaviour along one axis during data acquisition (Fig. 3.60a). One may imagine this oscillatory behaviour as being equivalent to the *net* magnetisation generated by the evolution of a doublet that is on-resonance in the rotating frame (Fig. 3.60b; see also Fig. 2.13). Indeed, FT of the nutation data leads to the generation of an anti-phase doublet in the resulting nutation spectrum with a splitting $\Delta\nu$ Hz (Fig. 3.61). The oscillation frequency reflects the mean nutation frequency of the vector during data collection and equates to the frequency of evolution of each half of the hypothetical doublet $\Delta\nu/2$. This provides the mean γB_1 value of the applied rf and hence the pulse width calibration under the conditions used.

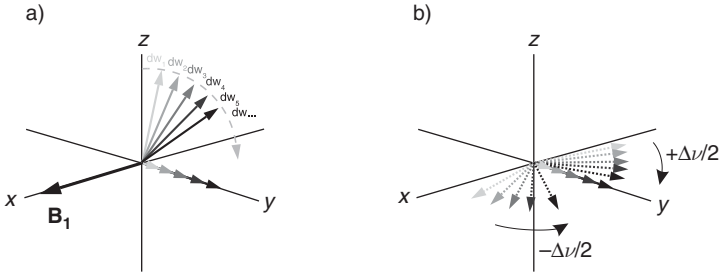


Figure 3.60. The behaviour of an on-resonance magnetisation vector during the nutation experiment. (a) As the rf field is applied during each of the FID dwell periods (dw_1, dw_2, dw_3, ...) the vector is driven stepwise about the B_1 field (shown in shades of grey). The sampled transverse component (likewise greyed) oscillates along the y axis as the data acquisition proceeds and the total nutation angle progressively increments. (b) Equivalent behaviour arises from the evolution of an anti-phase doublet of splitting $\Delta\nu$.

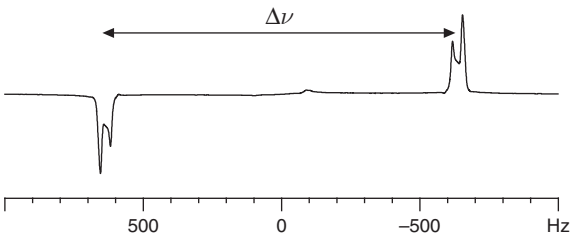


Figure 3.61. The single-scan nutation spectrum of an aqueous sample (12 dB attenuation, 10% duty cycle). The doublet splitting of 1268 Hz calibrates the high power 90° hard pulse at 9.86 µs.

To avoid possible probe damage and sample heating, it is safe practice during the nutation experiment to reduce the rf amplitude from that used for hard pulses by typically 12 dB, corresponding to a fourfold increase in pulse width. It also becomes necessary to account for the fraction of time during which the rf pulse is applied within each dwell period of the FID, the so-called duty cycle, d. Values may range from 10–80% ($d = 0.1$–0.8), and the measured rf amplitude must be corrected for this. With the low power calibration complete the required hard pulse calibration may be calculated directly when linearised rf amplifiers are in use, as found on modern instruments, otherwise additional calibrations may be required to correct for the power differences. Assuming amplifier linearisation and the use of 12 dB attenuation in the nutation experiment, the high power pulse width may be computed as

$$t_{360} = \frac{d}{4(\Delta\nu/2)} \quad \text{or} \quad t_{90} = \frac{d}{8\Delta\nu} \tag{3.16}$$

The method is most appropriate when the spectrum is dominated by a single resonance, so has obvious application to samples in heavily protonated water but can be equally effective for organic solutions when a single dominant peak can be identified, such as a resolved methyl resonance for example. This must be placed on-resonance, and the smallest frequency separation between the anti-phase peaks in the nutation spectrum is then used in the appropriate calculation. The whole procedure is most conveniently executed through appropriate software routines, which may be supplied as standard by your instrument vendor.

Observe pulses: low sensitivity

When the sample is too weak to allow its observation within a few scans, one possible approach is to use the method of signal averaging and to search for the 360° null as described above [60], although this could become a rather laborious affair. An alternative approach [61] requires that only two spectra are collected with the pulse width of the second being exactly double that of the first ($PW_2 = 2PW_1$). Once again, a sufficient delay between scans is required to avoid saturation effects, and the spectrometer must be operating in an 'absolute intensity' scaling mode so that the relative signal intensities from both experiments can be compared. Heteronuclear spectra should also be acquired without enhancement by the NOE, and the use of the inverse-gated decoupling scheme (Section 4.2.3) allows such data to be collected with broadband decoupling during acquisition to aid sensitivity. Following some simple arithmetic [61], it can be shown that

$$\theta_1 = \cos^{-1}\left(\frac{I_2}{2I_1}\right) \tag{3.17}$$

where I_1 and I_2 correspond to the measured signal intensities in the first and second experiment, respectively. The duration of the pulse in the first experiment corresponds to the pulse angle θ_1, from which the 90° pulse width can be derived. For optimum results, θ_1 should be greater than 30° and θ_2, the tip angle in the second experiment, should be less than 180°, so again some rough feel for the likely answer is required for this approach to be efficient.

A more convenient approach in many instances is to perform the calibrations not on the sample of interest itself but on a suitable sample made up for the purposes of calibration that is strong enough to permit the observation of the 180/360° null directly and quickly. The calibration can then be recorded and used in future experiments, as this should be quite reproducible if similar solvents are used and the probe correctly tuned. The largest discrepancies occur between aqueous and organic solvents, especially when the aqueous solutions are highly ionic, and separate calibrations may be required.

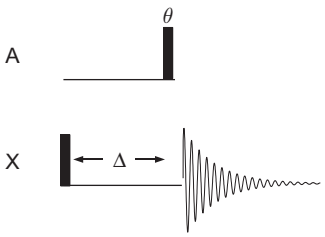


Figure 3.62. The sequence for the calibration of "indirect" transmitter pulses (those on nuclei other than that being observed). The period Δ is set to $1/2J_{AX}$.

Indirect pulses

In contrast to above, the calibration of pulses for nuclei other than that being observed must be performed 'indirectly'. This is often referred to as the calibration of the 'decoupler' pulses although this is something of a misnomer nowadays since the pulses may have nothing to do with decoupling itself but may be an integral part of the pulse sequence. Most commonly, the indirect channel is used in two modes; the first is to apply short, hard pulses typically at maximum power, the second is to apply a pulsed decoupling sequence for longer periods of time, at significantly less power. It is therefore commonplace to record indirect calibrations at these two power levels, for which the following method is equally applicable, and having performed the calibration at high power, the low power attenuation can be estimated from Eq. (3.14) prior to its accurate determination. Alternatively, the use of accurately linearised amplifiers can avoid the need for low power calibrations altogether. More recent spectrometers make use of linear amplifiers together with calibrated power correction look-up tables, meaning the pulse output power is linear over the whole attenuation range. This means low power calibrations become superfluous since all pulse widths can be calculated from a single high power calibration with Eq. (3.14), so it is worthwhile checking if your spectrometer is equipped for this. This also makes the application of more elaborate pulse profiles, such as the shaped and adiabatic pulses described in Chapter 10, far more straightforward since their calibrations may also be computed for both observe and indirect pulses.

Indirect pulses on high abundance nuclides

Calibrations in which the indirect nuclide exists at high abundance are generally more straightforward because of the clear appearance of heteronuclear coupling. Thus, to relate the behaviour of the indirect spin, A, to that being observed, X, one exploits the mutual heteronuclear scalar coupling that must exist between them, J_{AX}, for the calibration to be possible. The simplest calibration sequence [62, 63] suitable for spin-½ nuclei is shown in Fig. 3.62. The A and X pulses should be on resonance, the X-pulse calibration must already be known, and the value of J_{AX} measured directly from the (coupled) X spectrum. The delay Δ is set to $1/2J_{AX}$ s for an AX group so that the X spin doublet vectors are anti-phase after this evolution period. The subsequent A pulse renders the *x*-magnetisation *unobservable* when $\theta = 90°$ as this generates pure heteronuclear multiple-quantum coherence, which is unable to generate a detectable signal in the probe (see Chapter 5 for an explanation of this effect).

The approach is to begin with θ very small and to phase the spectrum so that the doublet lines are in anti-phase. As θ increases, the doublet intensity will decrease and become zero when θ is 90°, whilst beyond this the doublet reappears but with inverted phase (Fig. 3.63). If it is necessary to perform the calibration with an A_2X group, the delay Δ should be 1/4J and it is the outer lines of the triplet that behave as described above whilst the centre line remains unaffected. When calibrating lower powers for the purpose of broadband decoupling, it is usually more convenient to set the *duration* of the θ pulse according to the decoupler bandwidth required (Section 10.3) and to vary the output attenuation to achieve the null condition. As is often the way in experimental procedures, simplicity is best and the indirect pulse calibrations can most conveniently be performed on strong calibration samples using the sequence of Fig. 3.62. For ^1H–^{13}C systems, $CHCl_3$ and (^{13}C-labelled) methanoic acid are convenient materials for organic and aqueous calibrations respectively.

A readily available alternative that is particularly suitable for (although not restricted to) proton calibration in ^1H–X systems is the distortionless enhancement by polarisation transfer (DEPT) sequence (Chapter 4). This improves observation sensitivity by making use of polarisation transfer from ^1H to X and is dependent on the faster relaxing protons for repetition rates in signal averaging. The 90° proton pulse is achieved when XH_2 or XH_3 groups pass through a null whereas XH groups have maximum intensity at this point (Fig. 3.64). One caveat here is that for the purposes of pulse calibration, all the proton pulses in the sequence

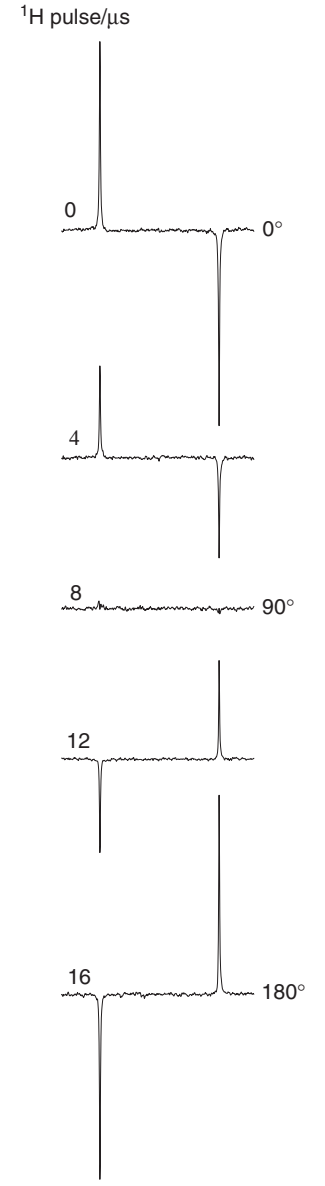


Figure 3.63. Indirect calibration of the proton pulse width with carbon observation using the sequence of Fig. 3.62. As the proton pulse width increases, the carbon signal diminishes until it disappears at the $^1H(90°)$ condition. Going beyond this causes the signal to reappear but with inverted phase (the same phase correction is used for all spectra). The sample is ^{13}C-labelled methanoic acid in D_2O.

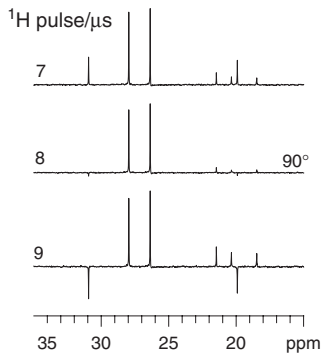


Figure 3.64. Calibration of indirect proton pulses using the DEPT sequence. When the proton editing pulse in the sequence is exactly 90°, only methine resonances are apparent (see Chapter 4 also). The sample is menthol in $CDCl_3$.

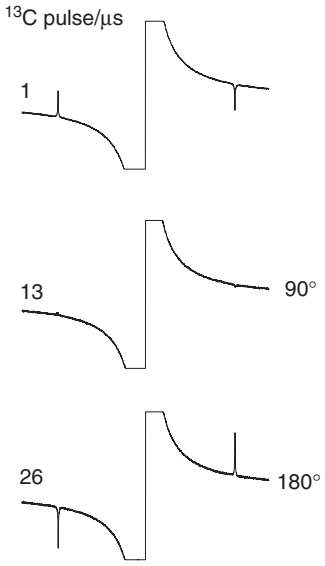


Figure 3.65. Calibration of indirect pulses (here ^{13}C) with proton observation using the sequence of Fig. 3.62. The parent ^{1}H(^{12}C) resonance appears as a large dispersive signal, but the behaviour of the satellites is clear, again disappearing with a 90° pulse on the indirect channel. The sample is $CHCl_3$.

need to be altered on each experiment, not just the final 'θ' proton pulse (see Chapter 4 for an explanation of this sequence).

Indirect pulses on low abundance nuclides

When directly observing protons and indirectly calibrating nuclei of low abundance such as ^{13}C and ^{15}N (the so-called inverse configuration), the dominant resonance will be of no use and may interfere since it is the much less intense coupled satellites that must be employed for the calibration. Thus, the major ^{12}C line of $CHCl_3$ will be apparent as a large dispersive signal between the ^{13}C satellites of interests when using the sequence of Fig. 3.62, but this should not pose too great a problem so long as the sensitivity is sufficient for the satellites to be readily observed (Fig. 3.65). In cases when the major resonance hampers observation of the satellites, a PFG-based variant may be employed that removes this. The 2CALIS sequences [64] (calibration for indirect spins) are shown in Fig. 3.66 and may be most readily understood by considering $\theta = 90°$, where the duration of θ is again the pulse calibration being sought and $\Delta = 1/2J_{HX}$. The initial ^{1}H–X spin-echo generates anti-phase ^{1}H doublet vectors that are then rendered longitudinal by the subsequent ^{1}H 90_y pulse (this is the same initial behaviour as in the insensitive nuclei enhanced by polarisation transfer (INEPT) sequence described in Chapter 4). The strong purge gradient G_2 then destroys all magnetisation arising from non X-coupled protons, that is, the ^{1}H–^{12}C resonance in the case of carbon calibration. Following this, the simultaneous 90° pulses on both ^{1}H and X transforms the longitudinal, anti-phase ^{1}H magnetisation into unobservable multiple quantum coherence and so eliminates the satellites when $\theta = 90°$, as for the simpler sequence of Fig. 3.62. Thus, as θ increases the satellite intensities decay toward this zero-crossing condition (Fig. 3.67). Improved suppression of the unwanted parent resonance may be achieved with additional X spin filtration by alternating the final X pulse between θ and 3θ (that is, 90° and 270° for the ideal calibration condition) combined with alternation of the receiver. The non-X-coupled proton magnetisation remains invariant to these pulses and so is cancelled by this receiver cycle. Optionally, an additional refocusing period may be employed, as in 2CALIS-1(Fig. 3.66b), to yield in-phase ^{1}H doublets that may, again optionally, be X spin decoupled. This approach may prove advantageous when signal overlap makes the presence of anti-phase signals susceptible to accidental cancellation with neighbouring resonances. Note that for these sequences to generate any observable signals, 2θ must be close to 180° so that the heteronuclear spin-echo generates anti-phase magnetisation; if no X spin 180° pulse is applied (that is if $\theta = 0°$), the heteronuclear coupling would be fully refocused. Thus, a reasonable estimate of θ is required before the calibration sequence begins as would typically be known from standard calibration samples. 2CALIS would then be employed for accurate calibration on the sample under study.

For the indirect calibration of natural abundance ^{15}N with ^{1}H observation, the dominant ^{14}N-bound proton resonance may be rather broad owing to the quadrupolar ^{14}N nucleus and may mask the small ^{15}N satellites. In such cases, the 2CALIS sequences may prove especially useful. Alternatively, the ^{14}N line can be preferentially suppressed without the need for PFGs, owing to its faster transverse relaxation rate, by the application of a spin-echo prior to the sequence of Fig. 3.62; see reference [65] for further details. A more convenient alternative is to use ^{15}N-labelled materials for calibration purposes. Calibration over long-range couplings is also possible, for example, when no one-bond coupling exist, for which the original CALIS-2 sequence [66] is best suited.

Homonuclear decoupling field strength

Homonuclear decoupling requires the decoupling (B_2) field to be on during the acquisition of the FID and, as described in Chapter 4, this demands the use of rapid gating of the decoupler during this period, so the methods presented above are therefore no longer applicable. The field strengths used in homonuclear decoupling are considerably less than those required for pulsing, and it is sometimes useful to have some measure of these when setting up decoupling experiments, to achieve the desired selectivity. A direct measure of the mean field strength produced during homonuclear decoupling can be arrived at through measurement of the Bloch–Siegert shift [67], which is the change in the resonance frequency of a line on the application of nearby irradiation. Provided the offset of the decoupler from the unperturbed resonance is greater than the decoupler field strength ($\nu_o - \nu_{dec} >> \gamma B_2$), then

$$\gamma B_2 = ((\nu_{BS} - \nu_o) \cdot 2(\nu_o - \nu_{dec}))^{1/2} \text{Hz} \tag{3.18}$$

where ν_{BS} is the signal frequency in the presence of the irradiating field, ν_o is that in the absence of the irradiating field and ν_{dec} the frequency of the applied rf. An example of such a calibration is shown in Fig. 3.68.

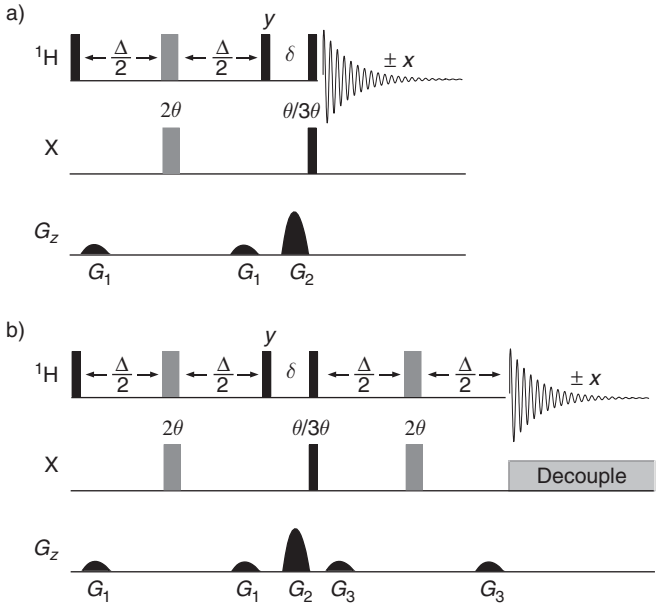


Figure 3.66. The 2CALIS sequences for the calibration of indirectly observed, low abundance nuclides. (a) 2CALIS-2 for anti-phase satellite display and (b) the 2CALIS-1 refocused version with X-spin decoupling. All pulse phases are x unless indicated and the optimum phase cycle is y, y,$-y$, $-y$ on the second proton 90° pulse with corresponding alternation of the receiver (x, $-x$, x, $-x$) as the final X pulse *duration* steps θ, 3θ, 3θ, θ.

Heteronuclear decoupling field strength

Complementary to the pulse methods described above for the measurement of the 'indirect' pulse widths is the measurement of the heteronuclear decoupling field strength. This method is applicable to the measurement of medium to low power outputs, such as those used in broadband decoupling, soft pulses and selective decoupling, and makes use of continuous, rather than pulsed, irradiation. Once again the method exploits the heteronuclear coupling between spins and observes the changes to this on the application of off-resonance heteronuclear decoupling [68]. Consider again the AX pair in which continuous A spin decoupling is applied so far from the A resonance that it has no effect; the X spectrum displays a doublet. As the decoupler frequency moves towards the A resonance, the observed coupling begins to collapse until eventually the X spectrum displays only a singlet when decoupling is on-resonance. Measurement of the reduced splitting in the intermediate stages provides a measure of the field strength according to [68]:

$$\gamma B_2 = \left[\left(\frac{J.\nu_{\text{off}}}{J_{\text{r}}}\right)^2 + \frac{\left(J_{\text{r}}^2 - J^2\right)}{4} - \nu_{\text{off}}^2\right]^{1/2} \text{Hz} \tag{3.19}$$

where J is the AX coupling constant, J_{r} the reduced splitting in the presence of the decoupling field and ν_{off} the decoupler offset from the A resonance. When the decoupler field strength is much larger than the decoupler resonance offset, this simplifies to:

$$\gamma B_2 \approx \frac{J.\nu_{\text{off}}}{J_{\text{r}}} \text{Hz} \tag{3.20}$$

Thus, by measuring the resonance splitting in the absence and presence of the irradiation, the field strength is readily calculated. For ^{1}H–^{13}C spectroscopy, the one-bond coupling is suitable for the medium to strong powers, whilst for very weak decoupler fields (<50 Hz) it may be necessary to utilise long-range ^{1}H–^{13}C couplings to obtain significant changes in splitting. The calibration of proton field strength with this method is illustrated in Fig. 3.69.

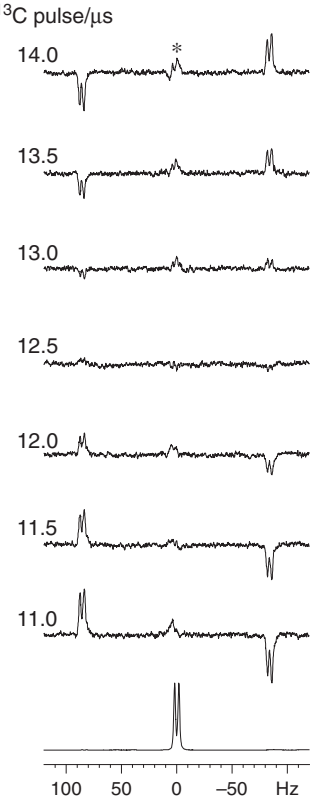


Figure 3.67. ^{13}C Indirect calibration with the 2CALIS-2 sequence performed on 50 mM sucrose. The natural abundance ^{13}C-satellites of the anomeric proton are shown and disappear when $\theta = 90°$, here 12.5 μs. The asterisk indicates the residual ^{12}C parent signal (shown in the lower trace).

3.5.2. Pulsed field gradients

Numerous modern NMR experiments utilise field-gradient pulses for signal selection or rejection (see Chapter 5), requiring these to be applied for well-defined durations at known gradient strengths. Experiments presented in the literature will (or should) state the gradient strengths required to achieve the desired results, so it is necessary to have some knowledge of the strengths provided by the instrument if wishing to implement these techniques. Practical procedures for determining gradient strengths and gradient recovery periods are described here, as it is these two parameters that must be defined by the operator when preparing a gradient-selected experiment. Discussions are restricted to static B_0 gradients rather than the far less widely used rf B_1 gradients [69, 70]. Despite the claimed advantages of B_1 gradients [71], these have yet to gain widespread acceptance.

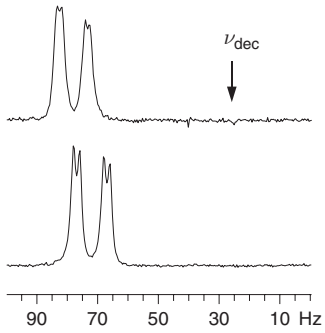


Figure 3.68. Calibration of the homonuclear decoupling field strength via the Bloch–Siegert shift. The decoupler offset from the unperturbed resonance was 47.5 Hz causing a shift of 5.5 Hz, indicating the mean rf field to be 23 Hz.

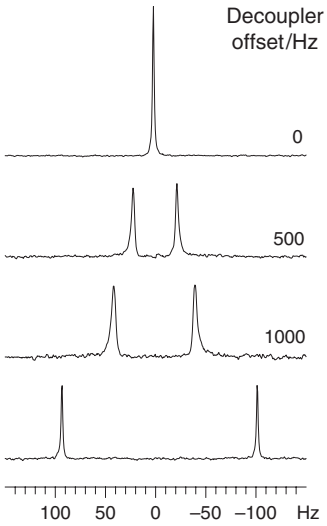


Figure 3.69. Calibration of the heteronuclear decoupler field strength via off-resonance proton decoupling. From the reduced splitting of the ^{13}C–^{1}H doublet, the proton rf field strength is calculated to be 2.2 kHz, a typical value for broadband proton decoupling.

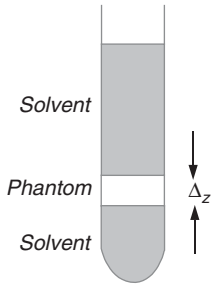


Figure 3.70. A phantom sample suitable for calibrating z axis gradient field strength. The teflon or rubber phantom has precise dimensions (Δ_z, typically 2–4 mm in height) and excludes solvent from a slice of the tube. It is positioned in the tube such that it sits at the centre of the rf coil when placed in the probe.

Gradient strengths

For the application of pulsed field gradients in high-resolution NMR, knowledge of the exact gradient strengths produced is not, in fact, crucial to the success of most experiments. This is in contrast to the application of rf pulses, where precise pulse widths are required. Provided the gradients are *sufficiently strong* for the experiment, the critical factors are pulse reproducibility and the ability to produce gradients pulses of precise amplitude *ratios*. The notable exception to this is in the measurement of self-diffusion coefficients by NMR spectroscopy (Chapter 9) for which very accurate gradient strengths calibrations are required. The procedures described below can be used to calibrate gradient strengths and to check the linearity of the gradient amplifier output.

Essentially there are two approaches to gradient calibration, one based on the measurement of molecular diffusion, the other using an image of a suitable calibration sample. The first requires precise experimental conditions, accurate temperature calibration and a reference solvent or solute with a known diffusion coefficient. Residual H_2O in D_2O is suitable for this, as described in Chapter 9 where diffusion NMR spectroscopy is considered in detail. With the relatively low gradient field strengths used in high-resolution NMR, the alternative imaging approach may be rather more straightforward and less dependent on solution conditions. The calibration of z-gradients is based on the measurement of a 1D image profile of a phantom of known spatial dimensions placed in a standard NMR tube containing H_2O (Fig. 3.70). The phantom is typically a disk 2–4 mm in length of rubber or teflon (a cut-down vortex suppresser for example), which is positioned in the tube such that it sits at the centre of the rf coil. The sample is not spun and need not be locked, in which case the lock field sweep must be stopped. The proton spectrum is first acquired in the absence of the field gradient and the transmitter placed on resonance. The spectrum is then acquired with a very large spectral width, say 50 kHz, in the presence of the field gradient. This can, in principle, be achieved by a simple pulse-acquire scheme with the gradient turned on throughout (Fig. 3.71a), although this approach has its limitations because the application of rf pulses in the presence of the gradient requires excitation over a very wide bandwidth. Thus, it is preferable to record the spectrum with a gradient spin-echo sequence (Fig. 3.71b), in which the second of the two gradients has the same strength but twice the duration of the first and is applied throughout the acquisition period during which the spin-echo refocuses. In either case, the acquisition time must be kept short (10 ms or less) and relatively low gradient strengths applied. The resulting data are processed with a *non-shifted* sine-bell window function (this matches the shape of the echo FID) and magnitude calculation. The spectrum then displays a dip in the profile corresponding to the region of the sample that contains the phantom, and hence no solution, (Fig. 3.72) from which the gradient strength may be calculated as:

$$G_z = \frac{2\pi\Delta\nu}{\gamma\Delta_z} \equiv \frac{\Delta\nu}{4358 \times \Delta_z}\,\mathrm{G\ cm^{-1}} \quad \text{or} \quad \frac{\Delta\nu}{4.358 \times 10^5 \times \Delta_z}\,\mathrm{T\ m^{-1}} \qquad (3.21)$$

where $\Delta\nu$ is the width of the dip in hertz and Δ_z is the height of the phantom in cm. The SI units for a field gradient are $\mathrm{T\,m^{-1}}$, although it is common for values quoted in the literature to be given in $\mathrm{G\,cm^{-1}}$ ($1\,\mathrm{G\,cm^{-1}} \equiv 0.01\,\mathrm{T\,m^{-1}}$). For calibrations of x and y gradients, a similar method may be used, in which case the phantom is not required and the internal diameter of the NMR tube defines the sample width, with the areas outside the tube providing the empty region, and the width of the resulting 'peak' used in the above expression.

Repeating the above procedure with increasing gradient field strengths indicates the linearity of the amplifier, with larger gradients producing a proportionally wider dip in the profile. Figure 3.73 shows the result obtained for a 5 mm inverse z-gradient probe. Assuming linearity over the whole amplifier range, the maximum gradient strength for this system is $0.51\,\mathrm{T\,m^{-1}}$ ($51\,\mathrm{G\,cm^{-1}}$).

Gradient recovery times

As is discussed in Chapter 5, significant distortions in the magnetic field surrounding a sample will result following the application of a gradient pulse, and to acquire a high-resolution spectrum, the eddy currents responsible for these field distortions must be allowed to decay before data are collected. A simple scheme for the measurement of the gradient recovery period is shown in Fig. 3.74, in which the initial gradient pulse is followed by a variable recovery period, after which the data are acquired. Distortions of the spectrum are readily seen by observing a sharp proton singlet, such as that of chloroform. Initially, the recovery period, τ, is set to a large value (1 s) to establish the reference spectrum, and the value progressively decreased until distortions appear in the spectrum

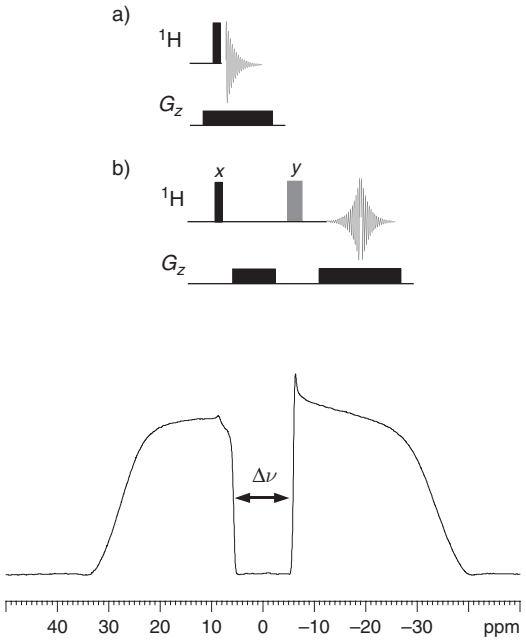


Figure 3.71. Sequences for collecting gradient profiles for the calibration of the z axis gradient field strength. Sequence (a) collects the FID directly in the presence of the applied gradient whilst sequence (b) makes use of a gradient echo (see text). In (b), the second gradient has the same strength but twice the duration of the first.

Figure 3.72. A typical gradient profile collected with the sequence of Fig. 3.71b. Here the dip is 5896 Hz wide, corresponding to a gradient strength of 0.051 T m^{-1} (10% of maximum).

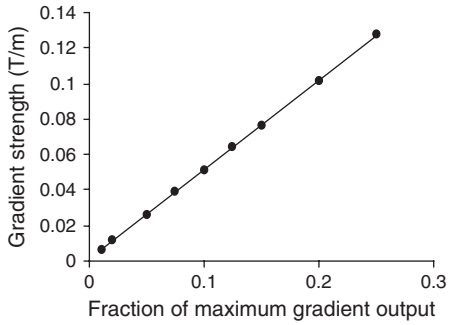


Figure 3.73. Calibration of the gradient amplifier linearity using the sequence of Fig. 3.71b. Assuming system linearity over the full output range and extrapolating the calibration, the maximum gradient strength provided by this system is 0.51 T m^{-1} (51 G cm^{-1}).

(Fig. 3.75). This provides some indication of the recovery period required following a gradient pulse, during which data should not be collected or further gradient or rf pulses applied. The success of high-resolution gradient-selected NMR is critically dependent on being able to use very short recovery periods, which are typically around 100 µs on modern instruments that have actively shielded probe heads. Whilst NMR experiments generally utilise *shaped* gradient pulses (as these help reduce eddy current generation), a more demanding test of hardware performance can be achieved by using *rectangular* gradients, as in Fig. 3.75. Performing this test with a half-sine-shaped gradient pulse showed complete recovery of the resonance within 10 µs.

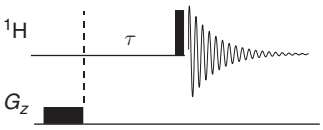


Figure 3.74. A simple scheme for investigating gradient recovery times.

3.5.3. Sample temperature

Many spectrometers are equipped with facilities to monitor and regulate the temperature within a probe head. Usually the sensor takes the form of a thermocouple whose tip is placed close to the sample in the gas flow used to provide temperature regulation. However, the readings provided by these systems may not reflect the true temperature of the sample

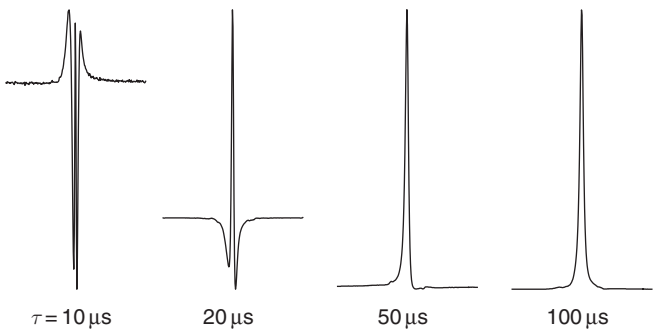


Figure 3.75. Gradient recovery after application of a 1 ms square gradient pulse of 0.25 T m^{-1} (50% of maximum output). The sample is chloroform, and data were acquired with an actively shielded 5 mm inverse z axis gradient probe head.

due to the finite heat capacity of the flowing gas [72] and must be subject to appropriate calibration for accurate temperature recording. One approach to such calibration is to measure a specific NMR parameter that has a known temperature dependence to provide a more direct reading of *sample* temperature. Whilst numerous possibilities have been proposed as reference materials [73], two have become accepted as the standard ^{1}H temperature calibration samples for solution spectroscopy with conventional probe heads. These are methanol for the range 175–310 K and 1,2-ethanediol (ethylene glycol) for 300–400 K. Such strong solutions are inappropriate for cryogenic probes however, and pure perdeuterated methanol offers a better alternative. For use as an *internal* temperature calibrant in ^{13}C measurements, tris(trimethylsilyl)methane is a robust material that may be used in a range of organic solvents, with the shift difference between the ^{13}C resonances of $Si(CH_3)_3$ and CH being indicative of sample temperature; see reference [74] for details of synthesis and calibration.

For the low temperature ^{1}H calibration, neat MeOH is used, with a trace of HCl to sharpen the resonances, for which the following equation holds [75, 76]:

$$T(\mathrm{K}) = 403 - 29.53\Delta\delta - 23.87(\Delta\delta)^2 \tag{3.22}$$

where $\Delta\delta$ is the OH–CH_3 chemical shift difference (ppm). This expression may be approximated to three linear equations of the form:

$$175 - 220\ \mathrm{K} : T(\mathrm{K}) = 537.4 - 143.1\Delta\delta \tag{3.23}$$

$$220 - 270\ \mathrm{K} : T(\mathrm{K}) = 498.4 - 125.3\Delta\delta \tag{3.24}$$

$$270 - 310\ \mathrm{K} : T(\mathrm{K}) = 468.1 - 108.9\Delta\delta \tag{3.25}$$

which are presented as calibration charts in Fig. 3.76. For calibrations with neat ethane-1, 2-diol, the appropriate equation is [76]:

$$T(\mathrm{K}) = 466.0 - 101.6\Delta\delta \tag{3.26}$$

where $\Delta\delta$ now represents the OH–CH_2 chemical shift difference (ppm) (Fig. 3.77). For both samples, the peak shift difference decreases as the sample temperature increases due to the reduced intermolecular hydrogen bonding, which in turn causes the hydroxyl resonance to move to lower frequency. Cryogenic probes are limited in their operating temperature range to within typically 280–330 K, but neat solutions produce too strong a signal for these probes, leading to radiation damping that both broadens and shifts the resonances. A more appropriate calibration sample is neat *perdeuterated* methanol for which the following holds [77] (Fig. 3.78):

$$T(\mathrm{K}) = 419.1381 - 52.5130\Delta\delta - 16.7467(\Delta\delta)^2 \tag{3.27}$$

Before making temperature measurements, the sample should be allowed to equilibrate for at least 5–10 min at each new temperature, and it is also good practice to take a number of measurements to ensure they do not vary before proceeding with the calculation.

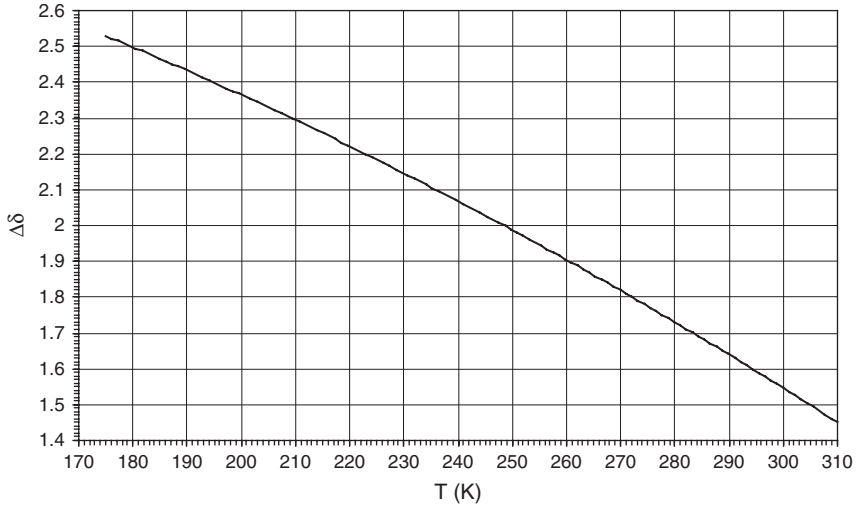


Figure 3.76. Temperature calibration chart for neat methanol. The shift difference $\Delta\delta$ (ppm) is measured between the CH_3 and OH resonances.

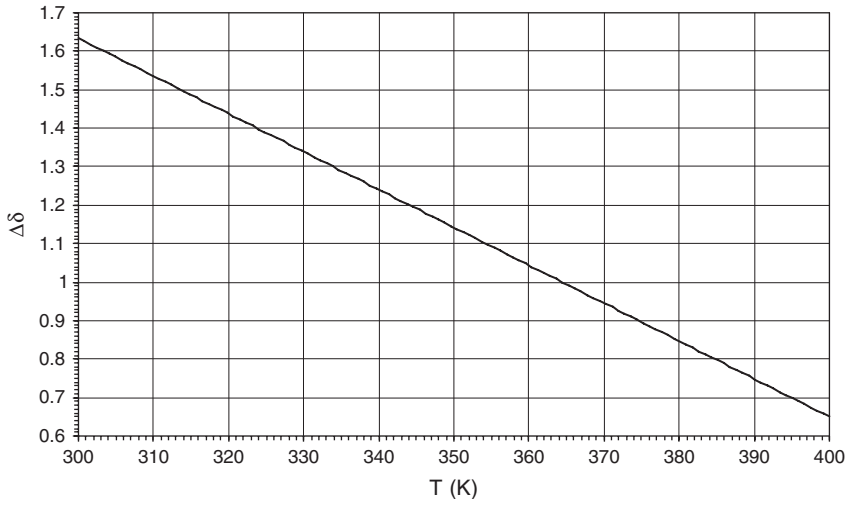


Figure 3.77. Temperature calibration chart for neat ethane-1,2-diol (ethylene glycol). The shift difference $\Delta\delta$ (ppm) is measured between the CH_2 and OH resonances.

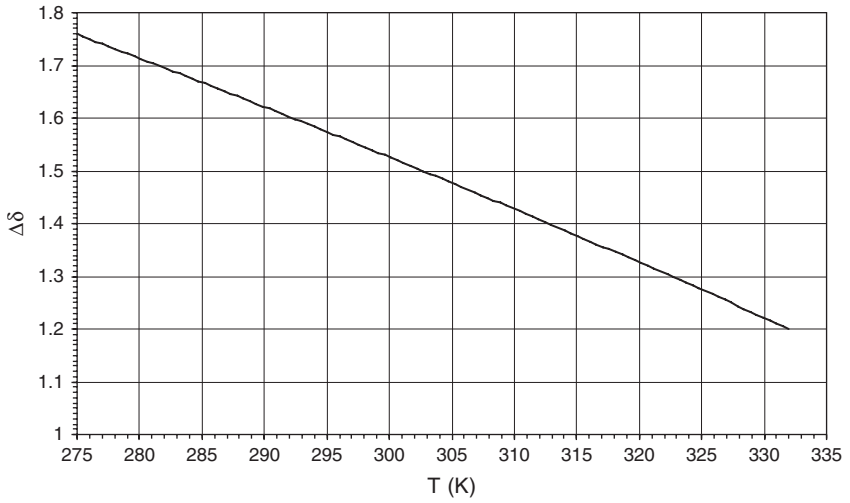


Figure 3.78. Temperature calibration chart suitable for cryogenic probes. The sample is neat d_4-methanol, and the shift difference $\Delta\delta$ (ppm) is measured between the residual CHD_2 and OH resonances.

3.6. SPECTROMETER PERFORMANCE TESTS

There exist a large number of procedures designed to test various aspects of instrument performance, and it would certainly be inappropriate, not to mention rather tedious, to document them in a book of this sort. Instead I want to introduce briefly only those that are likely to be of use in a routine spectrometer 'health check', and as such, this section will be of most interest to anyone with responsibility for the upkeep of an instrument. These tests are generally accepted in the NMR community and by instrument manufacturers since they provide a quantitative measure of instrument performance and, in theory at least, also provides a comparison between the capabilities of different spectrometers. Naturally, this leads to aggressive claims by instrument vendors regarding their results, and some caution is required when comparing advertised figures. This demands attention to the details of procedures used, for example, whether a lineshape test was performed with or without sample rotation or how wide a region of the spectrum was used for estimating the noise level in a sensitivity test. When utilising these tests for in-house checks, consistency in your methodology is the important factor if the results are to be compared with previous measurements. Ideally, these tests should be performed at regular intervals, say every six months or once a year, and appropriate records kept. Additionally more frequent updating of the 'master' shim settings may also be required for optimal performance, although this tends to be very dependent on instrument stability (as well as that of the surrounding environment). One should also be aware that, perhaps not surprisingly, the installation test results achievable on a modern state-of-the-art instrument will certainly be substantially better than the initial test data of an instrument installed some years ago (see, e.g., the discussions on probes in Section 3.4.2 and see Fig. 3.46). The idea of what is an acceptable test result for an older instrument should therefore be scaled accordingly, and reference to previous records become essential if one wishes to gauge instrument performance.

3.6.1. Lineshape and resolution

The most important test, in may be argued, is that for the NMR resonance lineshape, since only when this is optimised can other quantities also be optimum. The test measures the properties of a suitable singlet resonance that, in a perfectly homogeneous field and in the absence of other distortions, produces a precise Lorentzian lineshape, or in other words, its FID possesses an exact exponential decay. Naturally, the only way to achieve this is through careful optimisation of the magnetic field via shimming, which can take many hours to achieve good test results, depending on the initial state of the system. Nowadays, both the lineshape and the resolution measurements are often taken from a single set of test data. The *lineshape* is defined by the width of the resonance at 0.55 and 0.11% of the peak height (these numbers having evolved from measurement of proton linewidths at the height of the ^{13}C satellites and at one-fifth thereof). The *resolution* is defined by the half-height linewidth, $\Delta\nu_{½}$ of the resonance. In fact, by definition, this cannot be a true measure of resolution (how can one measure resolving power with only a single line?), rather the ability of the instrument to separate neighbouring lines is implied by the narrowness of the singlet. For a genuine Lorentzian line, the widths at 0.55 and 0.11% should be 13.5 and 30 times that at half-height respectively, although a check for this is often never performed as attempts are often made (erroneously) to simply minimise all measurements. The tests may be performed on both spinning and static samples (be sure you know which if comparing results), the first being of relevance principally for highest resolution proton measurements whereas the second indicates suitability for non-spinning 1D and 2D experiments.

The proton lineshape test uses chloroform in deuteroacetone typically at concentrations of 3% at or below 400 MHz, and 1% at or above 500 MHz. Older instruments and/or probes of lower sensitivity or observations via outer 'decoupler coils' may require 10% at 200 MHz and 3% at 500 MHz to prevent noise interfering with measurements close to the baseline. A single scan is collected and the data recorded under conditions of high digital resolution (acquisition time of 16 s ensuring the FID has decayed to zero) and processed without window functions. Do not be tempted to make measurements at the height of the satellites themselves unless these are confirmed by measurement to be 0.55%. Since these arise from protons bound to ^{13}C, which relax faster than those of the parent line, they may be relatively enhanced if full equilibrium is not established after previous pulses. The test result for a 400 MHz instrument is shown in Fig. 3.79. The traditional test for proton resolution, which dates back to the continuous-wave (CW) era (*o*-dichlorobenzene in deuteroacetone), is now no longer employed and has become part of NMR history.

Lineshape and resolution tests on other nuclei follow a similar procedure to that above. Not all nuclei available with a given probe need be tested, and typically, only tests for 'inner' and 'outer' coil observations on multinuclear probes are required. This means the second test will often involve carbon-13 for which two samples are in widespread use; the ASTM (American Society for Testing and Materials) test sample (40% *p*-dioxane in deuterobenzene, also used for the sensitivity test) or 80% benzene in deuteroacetone. In either case, on-resonance continuous-wave (CW) decoupling of protons should be used as this provides improved results for a single resonance relative to broadband decoupling. Rather long (30–40 s) acquisition times will be required for a well-shimmed system.

3.6.2. Sensitivity

A great disadvantage of NMR spectroscopy relative to many other analytical techniques is the intrinsically low sensitivity from which it suffers. This, of course, is greatly outweighed by its numerous benefits, yet is still one of the more likely causes of experiment 'failure' and so deserves serious attention; instrumental approaches to enhancing NMR sensitivity are described in Section 3.4.2. The term 'sensitivity' strictly defines a minimum amount of material that is detectable under defined conditions and so represents the limit of

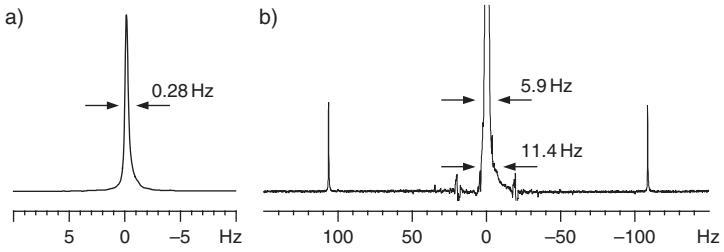


Figure 3.79. Proton resolution and lineshape tests for a static 3% $CHCl_3$ sample recorded on a 400 MHz spectrometer equipped with a dual $^{1}H(^{13}C)$ inverse probe head. The ^{13}C satellites are clearly seen in (b) at ±109 Hz, whereas the artefacts at ±20 Hz arise from floor vibrations.

detection for a measurement, but is used rather loosely throughout NMR and often interchangeably with 'signal-to-noise ratio'. The instrument sensitivity test is indeed a signal-to-noise measurement in which the peak height of the analyte is compared with the noise level in the spectrum. For maximum peak height, optimum lineshape is essential, and generally, the lineshape and resolution test should be performed first and the shim settings re-optimised on the sensitivity sample (for similar reasons, sensitivity tests are traditionally performed with sample spinning). Likewise, the probe must be tuned for the test sample and the 90° pulse determined accurately.

The definition of noise intensity is of prime importance for this measurement since different approaches may lead to differing test results. This has particular significance nowadays as the older method of 'manual' noise estimation has been superseded by automated computational routines provided by instrument vendors. The signal-to-noise measurement compares peak intensity, P, to twice the root mean square (rms) noise level, N_{rms}, in the spectrum. The rms noise is nowadays computed over a defined bandwidth for a region of the spectrum devoid of resonances with a 200 Hz window now most commonly employed. The traditional use of a 2 ppm bandwidth has fallen out of favour with instrument vendors and, perhaps not surprisingly, the narrower bandwidth usually provides a greater signal-to-noise measurement (be sure you know which method is used if comparing results). In cases when a computational determination of N_{rms} cannot be made, this may be estimated from the peak-to-peak noise of the baseline, N_{pp}, whereby N_{rms} is one-fifth N_{pp} as measured directly on screen or from paper plots. The signal-to-noise is then given by

$$\frac{S}{N} = \frac{2.5P}{N_{pp}} \tag{3.28}$$

The peak-to-peak measurement must include all noise bands within the defined frequency window but is somewhat susceptible to what one might call 'operator optimism' in the exclusion of 'spikes' or ''glitches''. Computational methods for determining noise have the advantage of removing such operator bias, and it is consistency of approach that is important for performance monitoring.

The proton sensitivity test uses 0.1% ethylbenzene in deuterochloroform. As for all sensitivity tests, a single-scan spectrum is recorded following a 90° pulse on a fully equilibrated sample (relaxation delay of 60 s in this case). The spectrum is processed with matched filtration corresponding to a line broadening of 1 Hz for the methylene group on which the signal intensity is measured (this resonance is broadened slightly by unresolved long-range couplings to the aromatic protons). Results for a 400 MHz instrument are presented in Fig. 3.80 and show a signal-to-noise ratio of 575:1 or 622:1 with a 2 ppm and a 200 Hz noise bandwidth respectively.

The carbon sensitivity test comes in two guises, without and with proton decoupling. These give rise to quite different results, so it is again important to be aware of which approach has been used. Whilst the first provides an indication of absolute instrument sensitivity, the second represents a more realistic test of overall performance since it also takes into account the efficiency of proton decoupling and is thus more akin to how one performs genuine experiments. Both approaches acquire carbon spectra under atypical conditions of high digital resolution to correctly define peak shapes, demanding acquisition times of around 4 s. The proton-coupled version makes use of the ASTM sample again (40% *p*-dioxane in deuterobenzene) but this time measures peak height of the deuterobenzene triplet. A line broadening of 3.5 Hz is applied and the noise measured over the 80 to 120 ppm window. The proton-decoupled version uses 10% ethyl benzene in deuterochloroform and employs broadband composite pulse decoupling usually via the

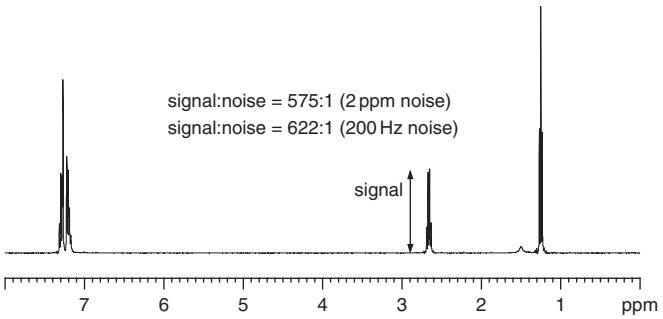


Figure 3.80. The proton sensitivity test for a 0.1% ethyl benzene sample recorded on a 400 MHz spectrometer equipped with a dual $^{1}H(^{13}C)$ inverse probe head.

Table 3.8. Standard sensitivity test samples of some common nuclei

Nucleus	Sensitivity test sample	Notes
^{1}H	0.1% Ethylbenzene in $CDCl_3$	
^{13}C	10% Ethylbenzene in $CDCl_3$	Use broadband proton decoupling
^{13}C	40% dioxane in C_6D_6 (ASTM sample)	No decoupling, C_6D_6 used for measurement
^{31}P	0.0485 M triphenylphosphate in d_6-acetone	No decoupling
^{19}F	0.05% trifluorotoluene in $CDCl_3$	No decoupling
^{15}N	90% formamide in d_6-dimethylsulfoxide	Use inverse gated decoupling to suppress negative NOE
^{29}Si	85% hexamethyldisiloxane in d_6-benzene	No decoupling

WALTZ-16 sequence (Chapter 10). A 0.3 Hz line broadening is used and the noise recorded over the same region, with the peak height determined for the tallest aromatic resonance. Tuning of the proton coil and appropriate calibration of the proton decoupling pulses are required in this case for optimum results. Test samples for other common nuclei are summarised in Table 3.8. Should you have frequent interests in other nuclei, a suitable standard should be decided upon and used for future measurements.

3.6.3. Solvent presaturation

Interest in biologically and medicinally important materials frequently demands NMR analysis to be undertaken on samples in 90% H_2O if solvent-exchangeable protons are also to be observed. The various tests presented so far have all made use of organic solvents; yet, the dielectric properties of these and of water are substantially different and a probe that performs well for, say, chloroform may not be optimum for an ionic aqueous solution, particularly with regard to sensitivity. Accurate tuning of the probe and pulse width calibration play an important role here, but protonated aqueous solutions also demand effective suppression of the solvent resonance if the analyte is to be observed (see Chapter 10 for discussions on suppression methods). The solvent suppression test makes use of solvent presaturation and can be used to measure a number of performance characteristics including the suppression capability, resolution and sensitivity. Good suppression performance places high demands on both static (B_0) and rf (B_1) field homogeneity. A narrow lineshape down to the baseline, and hence good B_0 homogeneity, is again a pre-requisite for good results and ensures all solvent nuclei resonate within a small frequency window. Good B_1 homogeneity means solvent nuclei in all regions of the sample experience similar rf power and are thus suppressed to the same degree. Much of the residual solvent signal that is observed following presaturation arises from peripheral regions of the sample that experience a reduced B_1 field.

The test sample is 2 mM sucrose with 0.5 mM sodium 2,2-dimethyl-2-silapentane-5-sulphonate (DSS) in 90 H_20/10% D_2O, plus a trace of sodium azide to suppress bacterial growth. In the absence of suppression, only the solvent resonance is observed. The test involves on-resonance presaturation of the solvent for a 2 s period, followed by acquisition with a 90° pulse. The presaturation power is selected to attenuate the resonance significantly ($\gamma B_1 \approx 25$ Hz) and this setting used in future comparative tests. Naturally, higher powers will produce greater suppression, but these should not be so high as to reduce the neighbouring sucrose resonances. Typically 8 transients are collected following two "dummy" scans, which achieve a steady-state. The suppression performance is judged by measuring the linewidth of the residual solvent signal at 50 and 10% of the height of the DSS resonance. With probes of recent design that have appropriate screening of the rf coil leads, the 50% linewidth should be somewhat less than 100 Hz and reasonably symmetrical. Much older probes may show poorer performance and may be plagued by wide unsymmetrical humps (despite careful attention to

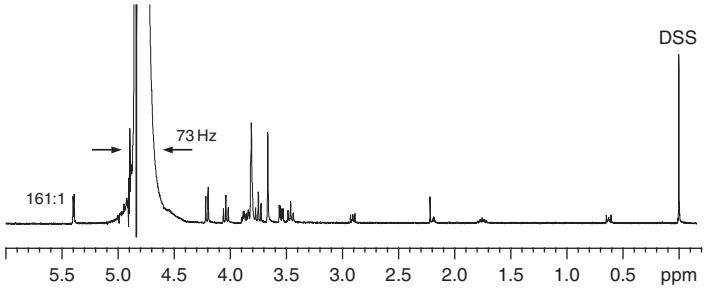


Figure 3.81. The proton solvent presaturation test. The data were acquired on a 400 MHz spectrometer equipped with a dual $^{1}H(^{13}C)$ inverse probe, and the sample was 2 mM sucrose and 0.5 mM DSS in 90% H_2O/10% D_2O.

shim optimisation), which result from signal pick-up in unscreened coil leads [78]. Resolution is judged by the splitting of the anomeric proton doublet at around 5.4 ppm, which should be resolved at least down to 40% of the anomeric peak height. Sensitivity may also be measured from this resonance with the noise determined for the 5.5–7.0 ppm region using either of the methods described above (baseline correction may be required). Test results for a 400 MHz $^{1}H(^{13}C)$ probe using a moderate presaturation power are shown in Fig. 3.81; further reductions in the water resonance could be achieved through the use of greater rf power.

REFERENCES

[1] J.N. Shoolery, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1995, **28**, 37–52.
[2] E. Bartholdi and R.R. Ernst, *J. Magn. Reson.*, 1973, **11**, 9–19.
[3] M.B. Comisarow, *J. Magn. Reson.*, 1984, **58**, 209–218.
[4] J.J. Led and H. Gesmar, *Chem. Rev.*, 1991, **91**, 1413–1426.
[5] J.C. Hoch and A.S. Stern, NMR Data Processing, Wiley-Liss, New York, 1996.
[6] W.F. Reynolds, M. Yu, R.G. Enriquez and I. Leon, *Magn. Reson. Chem.*, 1997, **35**, 505–519.
[7] C.J. Turner and H.D.W. Hill, *J. Magn. Reson.*, 1986, **66**, 410–421.
[8] D.I. Hoult and R.E. Richards, *Proc. Roy. Soc. (Lond.)*, 1975, **A344**, 311–340.
[9] M.A. Delsuc and J.Y. Lallemand, *J. Magn. Reson.*, 1986, **69**, 504–507.
[10] E. Kupce, J. Boyd and I. Campbell, *J. Magn. Reson. (A)*, 1994, **109**, 260–262.
[11] M.E. Rosen, *J. Magn. Reson. (A)*, 1994, **107**, 119–125.
[12] R.R. Ernst, *Adv. Mag. Reson*, 1966, **2**, 1–135.
[13] A.G. Ferrige and J.C. Lindon, *J. Magn. Reson.*, 1978, **31**, 337–340.
[14] G.A. Pearson, *J. Magn. Reson.*, 1987, **74**, 541–545.
[15] D.D. Traficante and G.A. Nemeth, *J. Magn. Reson.*, 1987, **71**, 237–245.
[16] C.R. Pacheco and D.D. Traficante, *J. Magn. Reson. (A)*, 1996, **120**, 116–120.
[17] J.N.S. Evans, *Biomolecular NMR Spectroscopy*, Oxford University Press, Oxford, 1995.
[18] J. Cavanagh, W.J. Fairbrother, A.G. Palmer, N.J. Skelton and M. Rance, *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed., Academic Press (Elsevier), San Diego, 2006.
[19] R.K. Harris, E.D. Becker, S.M. Cabral de Menezes, R. Goodfellow and P. Granger, *Pure Appl. Chem.*, 2001, **73**, 1795–1818.
[20] R.K. Harris and E.D. Becker, *J. Magn. Reson.*, 2002, **156**, 323–326.
[21] R.W. Dykstra, *J. Magn. Reson. (A)*, 1995, **112**, 255–257.
[22] R.J. Lewis, M.A. Bernstein, S.J. Duncan and C.J. Sleigh, *Magn. Reson. Chem.*, 2005, **43**, 783–789.
[23] V. Exarchou, M. Krucke, T.A. van Beek, J. Vervoort, I.P. Gerothanassis and K. Albert, *Magn. Reson. Chem.*, 2005, **43**, 681–687.
[24] K. Albert, On-line LC-NMR and Related Techniques, Wiley, Chichester, 2002.
[25] M.D. Grynbaum, D. Kreidler, J. Rehbein, A. Purea, P. Schuler, W. Schaal, H. Czesla, A. Webb, V. Schurig and K. Albert, *Anal. Chem.*, 2007, **79**, 2708–2713.
[26] D.I. Hoult and R.E. Richards, *J. Magn. Reson.*, 1976, **24**, 71–85.
[27] J.N. Shoolery, In: *Topics in Carbon-13 NMR Spectroscopy,* Vol. 3 (Ed.: G.C. Levy), Wiley Interscience, New York, 1979, pp. 28–38.
[28] R.C. Crouch and G.E. Martin, *Magn. Reson. Chem.*, 1992, **30**, S66–S70.
[29] G.E. Martin and C.E. Hadden, *Magn. Reson. Chem.*, 1999, **37**, 721–729.
[30] G.E. Martin, R.C. Crouch and A.P. Zens, *Magn. Reson. Chem.*, 1998, **36**, 551–557.
[31] G. Schlotterbeck, A. Ross, R. Hochstrasser, H. Senn, T. Kuhn, D. Marek and O. Schett, *Anal. Chem.*, 2002, **74**, 4464–4471.
[32] P.A. Keifer, L. Baltusis, D.M. Rice, A.A. Tymiak and J.N. Shoolery, *J. Magn. Reson. (A)*, 1996, **119**, 65–75.
[33] D.L. Olson, T.L. Peck, A.G. Webb, R.L. Magin and J.V. Sweedler, *Science*, 1995, **270**, 1967–1970.
[34] D.L. Olson, J.A. Norcross, M. O'Neil-Johnson, P.F. Molitor, D.J. Detlefsen, A.G. Wilson and T.L. Peck, *Anal. Chem.*, 2004, **76**, 2966–2974.
[35] A.G. Webb, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1997, **31**, 1–42.
[36] M. Gronquist, J. Meinwald, T. Eisner and F.C. Schroeder, *J. Am. Chem. Soc.*, 2005, **127**, 10810–10811.
[37] F.C. Schroeder and M. Gronquist, *Angew. Chem. Int. Ed.*, 2006, **45**, 7122–7131.
[38] H. Kovacs, D. Moskau and M. Spraul, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2005, **46**, 131–155.
[39] G.E. Martin, *Ann. Rep. NMR. Spectrosc.*, 2005, **56**, 1–96.
[40] P. Styles, N.F. Soffe and C.A. Scott, *J. Magn. Reson.*, 1989, **84**, 376–378.
[41] P. Styles, N.F. Soffe, C.A. Scott, D.A. Cragg, F. Row, D.J. White and P.C.J. White, *J. Magn. Reson.*, 1984, **60**, 397–404.
[42] A.E. Kelly, H.D. Ou, R. Withers and V. Dotsch, *J. Am. Chem. Soc.*, 2002, **124**, 12013–12019.
[43] T.M. de Swiet, *J. Magn. Reson.*, 2005, **174**, 331–334.
[44] N. Bloembergen and R.V. Pound, *Phys. Rev.*, 1954, **95**, 8–12.
[45] W.W. Conover, In: *Topics in Carbon-13 NMR Spectrosocopy,* Vol.4 (Ed.: G.C. Levy), Wiley, New York, 1983, pp. 37–51.
[46] G.N. Chmurny and D.I. Hoult, *Concept. Magn. Reson.*, 1990, **2**, 131–159.
[47] P.C.M. Vanzijl, S. Sukumar, M.O. Johnson, P. Webb and R.E. Hurd, *J. Magn. Reson. (A)*, 1994, **111**, 203–207.

[48] M. Weiger, T. Speck and M. Fey, *J. Magn. Reson.*, 2006, **182**, 38–48.
[49] H. Barjat, P.B. Chilvers, B.K. Fetler, T.J. Horne and G.A. Morris, *J. Magn. Reson.*, 1997, **125**, 197–201.
[50] S. Sukumar, M.O. Johnson, R.E. Hurd and P.C.M. van Zijl, *J. Magn. Reson.*, 1997, **125**, 159–162.
[51] G.A. Morris, *J. Magn. Reson.*, 1988, **80**, 547–552.
[52] G.A. Morris and D. Cowburn, *Magn. Reson. Chem.*, 1989, **27**, 1085–1089.
[53] G.A. Morris, H. Barjat and T.J. Horne, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1997, **31**, 197–257.
[54] K.R. Metz, M.M. Lam and A.G. Webb, *Concept. Magn. Reson.*, 2000, **12**, 21–42.
[55] H. Barjat, G.A. Morris, A.G. Swanson, S. Smart and S.C.R. Williams, *J. Magn. Reson. (A)*, 1995, **116**, 206–214.
[56] T.J. Horne and G.A. Morris, *J. Magn. Reson. (A)*, 1996, **123**, 246–252.
[57] A. Gibbs, G.A. Morris, A.G. Swanson and D. Cowburn, *J. Magn. Reson. (A)*, 1993, **101**, 351–356.
[58] H. Barjat, G.A. Morris, S. Smart, A.G. Swanson and S.C.R. Williams, *J. Magn. Reson. (B)*, 1995, **108**, 170–172.
[59] P.S.C. Wu and G. Otting, *J. Magn. Reson.*, 2005, **176**, 115–119.
[60] J.R. Wesener and H. Günther, *J. Magn. Reson.*, 1985, **62**, 158–162.
[61] E. Haupt, *J. Magn. Reson.*, 1982, **49**, 358–364.
[62] D.M. Thomas, M.R. Bendall, D.T. Pegg, D.M. Doddrell and J. Field, *J. Magn. Reson.*, 1981, **42**, 298–306.
[63] A. Bax, *J. Magn. Reson.*, 1983, **52**, 76–80.
[64] A.J. Benie and O.W. Sorensen, *J. Magn. Reson.*, 2006, **182**, 348–352.
[65] E. Kupce and B. Wrackmeyer, *J. Magn. Reson.*, 1991, **94**, 170–173.
[66] A.J. Benie and O.W. Sorensen, *J. Magn. Reson.*, 2006, **180**, 317–320.
[67] F. Bloch and A. Siegert, *Phys. Rev.*, 1940, **57**, 522–527.
[68] S.D. Simova, *J. Magn. Reson.*, 1985, **63**, 583–586.
[69] W.S. Price, *Annu. Rep. NMR. Spectrosc.*, 1996, **32**, 51–142.
[70] D.G. Cory, F.H. Laukien and W.E. Maas, *J. Magn. Reson. (A)*, 1993, **105**, 223–229.
[71] W.E. Maas, F. Laukien and D.G. Cory, *J. Magn. Reson. (A)*, 1993, **103**, 115–117.
[72] N.M. Loening and J. Keeler, *J. Magn. Reson.*, 2002, **159**, 55–61.
[73] M.L. Martin, J.-J. Delpeuch and G.J. Martin, *Martin, Practical NMR Spectroscopy*, Heydon, London, 1980.
[74] W.H. Sikorski, A.W. Sanders and H.J. Reich, *Magn. Reson. Chem.*, 1998, **36**, S118–S124.
[75] A.L. van Geet, *Anal. Chem.*, 1970, **42**, 679–680.
[76] D.S. Raidford, C.L. Fisk and E.D. Becker, *Anal. Chem.*, 1979, **51**, 2050–2051.
[77] M. Findeisen, T. Brand and S. Berger, *Magn. Reson. Chem.*, 2007, **45**, 175–178.
[78] R.W. Dykstra, *J. Magn. Reson.*, 1987, **72**, 162–167.