Part 2
The NMR Experiment

4 The NMR Spectrometer

The NMR experiment is technically difficult. First, the NMR signal is very weak. Second, the Larmor frequencies must be measured with extremely high accuracy (at least 1 part in 10^9). The twin challenges of *sensitivity* and *resolution* are the keynotes of NMR instrumentation.

The NMR signal is weak for two reasons. First, the individual nuclear magnetic moments are very small compared with those of electrons. Second, the distibution of nuclear magnetic moments is nearly isotropic. As described in Chapter 2, macroscopic nuclear magnetism depends on the very slight thermal imbalance of the distribution of magnetic moments. In ordinary circumstances, this imbalance represents only about 1 part in 10^5. Detection of the weak nuclear magnetism is a considerable instrumental challenge.

The problem of resolution has been mastered, and a great deal of progress has been made on the problem of sensitivity. Nevertheless, sensitivity remains a great limitation of the NMR technique, at least in its conventional form. Rather large amounts of sample are required, compared with other spectroscopies. Around 10^{14} nuclear spins are typically required to obtain a usable NMR signal.[1]

In the following sections, I outline the operation of a typical solution-state NMR spectrometer. Solid-state NMR spectrometers, and NMR imaging instruments, are based on similar principles, although with technical differences that are largely restricted to the probe and magnet.

4.1 The Magnet

Most NMR experiments require a magnetic field which is *homogeneous* (i.e. independent of position) within at least 1 part in 10^9. This extremely high magnetic field homogeneity must be maintained over the entire volume of the sample, i.e. around a cubic centimetre in solution NMR and many hundreds of cubic centimetres in NMR imaging experiments. The magnetic field must also be extremely stable with respect to time.

A perfectly uniform magnetic field avoids inhomogeneous broadening of the NMR signal (see Section 3.6) and allows the resolution of small differences in the nuclear Larmor frequency, due to chemical shifts, spin–spin couplings, or other molecular interactions.

Almost all NMR spectrometers employ *superconducting magnets*. Superconductors are capable of supporting large electric currents without any external source of power. Once charged with current, a superconducting magnet runs almost indefinitely, providing an extremely stable magnetic field with no outside interference.

At present (2007), superconducting materials must be cooled to around the boiling point of liquid He (4.18 K) in order to support large magnetic fields without electrical resistance.[2] The superconducting

Spin Dynamics: Basics of Nuclear Magnetic Resonance, Second Edition Malcolm H. Levitt

magnet windings are typically made from an alloy of Nb and Sn. The heart of an NMR magnet is a Nb–Sn coil immersed in a bath of liquid He. The liquid-He bath is itself insulated by a large reservoir of liquid N_2 (at a temperature of 77 K). The reservoirs are separated from each other, and from the outside environment, by evacuated barriers in order to reduce thermal leakage. The large insulated can of coolant is the visible portion of the magnet from outside:

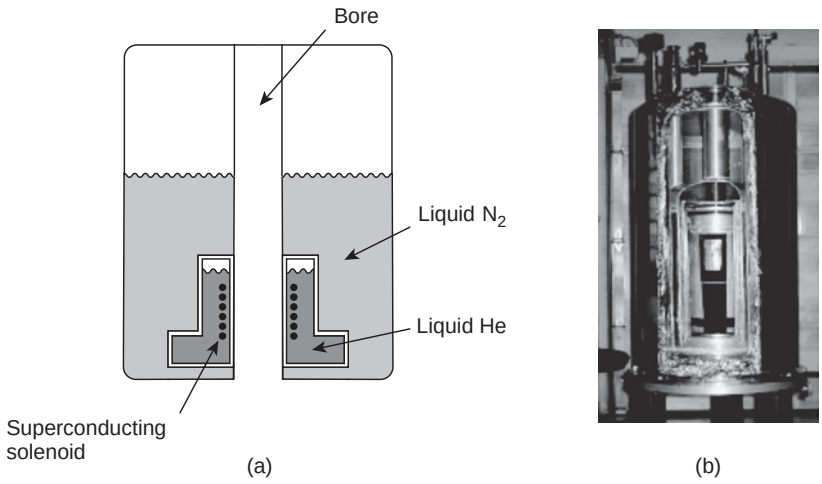


Figure 4.1
(a) Schematic picture of a superconducting NMR magnet. (b) A magnet that has been opened to reveal the superconducting shim coils and the solenoid around the bore (black tube). (Provided by JEOL, USA, Inc.)

Running through the centre of the cylindrical can, and through the centre of the superconducting coil, is a large hole called the *bore*. The sample is mounted in a cylindrical device called a *probe*, which is inserted into the bore so as to position the sample at the point of maximum field. The bore is separated from the superconducting coil by cooled evacuated barriers, allowing the probe and the sample to remain at room temperature.

Currently (2007), the largest commercially-available magnetic field for NMR use is around 22.3 T, corresponding to a proton Larmor frequency of $|\omega^0/2\pi| \cong 950$ MHz.

The magnet is also provided with two sets of additional coils, called *shims*, for adjusting the homogeneity of the magnetic field.[3] One set of coils, called the *superconducting shims*, is wound from superconducting material and immersed in the liquid-He bath. This set of coils is charged when the magnet is installed so as to provide a primary correction for the inhomogeneity of the magnetic field. The second set of shims is supported on a tube that is inserted into the magnet bore. These are called the *room-temperature shims*. Whenever the sample or the probe is changed, the currents in the room-temperature shims are adjusted in order to optimize the field homogeneity (a process called 'shimming the magnet'). In modern instruments, the shimming process is conducted semi-automatically.

4.2 The Transmitter Section

NMR instruments require irradiation of the sample with r.f. waves at the nuclear Larmor frequency and detection of radio signals emitted by the nuclei. The following sections describe the instrumentation used for generating and detecting these radio signals.

The *transmitter section* is that part of the spectrometer which produces the r.f. irradiation. In general, there are several transmitter sections, each dedicated to producing r.f. signals at frequencies close to the Larmor frequencies of different isotopes. For simplicity, the operation of a *single-channel* instrument is described here.

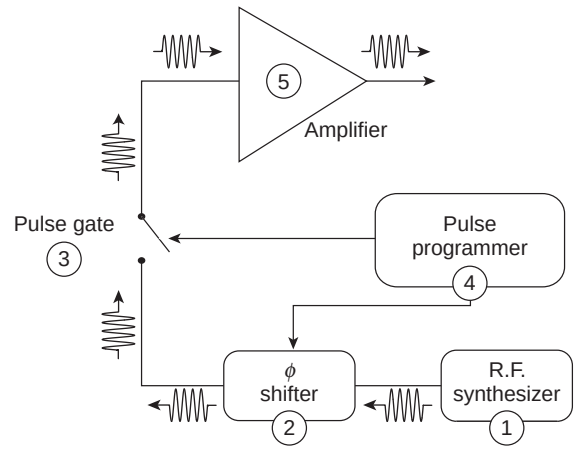


Figure 4.2
The transmitter section.

4.2.1 The synthesizer: radio-frequency phase shifts

The *r.f. synthesizer* ① produces an oscillating electrical signal with a very well-defined frequency. This electronic signal is usually subjected to a number of frequency conversion and electronic filtering steps, which we will not go into here. The result is an r.f. wave that oscillates at the *spectrometer reference frequency*,[4] denoted ω_{ref}.

As usual, the symbol ω specifies the frequency in units of radians per second. The reference frequency in hertz is therefore written $\omega_{\text{ref}}/2\pi$. A reference frequency of 100 MHz corresponds to $\omega_{\text{ref}}/2\pi = 100 \times 10^6$ Hz.

In general, the synthesizer output wave is given by

$$s_{\text{synth}} \sim \cos(\omega_{\text{ref}} t + \phi(t)) \tag{4.1}$$

where $\phi(t)$ is the r.f. phase and t is the time coordinate.

The *phase* of an oscillating wave indicates the position of the oscillation at the time origin $t = 0$:

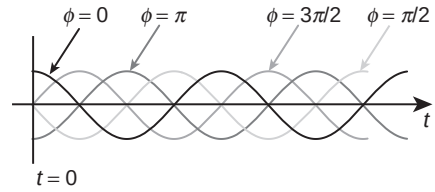


Figure 4.3
Four different phases of a wave.

In many NMR experiments, the r.f. phase $\phi(t)$ is jumped rapidly between different values. In the NMR spectrometer, these discontinuous phase jump events are controlled by a timing device called the *pulse programmer* (④ in Figure 4.2). The following figure shows the synthesizer output for a case in which the phase shift ϕ jumps from an initial value 0 to a value of $\pi/2$, and then back again to 0, after a certain interval:

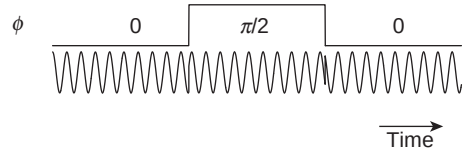


Figure 4.4
Changing the phase and changing it back.

Note the discontinuities in the synthesizer wave.

4.2.2 The pulse gate: radio-frequency pulses

The next item in the transmitter chain is the *pulse gate*, marked ③ in Figure 4.2. This is simply a fast switch that is opened at defined moments in order to allow the r.f. reference wave to pass through. The effect is to chop a 'time-slice' out of the r.f. waveform:

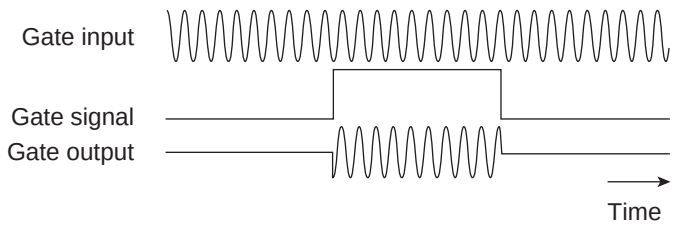


Figure 4.5
Constructing an r.f. pulse.

The duration of an r.f. pulse is sometimes referred to as the *pulse width*. In this book, the pulse duration (i.e. pulse width) is denoted by the symbol τ_p.

The gating events, just like the r.f. phase shifts, are controlled by the pulse programmer. By arranging appropriate timing of the phase shift and the gating events, it is possible to produce r.f. pulses of any desired phase ϕ. Normally, the phase shift is implemented a short time before the opening of the pulse gate and reset a short time after the pulse gate is closed, in order to generate clean transitions in the r.f. waveform:

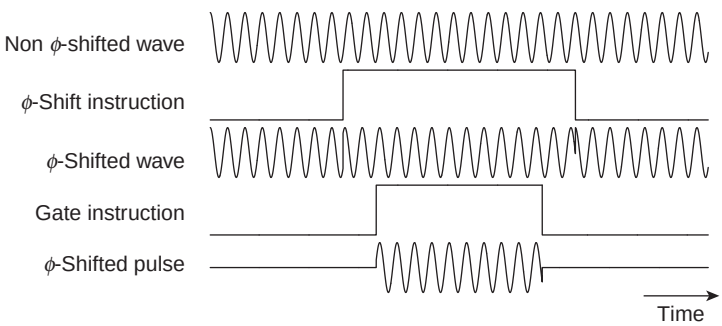


Figure 4.6
Detdetailed timing for an r.f. pulse.

⚠The phase of an r.f. pulse is governed by the underlying function $\cos(\omega_{ref}t + \phi)$ and not by the appearance of the waveform at the beginning of the pulse, which depends on the precise gating instant, a matter of little importance. The two pulses shown below both have phase 0, but look quite different:

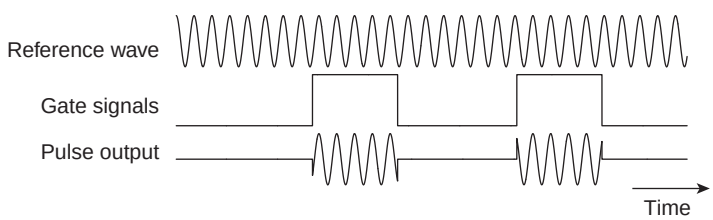


Figure 4.7
Two pulses with the same phase.

NMR spectroscopists often use a special jargon for the four phases that are multiples of $\pi/2$, summarized in Table 4.1. The symbols $\bar{x}$ and $\bar{y}$ are pronounced 'x-bar' and 'y-bar'. The reasons for this jargon are explained in Chapter 10.

The gating scheme described here generates a 'rectangular' r.f. pulse. Ideally, the waveform envelope rises infinitely fast at the beginning of the pulse, remains constant during the pulse, and falls off infinitely fast at

Table 4.1 Jargon used for r.f. phases.

R.f. phase	Jargon
$\phi = 0$	"x-pulse"
$\phi = \pi/2$	"y-pulse"
$\phi = \pi$	"$\overline{x}$-pulse" or "$-x$-pulse"
$\phi = 3\pi/2$	"$\overline{y}$-pulse" or "$-y$-pulse"

the end. There are many more complicated possibilities. Modern NMR experiments often use modulation schemes in which the amplitude, frequency, and phase of the r.f. pulses are varied smoothly.

4.2.3 Radio-frequency amplifier

The function of the r.f. amplifier ⑤ is to scale up the gated waveform so as to produce a large-amplitude r.f. pulse for transmission to the probe. Typical r.f. amplifiers for NMR applications have ratings between several watts and about 1 kW of peak output power.

4.3 The Duplexer

The amplified r.f. pulse travels down a cable into the *duplexer* ⑥. From the duplexer, two more cables emerge. One cable leads to the probe, which is mounted inside the magnet, and which contains the sample. The second cable leads to the receiver section, which is used to detect the weak r.f. signals generated by the nuclear spins.

The duplexer achieves an apparently impossible task: when a strong r.f. pulse arrives from the amplifier, the duplexer diverts it down the cable leading to the probe, not into the sensitive signal detection circuitry:

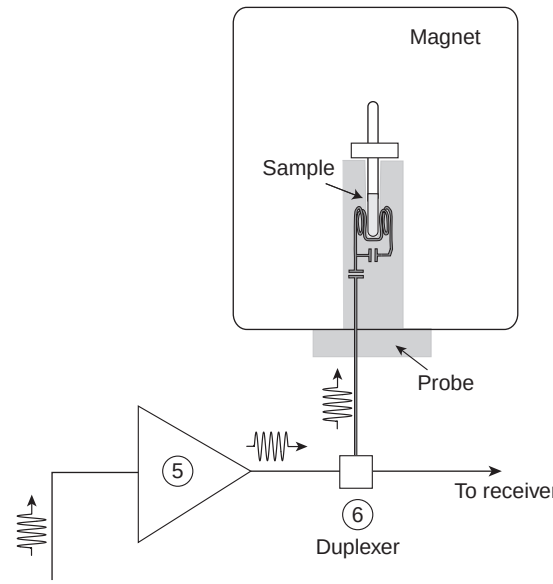


Figure 4.8
The duplexer in transmit mode.

When, on the other hand, the tiny NMR signal travels down the same cable in the opposite direction, i.e. from the probe to the duplexer, it is diverted into the signal detection path, not down the cable leading to the amplifier:

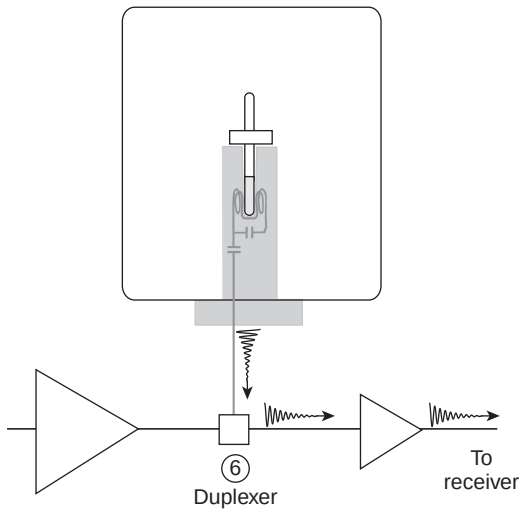


Figure 4.9
The duplexer in receive mode.

On some instruments, this demanding task is accomplished by a cunning arrangement of cables and diodes, without any external switching signals. In other instruments, the duplexer is switched rapidly between transmit and receive mode by a signal from the pulse programmer.

4.4 The Probe

The probe is a complex piece of apparatus that has several functions:

1. It locates the sample in the region of homogeneous magnetic field.

2. It includes r.f. electronic circuits for irradiating the sample with r.f. waves, and for detecting the subsequent r.f. emissions from the sample.

3. In some cases, the probe has a device for rotating the sample, in order to reduce the width of the NMR peaks.

4. In some cases, the probe has devices for stabilization of the temperature of the sample.

5. *Cryoprobes* employ cooling of the electronic circuits (not the sample!) to cryogenic temperatures (typically around 20 K) in order to improve the signal detection efficiency.

6. In some cases, the probe contains additional coils for creating magnetic fields with a controlled spatial inhomogeneity. These coils are used in NMR imaging experiments, and many other methods (see Section 4.7).

The probe is the most specialized part of the NMR spectrometer, and is often the only part that needs to be exchanged when switching between different NMR experiments, e.g. from liquid-state NMR experiments to solid-state NMR experiments.

No attempt is made here to survey the different types of NMR probe. For the present purposes, consider the following highly simplified representation of a liquid-state NMR probe:

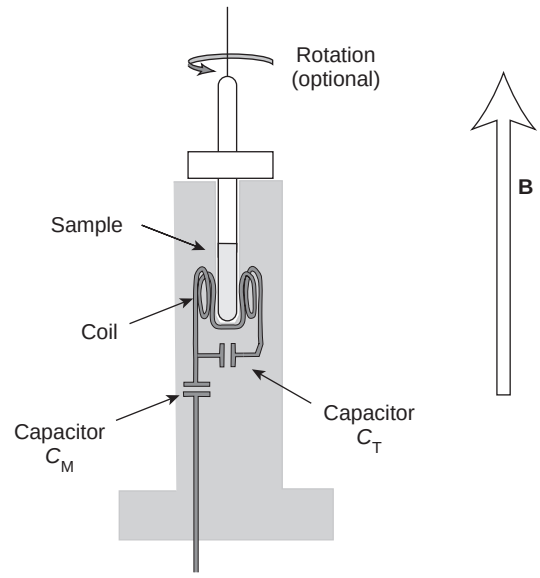


Figure 4.10
A single-channel solution NMR probe.

The sample is contained in a glass tube, which sits in a hole in the top of the probe so as to position the sample inside the r.f. coil. When the probe is mounted in the magnet bore, the sample is located in the region of homogeneous magnetic field.

For maximum resolution, NMR tubes are machined with high precision from very uniform glass, to avoid undesirable distortions of the magnetic field. All probe components in the vicinity of the sample are strictly non-magnetic. Even the electrical coil is usually made from a carefully chosen composite material with near-zero net magnetic susceptibility.

In liquid-state NMR, the sample tube is often rotated around its long axis at a frequency of around 10 Hz. This is a form of *motional averaging* of the nuclear spin interactions. The spatial inhomogeneities in the magnetic field are partially smoothed out by the sample motion, leading to enhanced resolution.

I now discuss the probe electronics. Figure 4.10 shows a highly simplified r.f. circuit consisting of a coil and two capacitors. The function of the coil is twofold. First, during the r.f. pulse, electrical currents in the coil generate an oscillating magnetic field at the sample. As shown later, this oscillating field rotates the nuclear spin polarizations and creates transverse magnetization. Second, when the r.f. pulse has terminated, the precessing nuclear magnetization generates electrical currents in the coil that give rise to a detectable signal travelling back down the cable again towards the duplexer. The duplexer diverts these weak NMR signals towards the receiver electronics.

The coil geometry is such that it generates an oscillating magnetic field that is predominantly *perpendicular* to the main magnetic field. This requires that the coil is *not* wound 'around' the sample tube, but perpendicular to it. The importance of this geometry will become apparent when we go into the theory.

The two capacitors in the probe circuit have the following functions. The one marked C_M is called the *matching capacitor*; this couples the external signals into the probe circuit with maximum efficiency. I will not discuss it further. The capacitor marked C_T is called the *tuning capacitor*; this is more important for our purposes. It is wired in parallel with the coil enclosing the sample.

The function of the tuning capacitor is to enhance the currents in the coil by electromagnetic resonance. According to elementary circuit theory, a parallel circuit of a capacitor of value C_T (in units of farads) and a coil of inductance L (in units of henrys) comprise an electromagnetic *oscillator*. The frequency of the oscillator, in radians per second is given approximately by

$$\omega_{osc} = (LC_T)^{-1/2} \tag{4.2}$$

Energy in the circuit is stored alternately as an electric field between the plates of the tuning capacitor, and as a magnetic field enclosed by the windings of the coil.

The tuned oscillator acts as an *accumulator of energy* for electromagnetic fields with a frequency matching the oscillation frequency ω_{osc}. When the r.f. pulse arrives from the amplifier via the duplexer, it sets up oscillations in the tuned circuit. If the frequency of the pulse $\omega_{carrier}$ is close to the resonant frequency of the circuit ω_{osc}, the oscillations build up. The phenomenon is similar to a child's swing: by giving the swing small pushes at instants matching the natural resonant frequency, a large-amplitude motion is built up. The swing accumulates mechanical energy. Similarly, the tuned circuit accumulates electrical energy, building up a large oscillating magnetic field in the coil. This magnetic field is much larger than would be achieved in the absence of the tuning capacitor C_T.

The reciprocal phenomenon takes place when the NMR signal is detected. In this case, the oscillating magnetic field produced by Larmor precession of the spins drives resonant electrical oscillations in the tuned circuit. The oscillating electrical signal induced by the spin precession is much larger than would be achieved if the circuit were not tuned. The capacitor and coil act together to accumulate the energy of the weak NMR signal.

The electrical properties of the tuned circuit are affected by the nature of the sample. It is necessary, therefore, to adjust the values of the capacitors C_M and C_T every time the sample is changed. Normally, this is done by manual adjustments of the capacitors – a process called 'tuning the probe'.

Once a circuit is tuned, it may only be used for the observation of a single nuclear isotope. R.f. signals intended for different nuclear isotopes have too large a frequency difference to build up appreciable amplitude in the same tuned circuit. However, it is technically feasible to share the same coil between separate tuned circuits, each operating on a different frequency. This is called *multiple tuning*, and requires careful electronic construction to isolate each circuit from the others. A multiple-tuned coil is simultaneously involved in circuits resonating at a number of different frequencies. Each circuit may be tuned independently to resonate at the Larmor frequency of a particular nuclear isotope.

4.5 The Receiver Section

We now trace the fate of the r.f. NMR signal generated by the nuclear spins, using instrumentation typical for a relatively old-fashioned NMR instrument. Modern instruments employ a more sophisticated and integrated processing of the NMR signal, but the principles are basically the same.

The immediate fate of the NMR signal is as follows:

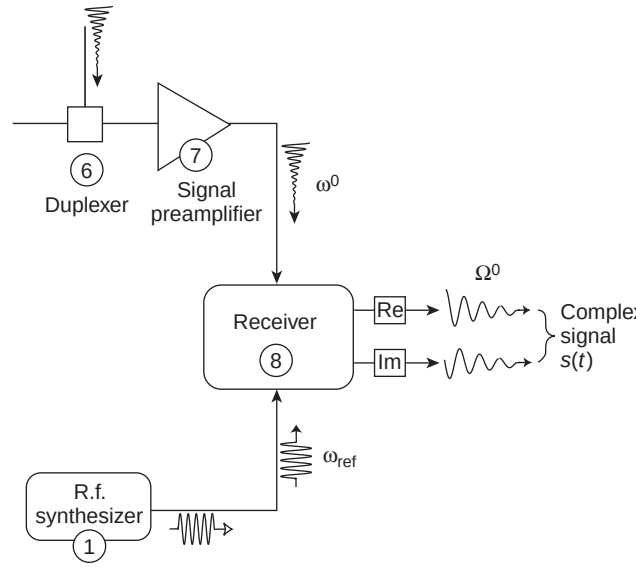


Figure 4.11
The receiver section.

4.5.1 Signal preamplifier

The NMR signal first arrives at the duplexer ⑥, which diverts it down the cable towards the *signal preamplifier* (often simply called the *preamplifier*) ⑦. This is a low-noise r.f. amplifier that scales up the tiny signal to a more convenient voltage level.

4.5.2 The quadrature receiver

The NMR signal must be fed into the computer for interpretation and presentation. Therefore, it is necessary to convert the NMR signal (an oscillating electrical current) into digital form (a sequence of 'ones' and 'zeros').

The process of converting a continuous current or voltage into digital form is called *analogue-to-digital conversion*. This task is accomplished by specialized electronic circuits called *analogue-to-digital converters* (ADCs).

One technical problem must first be faced. The 'raw' NMR signal oscillates at many hundreds of megahertz, which is too fast for current ADCs (as of 2007). Therefore, it is necessary to 'down-convert' the frequency of the NMR signals, so that they may be handled by the digital electronics at the next stage.

The *quadrature receiver* ⑧ accomplishes the signal frequency conversion by comparing it with a reference wave of frequency ω_{ref}, as supplied by the r.f. synthesizer (see Equation 4.1). The quadrature receiver combines the NMR signal, which oscillates at the Larmor frequency ω^0, with the reference signal, oscillating at the frequency ω_{ref}, to generate a new signal that oscillates at the *relative Larmor frequency*:

$$\Omega^0 = \omega^0 - \omega_{\text{ref}}$$

This is the frequency conversion process encountered in Sections 3.4 and 3.5.

The offset frequency Ω^0 is usually of the order of 1 MHz or less. The signal emerging from the quadrature receiver varies slowly enough to be handled accurately by the digital electronics at the next stage.

A similar procedure takes place in an ordinary *radio receiver*. In that case, the modulated r.f. waves, travelling through space, are detected and down-converted to the audible frequency range, where they are transformed into mechanical oscillations to generate sound.

Usually, *two* output signals emerge from the receiver. The use of two outputs, rather than one, will now be motivated.

Consider the r.f. signal from nuclear spins with Larmor frequency ω^0 (for simplicity, suppose that there is only one peak in the NMR spectrum). The NMR signal (FID) s_{FID} has the form

$$s_{\mathrm{FID}}(t) \sim \cos\left(\omega^0 t\right) \exp\{-\lambda t\}$$

including a damping factor with the rate constant $\lambda = {T_2}^{-1}$.

Now consider the output of the receiver. The simplest form of frequency down-conversion would substitute ω^0 by Ω^0, leading to

$$\cos\left(\Omega^0 t\right) \exp\{-\lambda t\} \tag{4.3}$$

However, there is a problem. Equation 4.3 does not distinguish between signals generated by spins whose precession frequencies ω^0 are *larger* than ω_{ref} from those whose precession frequencies ω^0 are *smaller* than ω_{ref}. For example, suppose that protons are observed and that the proton Larmor frequency is close to -500 MHz. Suppose that the spectrometer reference frequency is set to exactly $\omega_{\mathrm{ref}}/2\pi = -500.000\,000$ MHz. The precise precession frequency of the nuclei depends on the molecular environment. Nuclear spins in environments with slightly stronger magnetic fields precess slightly faster, e.g. $\omega^0/2\pi = -500.001\,000$ MHz; nuclear spins in environments with slightly smaller magnetic fields precess slightly slower, e.g. $\omega^0/2\pi = -499.999\,000$ MHz. The difference frequency for spins of the first type is $\Omega^0/2\pi = -1.000$ kHz, whereas the difference frequency for spins of the second type is $\Omega^0/2\pi = +1.000$ kHz. Equation 4.3 does not distinguish the signals from these physically distinct situations.

To resolve this ambiguity, the receiver supplies *two* output signals, of the form

$$s_{\mathrm{A}}(t) \sim \cos\left(\Omega^0 t\right) \exp\{-\lambda t\}$$

$$s_{\mathrm{B}}(t) \sim \sin\left(\Omega^0 t\right) \exp\{-\lambda t\}$$

These two signals may be interpreted as the real and imaginary components of a single complex signal $s(t)$, i.e.

$$s_{\mathrm{A}}(t) = \mathrm{Re}\{s(t)\} \quad s_{\mathrm{B}}(t) = \mathrm{Im}\{s(t)\}$$

where

$$s(t) = s_{\mathrm{A}}(t) + \mathrm{i}s_{\mathrm{B}}(t) \sim \exp\{(\mathrm{i}\Omega^0 - \lambda)t\} \tag{4.4}$$

This complex NMR signal distinguishes between positive and negative values of Ω^0 and retains full information as to the magnitude of the Larmor frequency relative to that of the reference wave.

This two-output scheme is called *quadrature detection*. The electronic configuration and operation of a typical quadrature detector is discussed in Appendix A.5.

The representation of an NMR signal as *complex*, with both a real and an imaginary part, is mathematically very convenient. NMR signals with a *decaying complex exponential* form, as in Equation 4.4, will be encountered a lot from now on.

4.5.3 Analogue–digital conversion

Each of the two outputs of the quadrature receiver ⑧ are connected to their own ADC ⑨:

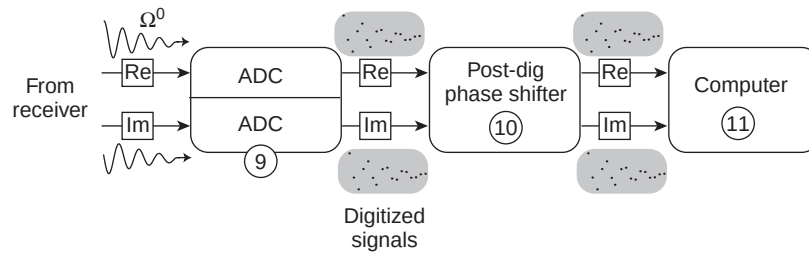


Figure 4.12
The digitizer and phase shifter chain.

An ADC is an electronic circuit that rapidly measures the voltage level of an input signal and presents the information as a string of 'ones' and 'zeros'. The NMR signal is converted to digital form by repeating the measurement at a set of consecutive time points and storing the information in the computer as an ordered set of values. This is called *digitization*:

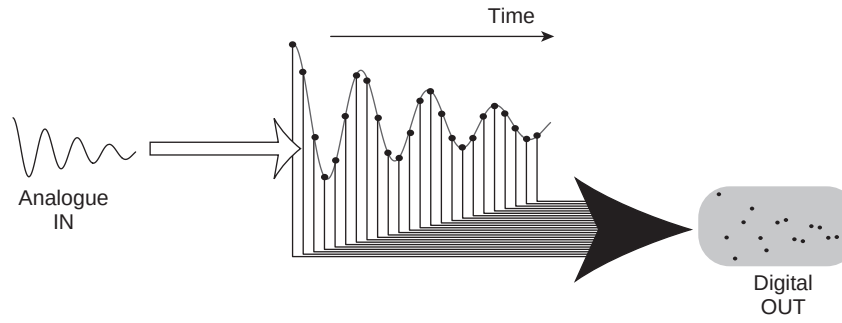


Figure 4.13
Digitization.

The time separation between the sampling points of the ADCs is called the *sampling interval*, or the *dwell time* in older literature (symbol τ_{sample}). The inverse of the sampling interval is the *sampling bandwidth*, often referred to (somewhat misleadingly) as the *spectral width*. The sampling bandwidth sets the maximum range of signal frequencies that is represented accurately by the sampling process (see *Further Reading*). NMR signals with a wide range of frequencies (common in solid-state NMR) require more rapid sampling than signals that span a narrow frequency range (common in liquid-state NMR). Typical sampling bandwidths are around 4 MHz in solid-state NMR and around 250 kHz for liquid-state NMR. This corresponds to sampling intervals τ_{sample} of 250 ns and 4 μs respectively.

For technical reasons, the number of sampled points n_{sample} is usually an integer power of 2. There is a special jargon for the larger powers, as summarized in Table 4.2. For example, $n_{\text{sample}} = 2^{10} = 1024$ is referred to as '1k', etc.

Table 4.2 Jargon used for the number of digital sampling points.

Number of points	Power of 2	Jargon
1 024	10	1k
2 048	11	2k
4 096	12	4k
8 192	13	8k
16 384	14	16k
32 768	15	32k

The total duration over which the signal is sampled is called the *acquisition time* and is given by

$$\tau_{\text{acq}} = n_{\text{sample}}\tau_{\text{sample}} \tag{4.5}$$

For reasons of sensitivity and resolution, it is usually desirable to sample the NMR signal until it has died out completely. The digitization of 32k points, at 20 μs per point, covers a time-span of about 0.65 s, which is adequate in most circumstances (see *Further Reading*).

4.5.4 Signal phase shifting

In many experiments, the phases of the r.f. pulses, and that of the NMR signal, are changed in a dynamic way as the experiment proceeds. As discussed later, this allows NMR signals to be distinguished from experimental artefacts, and also allows different types of NMR signal to be distinguished from each other.

The phases of r.f. pulses are changed as described in Section 4.2.2. The phase of the NMR signal may also be changed as it passes through the receiver/digitizer chain. There are two ways to give the signal a controllable phase shift. Both are in common use.

1. *Receiver reference phase.* As described above, the quadrature receiver compares the NMR signal with a reference wave from the synthesizer. If the phase of the synthesizer reference wave is changed during the entire period of signal detection, then this phase shift is transferred to the signal emerging from the receiver.

2. *Post-digitization phase.* The second method for changing the signal phase operates on the *digitized* signal emerging from the ADCs. The digitized complex signal is passed into a device called a *post-digitization phase shifter* ⑩, which multiplies the signal by a complex phase factor before it is passed to the computer ⑪.

In modern instruments, both types of phase shifting are essentially equivalent and I will not distinguish between them.[5] The overall phase shift imparted to the NMR signal as it passes through the receiver–digitizer chain is called the *receiver phase* and is denoted ϕ_{rec}.

Mathematically, the application of a receiver phase ϕ_{rec} multiplies the complex NMR signal by the factor $\exp\{-i\phi_{\text{rec}}\}$. This property will be used later on, when we analyse some actual NMR experiments.

4.6 Overview of the Radio-Frequency Section

Figure 4.14 gives an overview of the r.f. generation and data processing pathway of a single-channel NMR spectrometer.

The different electronic circuits are orchestrated by the pulse programmer, which receives its directions from the computer. The computer is, in turn, controlled by the spectrometer operator.

The digitized NMR signal is returned to the computer, where it may be processed and displayed, as described in Chapter 5.

A block diagram of a multiple-channel spectrometer, capable of excitation and detection at the Larmor frequency of several isotopes simultaneously, may be derived from this one by duplicating the r.f. network ①–②–③–⑤–⑥–⑦–⑧ shown in Figure 4.14. The probe is the only r.f. component that is not duplicated: as described in Section 4.4, a multiple-channel probe accomodates many r.f. channels at the same time. Sometimes, the receiver is not duplicated either, but assigned to one of the r.f. channels before the experiment starts.

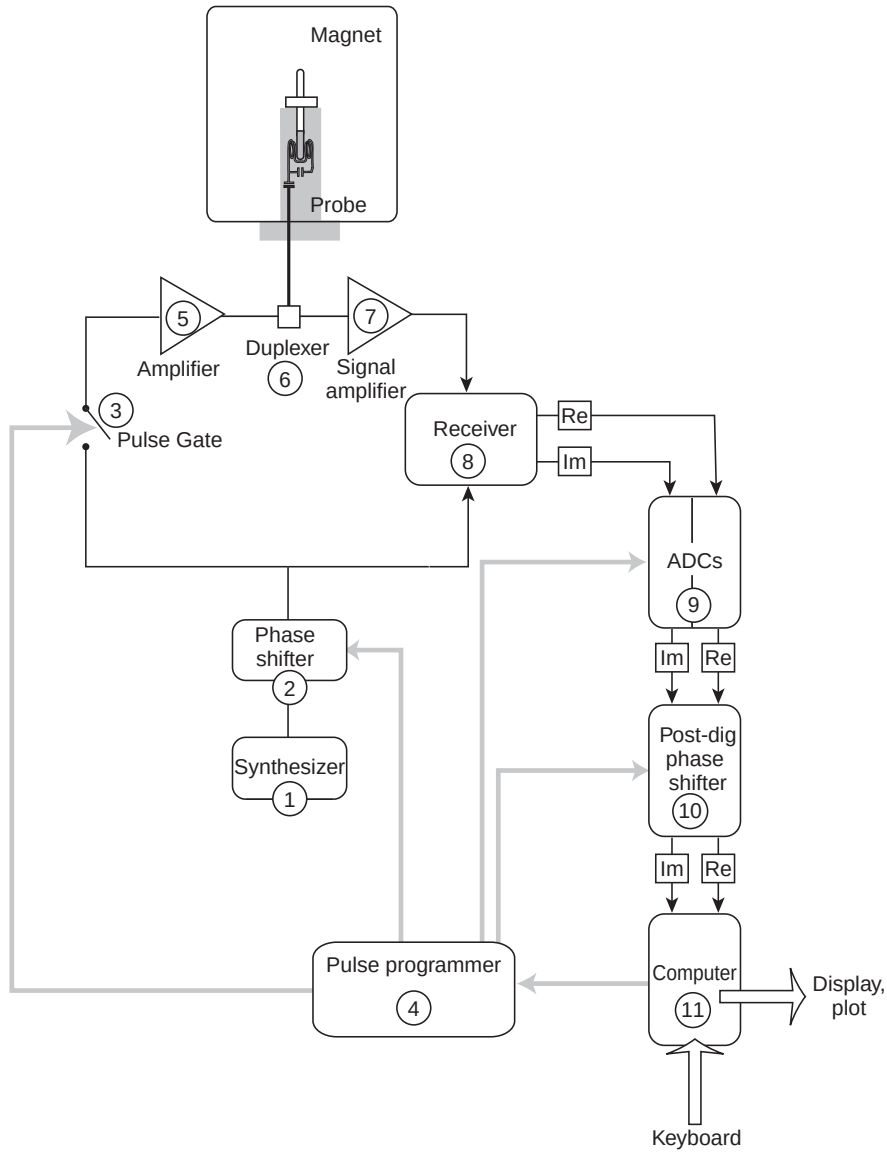


Figure 4.14
Schematic overview of a single-channel spectrometer, without pulsed field gradient equipment.

4.7 Pulsed Field Gradients

Many NMR experiments require the application of pulsed static magnetic fields, as well as r.f. fields. As described in Section 3.6, NMR spectra taken in the presence of a magnetic field gradient contains information on the spatial distribution of nuclei; in Section 12.6, we will see how the switching of field gradients along different axes may be used to construct full three-dimensional images. Switched magnetic field gradients are also used for the study of flow and diffusion by NMR (see Chapter 19), and for selecting certain classes of NMR signals while suppressing others (see Appendix A.12). Pulsed magnetic field gradients may also

be used to perform many different NMR experiments at the same time in different parts of the sample tube.[6]

4.7.1 Magnetic field gradients

In the absence of a magnetic field gradient, the magnetic field is homogeneous. It has the same magnitude, and same direction, at all points in space. If the strength of the magnetic field is B^0, and its direction is defined as the z-axis, the magnetic field at point $\mathbf{r}$ is

$$\mathbf{B}(\mathbf{r}) = B^0\mathbf{e}_z$$

where $\mathbf{e}_z$ is a unit vector along the z-axis.

In the presence of a field gradient, on the other hand, the magnitude of the field varies in a controlled fashion along the gradient direction. Consider, for example, a field gradient of magnitude G_x applied along the x-axis, i.e. perpendicular to the main magnetic field. The magnetic field at point $\mathbf{r}$ in the presence of this gradient is

$$\mathbf{B}(\mathbf{r}) = B^0\mathbf{e}_z + G_x x\mathbf{e}_z \tag{4.6}$$

where x is the coordinate along the x-axis, i.e.

$$x = \mathbf{r} \cdot \mathbf{e}_x$$

Note carefully that Equation 4.6 describes a magnetic field that still points along the z-axis, but whose *magnitude* varies as a function of x. The magnetic field in the presence of G_x may be visualized as follows:

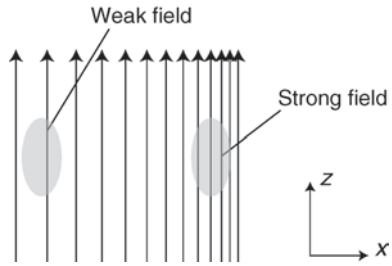


Figure 4.15
Magnetic flux lines in the presence of a strong magnetic field along the z-axis and a field gradient along the x-axis.

The magnitude of the field is proportional to the density of field lines. Note the variation in the strength of the field, but not its direction, along the x-axis.

For clarity, the sketch in Figure 4.15 greatly exaggerates the gradient. In practice, the variation of the magnetic field across the sample is many orders of magnitude less than the size of the main magnetic field.

Similarly, the magnetic field in the presence of a gradient G_z along the z-axis is described by

$$\mathbf{B}(\mathbf{r}) = B^0\mathbf{e}_z + G_z z\mathbf{e}_z \tag{4.7}$$

where $z = \mathbf{r} \cdot \mathbf{e}_z$. The magnetic field lines in the presence of G_z are visualized in Figure 4.16. This time, the magnitude of the magnetic field varies along the z-axis, while its direction remains almost unchanged.[7] Once again, the magnitude of the gradient is greatly exaggerated in this sketch.

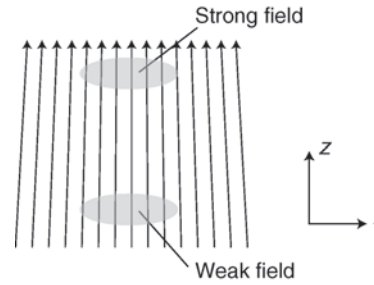


Figure 4.16
Magnetic flux lines in the presence of a strong magnetic field along the z-axis and a field gradient along the z-axis.

A field gradient G_y along the y-axis is described by the equation

$$\mathbf{B}(\mathbf{r}) = B^0\mathbf{e}_z + G_y y\mathbf{e}_z \tag{4.8}$$

where $y = \mathbf{r} \cdot \mathbf{e}_y$. It may be visualized by constructing a picture as in Figure 4.15, but where the field lines get closer as one 'goes into the paper'.

4.7.2 Field gradient coils

Magnetic field gradients are imposed in NMR experiments by passing electrical currents through carefully designed coils near the sample. In some cases, these gradient coils are incorporated into the NMR probe itself; in other cases, they are part of the magnet and shim assembly.

For example, a gradient G_z may be generated by passing currents through opposed coils above and below the sample:

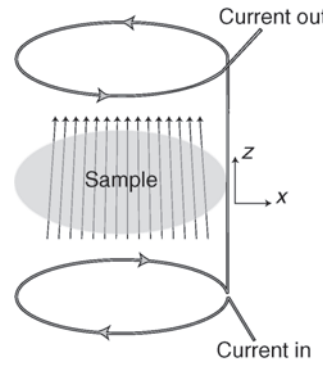


Figure 4.17
Coil and current configuration used to generate a magnetic field gradient along the z-axis. The flux lines shown in the sample region assume that a strong homogeneous magnetic field is applied at the same time along the z-axis.

The field generated by the upper coil enhances the magnetic field in the sample, while the field generated by the lower coil opposes it. Careful design of the geometry and current paths can lead to an accurate linear field gradient G_z across the sample volume.

Similarly, a gradient G_x may be generated by the coil configuration sketched in Figure 4.18. The currents running down each arm of the coils generate a field that enhances the magnetic field on the right-hand side of the sample, but which opposes it on the left-hand side.

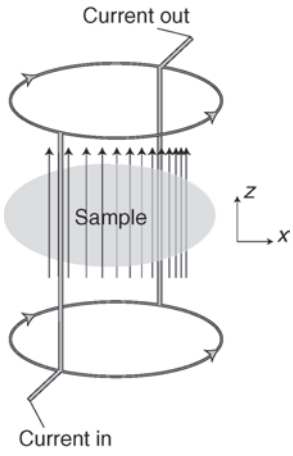


Figure 4.18
Coil and current configuration used to generate a magnetic field gradient along the *x*-axis. The flux lines shown in the sample region assume that a strong homogeneous magnetic field is applied at the same time along the *z*-axis.

A gradient along the *y*-axis is generated by rotating the coil arrangement in Figure 4.18 by 90° around the *z*-axis.

Field gradient coils are usually carefully shielded, in order to minimize the magnetic fields they generate in regions remote from the sample. This is because pulsed magnetic fields induce transient electric 'eddy currents' in metal parts of the probe and magnet assembly. Such eddy currents take some time to dissipate and themselves generate magnetic fields that can interfere with the NMR experiment.

Typical field gradient strengths in human-scale NMR imaging applications are around ~10 mT m^{-1}. For diffusion studies, and microimaging applications, gradients of up to ~1000 mT m^{-1} are used.

4.7.3 Field gradient control

In most spectrometers, each magnetic field gradient is controlled by a separate channel, which resembles that used in the r.f. section. Each gradient channel consists of a *digital-to-analogue converter* (DAC) and a *gradient driver*. The DAC converts timed digital instructions from the pulse programmer into an analogue voltage level. It therefore performs the inverse function of the ADCs used in the receiver circuit. The gradient driver is a powerful amplifier that converts the analogue input voltage into a strong, well-controlled current. This electric current is driven through the gradient coils, generating the desired magnetic field gradient pulse (see Figure 4.19). These circuits allow each gradient pulse to be shaped accurately as a function of time.

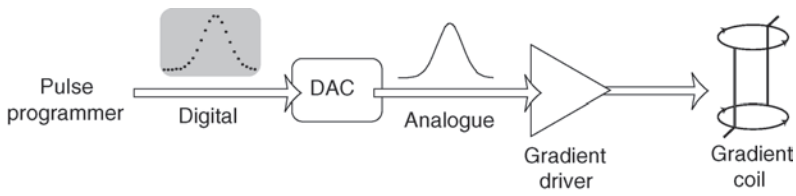


Figure 4.19
The circuits involved in driving a gradient coil.

An overview of a spectrometer with a single r.f. channel and one pulsed field gradient channel is given in Figure 4.20. The pulsed field gradient channels are duplicated if the probe or magnet assembly contains more than one pulsed field gradient coil, and the r.f. channels are duplicated if the probe allows irradiation and/or detection at the Larmor frequency of more than one nuclear isotope.

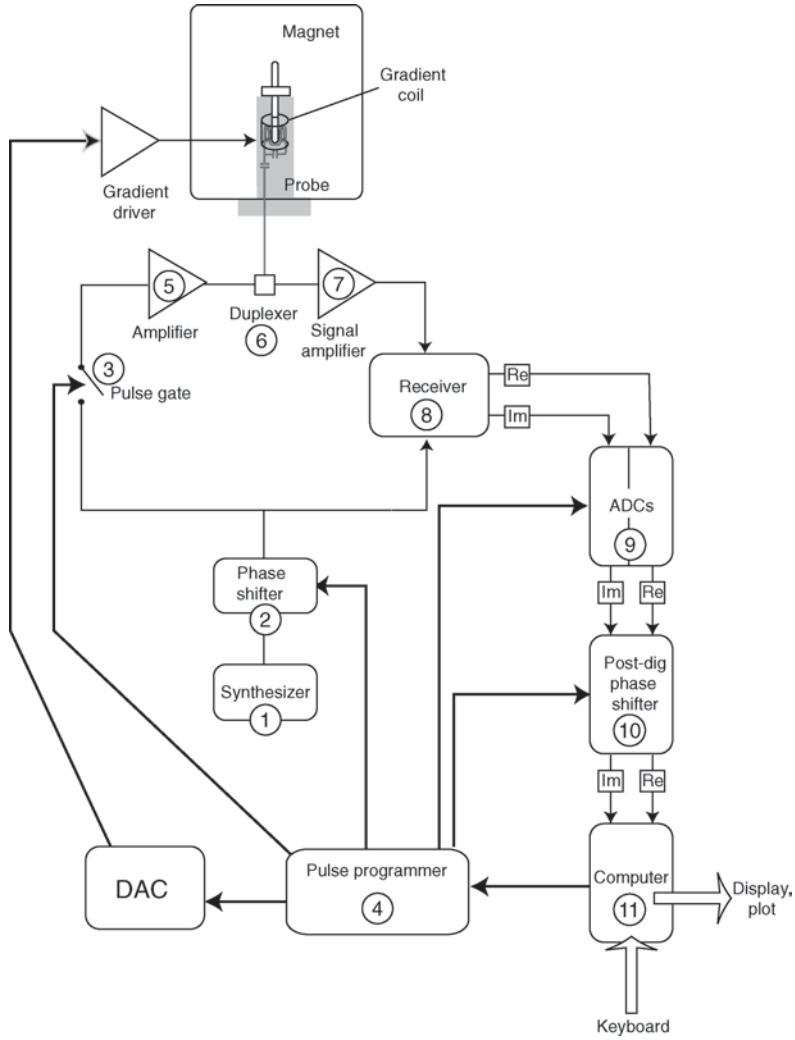


Figure 4.20
Schematic overview of a single-channel spectrometer, incorporating one pulsed field gradient channel.

Notes

1. Various methods are available for transcending the sensitivity limitations of conventional NMR. These schemes are aimed either at enhancing the net polarization of the nuclei or at improving the detection strategy, or both. In *dynamic nuclear polarization* (DNP), the nuclear spin polarization of paramagnetically doped materials is greatly enhanced by using microwave irradiation (e.g. see D. A. Hall *et al., Science*, **276**, 930 (1997) and J. H. Ardenkjaer-Larsen *et al., Proc. Nat. Acad. Sci. USA,* **100,** 10 158–10 163 (2003). Other polarization enhancement techniques include optical pumping of noble gas nuclei (e.g. see B. M. Goodson, *J. Magn. Reson.* **155**, 157–216 (2002)), chemical reactions of parahydrogen (e.g. see S. B. Duckett, C. J. Sleigh, *Prog. NMR Spectrosc.*, **34**, 71 (1999)), and photo-induced radical-pair creation (M. G. Zysmilich, A. E. McDermott, *Proc. Natl. Acad. Sci. USA,* **93**, 6857(1996)). Alternative detection methods have been proposed, such as superconducting quantum

interference devices (SQUIDs) (see C. Connor, *Adv. Magn. Opt. Reson.*, **15**, 201 (1990)) and the rotation of polarized light by the bulk nuclear magnetization (see I. M. Savukov *et al.*, *Nature* **442**, 1021–1024 (2006)). Mechanical oscillators are already sensitive enough to detect single *electron* spins (see D. Rugar *et al.*, *Nature*, **430**, 329–*332* (2004)), and have been used to detect small numbers of nuclear spins (see D. Rugar *et al.*, *Nature*, **360**, 563 (1992)). The magnetic properties of individual nuclear spins have been probed by optical techniques (see J. Kohler *et al.*, *Science*, **268**, 1457 (1995), J. Wrachtrup *et al.*, *Nature*, **363**, 244 (1993), and F. Jelezko *et al.*, *Phys. Rev. Lett.*, **93**, 130501 (2004). None of these methods has proven generally applicable, although many are very promising in specialized circumstances.

2. At the moment (year 2007) the materials known as high-T_c superconductors have not proven to be suitable for the main coils of NMR magnets.

3. The English word *shim* means a small piece of metal. The field adjustment coils are called shim coils because the original method for adjusting the field inhomogeneity involved moving small metal pieces around. Apart from the shims, there is another coil that is used for compensating any small undesired changes in the magnitude of the magnetic field, due to fluctuations in air pressure and temperature, slight imperfections in the superconducting material, or other magnetic disturbances. A feedback circuit continuously monitors the magnetic field in the sample and compensates any changes by increasing or decreasing the current in the extra coil. In order for this to work, this compensation circuit must obtain accurate, up-to-date, information as to the actual magnetic flux density in the sample. Such information can only be provided with sufficient accuracy by one method, namely NMR itself. Commercial spectrometers often run a separate NMR experiment on ^{2}H spins, deliberately introduced into the sample, in order to monitor the magnetic field. This feedback system is called the *field-frequency lock*. It normally runs 'in the background', without intervention from the operator.

4. There are some subtle issues concerning the signs of the frequencies and phase shifts used in the spectrometer. For the sake of simplicity, I have ignored these complications. A complete treatment may be found in M. H. Levitt, *J. Magn. Reson.*, **126,** 164 (1997) and M. H. Levitt, O. G. Johannessen, *J. Magn. Reson.*, **142,** 190–194 (2000).

5. Until recently, it was advisable to distinguish between the receiver reference phase and the post-digitization phase, since they have slightly different characteristics: (i) the post-digitization phase shift occurs *after* the quadrature detection of the NMR signal and is therefore immune to instrumental imperfections such as misbalance of the two sections of the quadrature detector; (ii) in older instruments, the choice of post-digitization phase shifts was frequently limited to multiples of $\pi/2$; (iii) in some NMR instruments, the hardware does not permit a receiver reference phase shift, whereas any value of the receiver reference phase is permitted in others. In order to cater for the variety of spectrometer configurations, the first edition of this book distinguished sharply between these two types of signal phase shift. Fortunately, advances in receiver technology, such as oversampling and digital signal post-processing, have largely erased these distinctions in modern instruments.

6. For example, see L. Frydman, T. Scherf, A. Lupulescu, *Proc. Natl. Acad. Sci. USA*, **99,** 15 858–15 862, (2002).

7. Because of the geometrical properties of magnetic fields, it is strictly impossible to construct field gradients that involve a pure variation of only one component of the magnetic field in only one direction, while keeping the direction of the magnetic field constant. Figure 4.16 illustrates this difficulty: the direction of the field is not strictly constant, as its magnitude is varied along the z-axis. Fortunately, the errors caused by these inevitable field distortions are negligible in ordinary high-field NMR and NMR imaging applications.

Further Reading

- For more on the practical aspects of the spectrometer, see E. Fukushima, S. B. W. Roeder, *Experimental Pulse NMR. A Nuts and Bolts Approach*, Perseus Press, Cambridge, MA, USA, 1986.
- For discussion of the appropriate choice of sampling bandwidth and the total acquisition duration, see A. E. Derome, *Modern NMR Techniques in Chemistry Research*, Pergamon Press, Oxford, 1990.

5 Fourier Transform NMR

In this chapter, I describe how the different components of the spectrometer are coordinated so as to produce an NMR signal. The fate of the NMR signal is followed as it is processed and converted into an NMR spectrum by the computer.

5.1 A Single-Pulse Experiment

Consider a simple NMR experiment consisting of a single r.f. pulse on a single r.f. channel, followed by signal detection. The sequence of events is as follows:

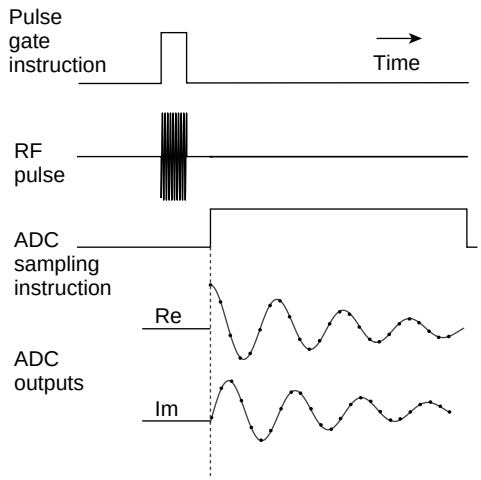


Figure 5.1
Timing sequence for a simple one-pulse experiment.

1. *Initialization.* Before the experiment starts, the computer downloads instructions to the pulse programmer and to other pieces of hardware, such as the synthesizer and the ADCs, setting up the carrier frequency, the sampling frequency and the number of sampled points, etc.

2. *Excitation.* On initiation, the pulse programmer executes a timed sequence of instructions to set the phase of the r.f. synthesizer and open the pulse gate. An r.f. pulse travels into the probe from the amplifier, via the duplexer. This r.f. pulse sets up resonant oscillations in the tuned circuit of the probe, irradiating the sample with an r.f. field close to the Larmor frequency of the chosen nuclear isotope. As decribed

Spin Dynamics: Basics of Nuclear Magnetic Resonance, Second Edition Malcolm H. Levitt

in Section 2.7, this r.f. pulse disturbs the equilibrium of the nuclear spin system and creates transverse nuclear magnetization.

3. *Detection.* The pulse is switched off. After some microseconds, the pulse energy in the tuned circuit dissipates and the system is ready to detect the NMR signal. The precession of the nuclear spin magnetization sets up oscillations in the tuned circuit, which give rise to an r.f. NMR signal (the FID), travelling back down the cable from the probe to the duplexer. This signal is amplified by the signal preamplifier and converted down in frequency by the quadrature receiver. The pulse programmer issues an instruction to the ADCs, which proceed to digitize the two quadrature receiver outputs. The digitized complex signal $s(t)$ is stored in the computer memory.

4. *Processing and display.* The digital complex signal is subjected to various mathematical operations, including the numerical calculation called *Fourier transformation* (FT), discussed in Section 5.8.1. FT converts the NMR signal, which is a function of time, into an NMR *spectrum*, which is a function of frequency. The NMR spectrum may be displayed on the computer screen, printed or plotted on a sheet of paper, or subjected to any of the chicanery of digital communications or processing.

In practice, there are additional timing signals for blanking the r.f. amplifiers, and switching off the receiver during the r.f. pulses, so as to avoid undesirable electronic interference. These complications are ignored here.

In pulse sequence diagrams, the above timing steps are often condensed to the following iconic form:

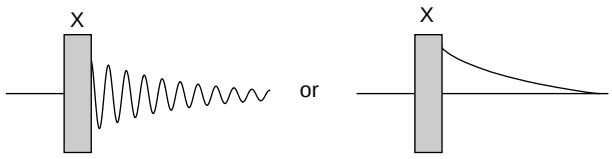


Figure 5.2
Icons representing a one-pulse sequence.

The symbol 'x' over the pulse indicates that the phase of the synthesizer is set to zero during the pulse (see Table 4.1).

Pulse sequence timing diagrams are rarely drawn to scale along the time axis. In reality, each r.f. pulse usually lasts only a few microseconds, whereas the signal acquisition interval may last several seconds.

5.2 Signal Averaging

The NMR signal is very weak. The signal emerging from the probe also contains uncontrolled random signals called *noise*. The battle against noise is a persistent feature of NMR spectroscopy.

The complex signal $s(t)$ at the output of the digitizer is given by

$$s(t) = s_{\mathrm{NMR}}(t) + s_{\mathrm{noise}}(t) \tag{5.1}$$

where s_{NMR} is the NMR signal and s_{noise} is the noise contribution. Typically, r.f. noise looks as follows:

Figure 5.3
Typical appearance of r.f. noise.

Noise is analogous to the 'hiss' in an old-fashioned transistor radio.

In a well-designed spectrometer, the main origins of r.f. noise are as follows: (i) thermal motion of charged particles in the sample, such as dissolved ions; (ii) thermal motions of electrons in the receiver coil. In most cases, the latter is the dominant noise source.[1]

The NMR spectrocopist has three ways of tackling the problem of noise. It is possible to exploit: (1) the narrow frequency spectrum of the NMR signals, compared with the noise; (2) the different time-profiles of the NMR signal and the noise – whereas the NMR signal tends to damp out as a function of time, the noise remains time independent, on the average; and (3) the different statistical properties of the NMR signal and the noise.

Property (1) may be exploited by good design of the r.f. hardware. It is not discussed further here. Property (2) may be exploited by manipulating the data in the computer, after the NMR experiment is finished.[2] See *Further Reading* for some practical recommendations in this respect.

In this section, we will consider how to exploit property (3). The NMR signals are reproducible (to a good approximation), whereas noise signals are random. If the same experiment is repeated, then the NMR signal is identical (assuming that the sample is unchanged, the spins are in similar states at the beginning of each experiment, and all the experimental parameters are well controlled). The noise, on the other hand, varies in an irreproducible way if the experiment is repeated, since it comes from sources other than nuclear spins. This property may be used to enhance the NMR signal at the expense of the noise. All that is needed is to repeat the same experiment many times and to add the signals together. This procedure is called *signal averaging*. (This term is not used very strictly. Often, the signal is 'summed' rather than 'averaged'. Nevertheless, I will continue to use the more familiar term.)

To see how signal averaging works, consider two independent, identical, NMR experiments. The true NMR signal $s_{\mathrm{NMR}}(1)$ from the first experiment is the same as the NMR signal $s_{\mathrm{NMR}}(2)$ from the second experiment. The sum of the two NMR signals, denoted $s_{\mathrm{NMR}}(1+2)$, is simply twice the signal from any individual experiment:

$$s_{\mathrm{NMR}}(1+2) = s_{\mathrm{NMR}}(1) + s_{\mathrm{NMR}}(2) = 2\,s_{\mathrm{NMR}}(1)$$

This simple relationship does not apply to the noise contributions, which are random. A suitable definition of the noise amplitude in a single experiment is provided by the *root-mean-square* (RMS) *noise*, defined as follows:

$$\sigma_{\mathrm{noise}}(1) = \left\langle s_{\mathrm{noise}}(1)^2 \right\rangle^{1/2}$$

where the angled bracket indicates the average over all the sampling points. Note that the simpler expression $\langle s_{\mathrm{noise}} \rangle$ would not do to quantify the noise, since the average of the noise points is approximately zero (the noise is just as likely to be positive as to be negative).

This RMS noise statistic is the same for the two experiments, to a good approximation, assuming that the noise is *stationary* (i.e. the amount of noise does not change from time to time):

$$\sigma_{\mathrm{noise}}(1) \cong \sigma_{\mathrm{noise}}(2)$$

However, this does not imply that the sum of the noise signals from two experiments has twice the RMS value. The RMS value of the summed noise is

$$\begin{aligned}
\sigma_{\mathrm{noise}}(1+2) &= \left\langle (s_{\mathrm{noise}}(1) + s_{\mathrm{noise}}(2))^2 \right\rangle^{1/2} \\
&= \left\langle s_{\mathrm{noise}}(1)^2 + 2s_{\mathrm{noise}}(1)s_{\mathrm{noise}}(2) + s_{\mathrm{noise}}(2)^2 \right\rangle^{1/2} \\
&= \left\{ \left\langle s_{\mathrm{noise}}(1)^2 \right\rangle + \left\langle 2s_{\mathrm{noise}}(1)s_{\mathrm{noise}}(2) \right\rangle + \left\langle s_{\mathrm{noise}}(2)^2 \right\rangle \right\}^{1/2}
\end{aligned}$$

Now, since the noise is *uncorrelated* between the two experiments, the 'cross-term' $\left\langle s_{\mathrm{noise}}(1)s_{\mathrm{noise}}(2) \right\rangle$ vanishes. The summed noise over two experiments, therefore, has the RMS value

$$\sigma_{\mathrm{noise}}(1+2) \cong \sqrt{2}\,\sigma_{\mathrm{noise}}(1)$$

It is larger than the noise over a single experiment, but only by the approximate factor $\sqrt{2}$. The *signal-to-noise ratio* for the sum of the two experimental signals is $\sqrt{2}$ larger than that for one experiment:

$$\frac{s_{\mathrm{NMR}}(1+2)}{\sigma_{\mathrm{noise}}(1+2)} \cong \frac{2s_{\mathrm{NMR}}(1)}{\sqrt{2}\sigma_{\mathrm{noise}}(1)} = \sqrt{2}\frac{s_{\mathrm{NMR}}(1)}{\sigma_{\mathrm{noise}}(1)}$$

The arguments above are easily extended to show that the signal-to-noise ratio accumulated by adding together $\mathfrak{N}$ transients is a factor $\sqrt{\mathfrak{N}}$ larger than the signal-to-noise ratio for a single transient:

$$\frac{s_{\mathrm{NMR}}(1+2+\ldots\mathfrak{N})}{\sigma_{\mathrm{noise}}(1+2+\ldots\mathfrak{N})} \cong \sqrt{\mathfrak{N}}\frac{s_{\mathrm{NMR}}(1)}{\sigma_{\mathrm{noise}}(1)} \tag{5.2}$$

The signal-to-noise ratio is approximately proportional to the square root of the number of summed transients.

It follows that repetition of experiments, and summation of signals, gradually distinguishes the NMR signals from the noise. The NMR signals and the noise both increase on signal averaging, but the NMR signals increase faster:

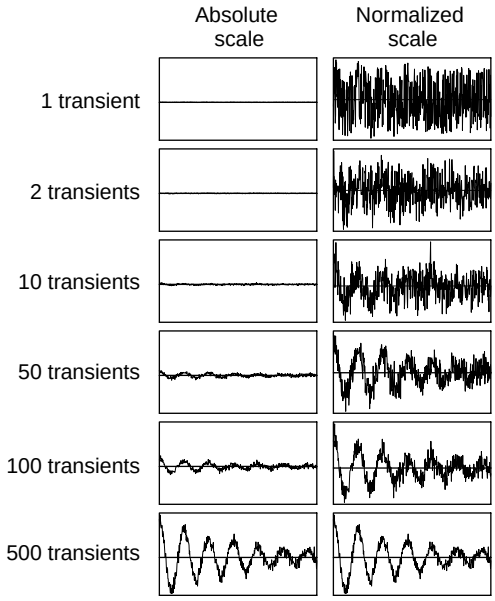


Figure 5.4
Adding together transients causes the NMR signals to build up faster than the noise.

The plots in the left column employ a uniform vertical scale, whereas the plots in the right column are rescaled so that the highest peak fills the box.

In NMR jargon, the separate experiments that compose a signal averaging scheme are called *transients*. The terms 'shots' and 'scans' are also used.

In principle, signal averaging always allows the NMR signals to be 'pulled out' of the noise. However, signal averaging is very time consuming. In order to repeat an experiment precisely, it is necessary to wait for the spin system to attain thermal equilibrium again. The different NMR experiments must be separated by an interval that is longer than the spin–lattice relaxation time constant T_1, several seconds in most cases. The available instrumental time and the limitations on the long-term stability of the instrument or the sample restrict the possibilities for distinguishing NMR signals from noise. For example, suppose that the signal-to-noise ratio in the spectrum obtained from a single transient is 1/10 (meaning that the signal is

completely buried in the noise). The signal-to-noise ratio may be transformed into 10:1 (which is usable), by averaging the results of 10 000 transients. If each transient takes 1 s, then collecting the whole set requires about 3 h of instrument time, which is acceptable. If, on the other hand, the signal-to-noise ratio in the spectrum obtained from a single transient is only 1/100, then 300 h of instrument time would be required to obtain an acceptable result. In many cases, this is not feasible.

5.3 Multiple-Pulse Experiments: Phase Cycling

Most NMR experiments employ more than one r.f. pulse. Consider, for example, the following pulse sequence:

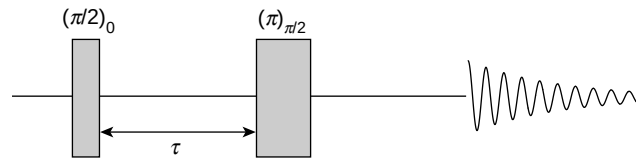


Figure 5.5
A two-pulse sequence.

This consists of two r.f. pulses, with the second pulse twice as long as the first. The time interval between the pulses, marked by a *double-headed arrow*, is denoted τ.

This pulse sequence is used to observe a phenomenon called a *spin echo*, discussed in Section 12.2. At the moment, I am not going to explain this pulse sequence, but merely discuss the notation.

The two pulses are annotated in the form β_ϕ. For the first pulse, β has the value $\pi/2$; for the second pulse, $\beta = \pi$. The value of β is called the *flip angle* of the r.f. pulse. The flip angle is a specialized notation for the *time duration* of the pulse, as described in Section 10.8.2. The fact that β for the second pulse is twice as large as for the first pulse conveys the same information as the timing diagram, where the second pulse is drawn twice as thick along the time axis as the first. As usual, the time axis is not realistic: the interval between the pulses is usually several orders of magnitude larger than the pulse durations.

The value of ϕ in the notation β_ϕ describes the *phase* of the pulse, as described in Section 4.2.1.

In the scheme above, the first pulse has phase $\phi = 0$, and the second pulse has phase $\phi = \pi/2$. This explicit notation implies that the pulse phases are fixed. If the experiment is repeated for the purposes of signal averaging, then the same phases are used on each transient.

The following pulse sequence diagram means the same thing:

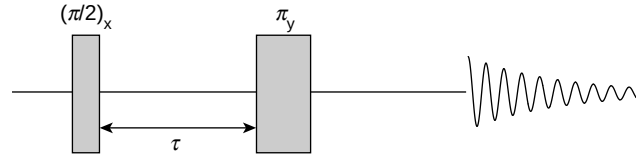


Figure 5.6
The same two-pulse sequence.

The subscripts 'x' and 'y' indicate r.f. phases according to Table 4.1.

Consider now the following pulse sequence diagram:

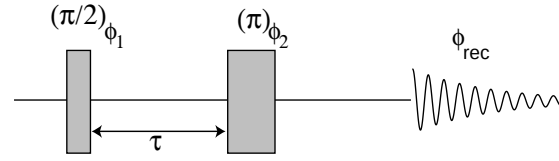


Figure 5.7
Two-pulse sequence with symbolic phases.

Table 5.1 A four-step phase cycle ($\mathfrak{n} = 4$) appropriate for a spin echo experiment.

Cycle counter $\mathfrak{m}$	ϕ_1	ϕ_2	ϕ_{rec}
0	0	0	0
1	0	$\pi/2$	π
2	0	π	0
3	0	$3\pi/2$	π

Here, the two pulse phases have *symbolic* values ϕ_1 and ϕ_2. In addition, the signal acquisition is labelled with the receiver phase ϕ_{rec}. The use of symbolic phases, rather than explicit phases, usually indicates the execution of a *phase cycle*. This means that, when the experiment is repeated, the phases of the pulses are not constant, but run through a set of values, often summarized in a table. For example, a suitable phase table for the spin echo experiment is given in Table 5.1.

This phase table should be understood as follows. Suppose each transient is specified by the value of a *transient counter* $\mathfrak{M}$, which starts at the value 0 and increases to the value $\mathfrak{N} - 1$, where $\mathfrak{N}$ is the total number of co-added transients. The phases to be used depend on the value of the *cycle counter* $\mathfrak{m}$, which is calculated as

$$\mathfrak{m} = \text{mod}(\mathfrak{M}, \mathfrak{n})$$

Here, $\mathfrak{n}$ is the number of steps in the phase cycle, equal to four in the example shown. The modulo function calculates the remainder when $\mathfrak{M}$ is divided by $\mathfrak{n}$. In the example, as $\mathfrak{M}$ increases 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 . . ., the counter $\mathfrak{m}$ takes the values 0, 1, 2, 3, 0, 1, 2, 3, 0, 1 . . .. The phases $\{\phi_1, \phi_2, \phi_{\text{rec}}\}$ for a given transient $\mathfrak{M}$ may be read from the table by calculating $\mathfrak{m}$ and taking the appropriate row. The phase table should, therefore, be read from top to bottom, and then from top to bottom again, etc. until the number of transients $\mathfrak{N}$ is completed. The number of averaged transients $\mathfrak{N}$ should be an integer multiple of the number of steps in the phase cycle $\mathfrak{n}$.

In the example shown, the phase ϕ_1 is always zero. In more complicated cases, any combination of phases is cycled.

Phase cycles are used for two main purposes. First, they are used to *select* NMR signals that have certain properties of interest, while removing other types of NMR signal. For example, Section 16.2 describes a technique in which phase cycling is used to select signals from nuclear spins that have J-couplings to other nuclear spins, while suppressing signals from spins with no coupling partners. Second, phase cycles are used to suppress spurious signals generated by imperfections in the spectrometer hardware. I concentrate on the first use of phase cycling in this book.

5.4 Heteronuclear Experiments

Most NMR experiments are conducted with r.f. irradiation of more than a single isotope, using a multiple-channel spectrometer and a multiply-tuned probe. A simple example, already encountered in Section 3.9, is shown below:

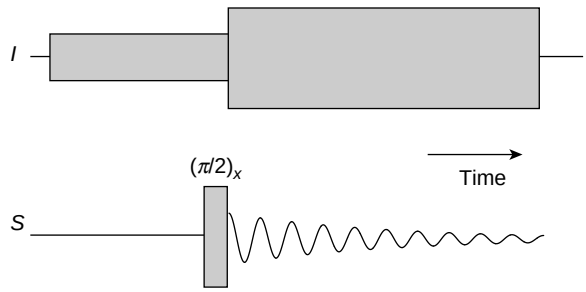


Figure 5.8
A heteronuclear pulse sequence.

The synchronized r.f. sequences on the two channels are shown one above the other and are to be read in parallel, like a musical score.

In the example shown, one-pulse excitation and detection is performed at the Larmor frequency of a nuclear isotope *S*. Before the application of the *S*-spin pulse, continuous r.f. irradiation is applied at the Larmor frequency of a different nuclear isotope *I*. After the *S*-spin pulse is applied, the *I*-spin irradiation continues, with increased amplitude, while the *S*-spin signal is detected.

As mentioned already in Section 3.9, the purpose of the *I*-spin irradiation *before* the *S*-spin pulse is to enhance the *S*-spin magnetization through the NOE. The purpose of the *I*-spin irradiation *during* the *S*-spin observation is to *decouple* the two spin species, simplifying the spectrum of the observed species.

This is only one example of a large class of techniques that involves synchronized r.f. irradiation at the Larmor frequencies of more than one nuclear isotope. Another important class of methods involves *heteronuclear polarization transfer*. In this case, magnetization is transferred from one set of nuclear spins to another, in order to enhance the NMR signals or to trace out the network of magnetic spin–spin interactions. Some examples of this method are treated in Sections 16.3 and 18.12.

5.5 Pulsed Field Gradient Sequences

Many NMR experiments involve magnetic field gradient pulses as well as r.f. pulses, arranged in a strict time sequence. An example is shown below:

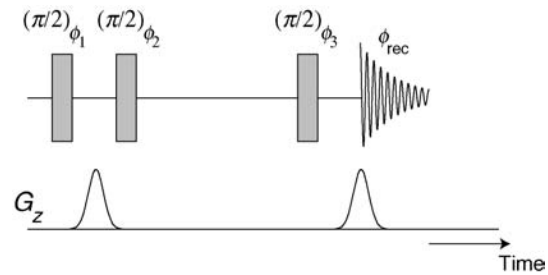


Figure 5.9
A pulse sequence involving two field gradient pulses and three r.f. pulses.

This is a *stimulated echo* pulse sequence, used for the study of molecular diffusion and flow (see Section 19.8). In the case shown, the field gradient pulses have identical shape and length, and both implement a field gradient along the *z*-axis (see Section 4.7). In general, the pulses may have any shape in time (including a change in sign), depending on the capabilities of the probe and spectrometer.

NMR imaging experiments often involve pulses on all three gradient channels (G_x, G_y and G_z). Some simple examples will be encountered in Section 12.6.

5.6 Arrayed Experiments

So far, we have assumed that the acquired NMR signals are simply added together, in order to enhance the signal-to-noise ratio. We now encounter some more sophisticated data acquisition motifs.

The pulse sequence shown below is used in *two-dimensional correlation spectroscopy*, described in Section 16.1:

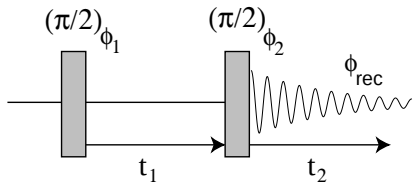


Figure 5.10
A two-dimensional pulse sequence.

The sequence consists of two r.f. pulses with equal durations (both have flip angles $\beta = \pi/2$), and with phases ϕ_1 and ϕ_2. Signal acquisition is conducted using a receiver phase ϕ_{rec}. As above, the use of these symbolic phase values indicates the use of a phase cycle. A number of different transients are added together, varying the pulse phases between consecutive pulse sequences, according to a phase table.

In Figure 5.10, the interval between the pulses is labelled with a *single-headed arrow*, marked t_1. By convention, the single-headed arrow indicates that a series of separate experiments is conducted, with the interval t_1 taking a set of different values. The digitized FIDs generated by pulse sequences with different values of t_1 are not added together, but stored in separate locations in the computer.

The data acquisition is marked with another single-headed arrow, labelled t_2. The symbol t_1 indicates the variable interval between the two pulses, and the symbol t_2 indicates the time coordinate of the digitized signal.

The variation of both time variables, with separate data storage for each value of t_1, leads to the compilation of a *data matrix*. This may be regarded as the compilation of a two-dimensional signal surface $s(t_1, t_2)$, one row at a time:

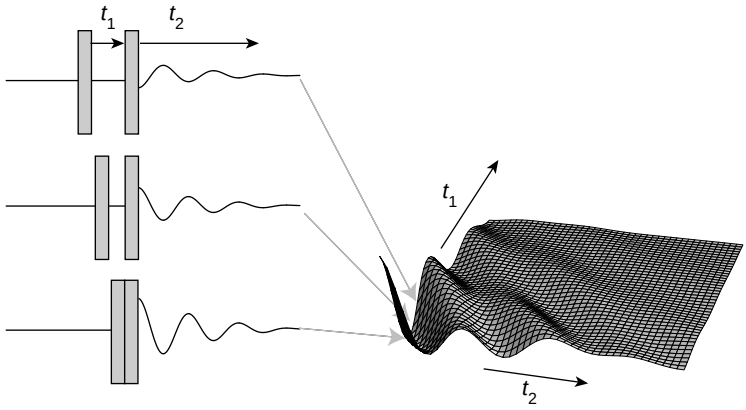


Figure 5.11
Compilation of a two-dimensional signal surface by arrayed data acquisition.

Each row of the data matrix is the result of a complete set of phase cycled experiments, all with the same value of t_1, but with cycling of the phases ϕ_1, ϕ_2 and ϕ_{rec} according to the phase table. When acquisition of one row is completed, the variable delay t_1 is changed, and the acquisition procedure is repeated. The iconic pulse sequence diagrams imply a timing hierarchy, with time interval incrementation enclosing an inner level of phase cycling.

Arrayed signal acquisition is the basis of most forms of *two-dimensional spectroscopy*.[3]

Arraying may be extended to three dimensions by introduction of an additional variable parameter. For example, a three-dimensional experiment may involve two variable delays, t_1 and t_2, in addition to the ordinary time coordinate for the data acquisition (usually renamed t_3 in this type of experiment). For each value of t_1 and t_2, a complete phase cycle is performed, compiling a single line through a 'data cube' $s(t_1, t_2, t_3)$. This entire procedure is repeated, keeping t_1 constant, but incrementing t_2, building up a sampled 'plane' of data. Repetition for all values of t_1 constructs the entire three-dimensional data set, plane by plane. The experimental hierarchy reads as follows: an upper level of t_2 incrementation; inside this a level of t_1

incrementation; and the innermost loop, a phase cycle for each set of averaged transients. The hierarchical experimental structure often means that a single data set is compiled by performing thousands of separate NMR pulse sequences, none of them identical.

The arraying concept may be extended to any number of dimensions. High-dimensional experiments of this type are frequently performed in biomolecular solution NMR.

The degree of complexity attained by some modern NMR pulse sequences is illustrated in the example shown in Figure 5.12, which is a fairly typical excerpt from the *Journal of Magnetic Resonance*. The increase in the number of spectrometer channels has led to an elongation of pulse sequence diagrams in the vertical direction, with a growing resemblance to orchestral musical scores.

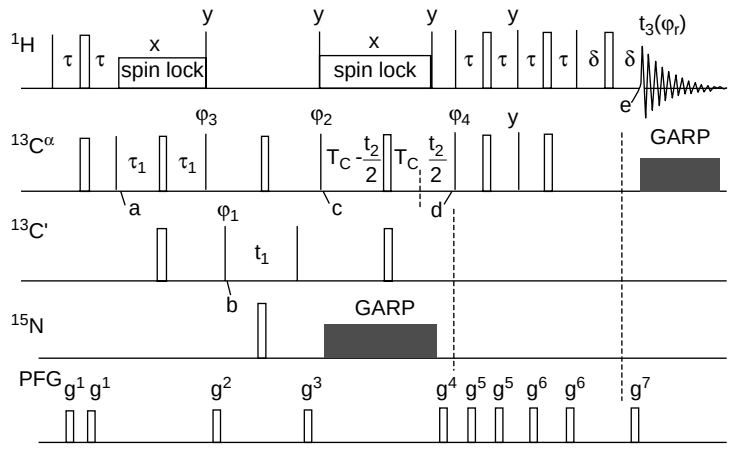


Figure 5.12
A typical example of a complicated r.f. pulse sequence. Note the pulsed field gradient pulses on the lowest line. From Y. Xia *et al.*, *J. Magn. Reson.* **143**, 407 (2000). (Reproduced by permission of Academic Press.)

5.7 NMR Signal

In Section 4.5.2, I introduced the following expression for the quadrature-detected NMR signal:

$$s(t) \sim \exp\{(\mathrm{i}\Omega^0 - \lambda)t\}$$

This expression corresponds to a spectrum with just one peak. A more generally applicable equation is

$$s(t) = \sum_{\ell} s_\ell(t) \tag{5.3}$$

where

$$s_\ell(t) = a_\ell \exp\{(\mathrm{i}\Omega_\ell - \lambda_\ell)\, t\} \tag{5.4}$$

This describes a superposition of many different signal components s_ℓ. In general, each component s_ℓ has a different frequency Ω_ℓ, a different damping rate constant λ_ℓ, and a different amplitude a_ℓ.

Equations 5.3 and 5.4 apply to the NMR signal generated by *any* pulse sequence. The peak frequencies Ω_ℓ and damping constants λ_ℓ depend only on what happens to the spins *when* the signal is detected. The peak amplitudes a_ℓ depend only on what happens to the spins *before* the signal is detected[4] (see Figure 5.13):

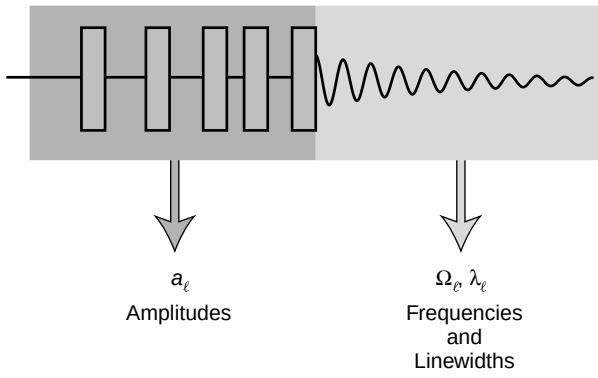


Figure 5.13
The relationship between the pulse sequence, the amplitudes, the frequencies, and the linewidths.

Most of the rest of this book is concerned with how to calculate a_ℓ, Ω_ℓ and λ_ℓ in specific cases.

The amplitudes a_ℓ are complex numbers in general. Each *complex amplitude* may be written as follows:

$$a_\ell = |a_\ell| \exp\{i\phi_\ell\} \tag{5.5}$$

The magnitude $|a_\ell|$ is called the *intensity* of the signal component. The factor ϕ_ℓ is the *phase* of the signal component.

The real and imaginary parts of the signal s are shown below for four different cases, all of which have only one signal component. The corresponding values of the complex amplitude, the frequency, and the damping rate constant are shown next to each plot:

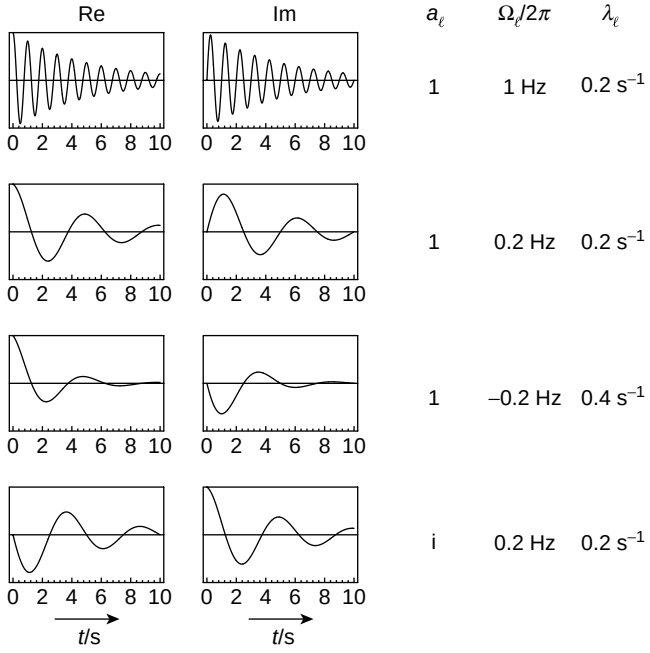


Figure 5.14
Four time-domain NMR signals, each with a single component.

The second case has a much lower frequency than the first case. The second and third cases differ in the sign of the frequency Ω_ℓ, which is reflected in the sign of the imaginary component. In addition, the third case has a larger value of λ_ℓ, which gives rise to a steeper decay as a function of time. The fourth case has an

imaginary complex amplitude a_ℓ. This interchanges the types of modulation seen in the real and imaginary parts of the signal.

A signal with two components:

$$s(t) = 0.5\exp\{(2\pi\mathrm{i} - 0.2)\,t\} + 0.5\exp\{(0.4\pi\mathrm{i} - 0.2)\,t\} \tag{5.6}$$

has the following appearance:

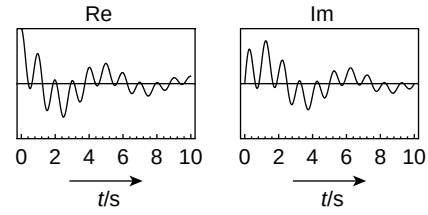


Figure 5.15
A complex NMR signal with two components.

which is much more difficult to interpret by eye.

In the last example, there are four signal components:

$$\begin{aligned} s(t) = {} & 0.25\exp\{(2\pi\mathrm{i} - 0.2)\,t\} + 0.25\exp\{(0.4\pi\mathrm{i} - 0.2)\,t\} \\ & +0.25\exp\{(-0.4\pi\mathrm{i} - 0.4)\,t\} + 0.25\mathrm{i}\exp\{(-3.6\pi\mathrm{i} - 0.2)\,t\} \end{aligned}$$

leading to a very complicated appearance:

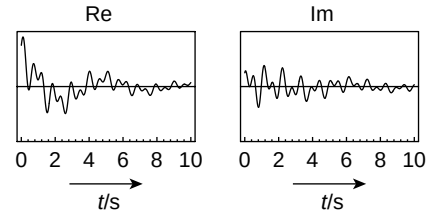


Figure 5.16
A complex NMR signal with four components.

The experimental FID from the proton spins in a solution of a protein is shown in Figure 5.17a. In this case, there are hundreds of signal components.

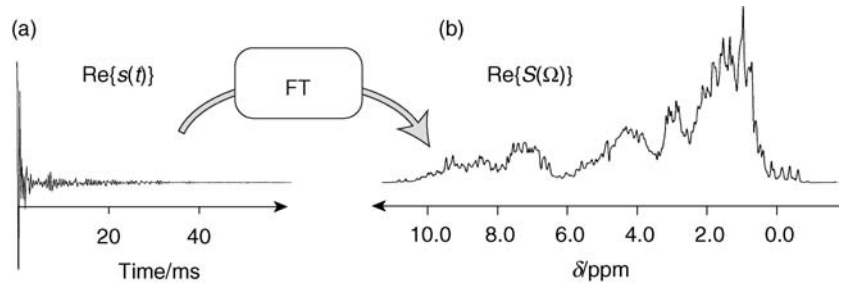


Figure 5.17 Proton NMR signal $s(t)$ for a solution of a protein (a) and its Fourier transform (b). Only the real parts of the signal and the spectrum are shown – in fact, the imaginary part of the signal $s(t)$ is also required for the Fourier transform. Thanks to Jianyun Lu for supplying these data.

5.8 NMR Spectrum

We now investigate the fate of the NMR signal after it has been stored in the computer.

5.8.1 Fourier transformation

Fourier transformation (FT) is a mathematical technique that converts a function of time into a function of frequency. The effect of FT is to make visible the individual components of the signal and plot their frequencies in a visually accessible form. FT does not enhance the theoretical information content, but makes the information in the signal more accessible to the human eye.

The mathematical definition of FT[5] is

$$S(\Omega) = \int_0^\infty s(t) \exp\{-i\Omega t\} \, dt \tag{5.7}$$

The 'input' s is a function of time t; the 'output' S is a function of a frequency variable Ω, and is called the *spectrum*.

Both the time-domain signal $s(t)$ and the spectrum $S(\Omega)$ are complex functions. The functional relationship between the real and imaginary parts of $s(t)$ and $S(\Omega)$ may be depicted as follows:

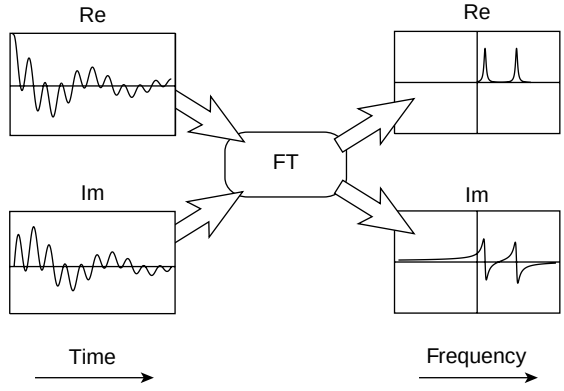


Figure 5.18
The relationship between the real and imaginary parts of $s(t)$ and $S(\Omega)$.

Note that the real and imaginary parts of the spectrum are each derived from *both* the real part and the imaginary part of the signal.

Figure 5.18 shows what happens when the signal has two components, as in Equation 5.6. Fourier transformation renders the two components clearly visible as two spectral peaks.

5.8.2 Lorentzians

FT may be understood either on a purely mathematical level or by using physical arguments. I first give a mathematical interpretation, and return later to the 'physical explanation'.

If Equation 5.7 is applied to Equation 5.3, we obtain

$$S(\Omega) = \sum_\ell S_\ell(\Omega) \tag{5.8}$$

where each spectral component $S_\ell(\Omega)$ is the Fourier transform of the corresponding signal component:

$$S_\ell(\Omega) = \int_0^\infty s_\ell(t)\, \exp\{-\mathrm{i}\Omega t\}\, \mathrm{d}t$$

From Equation 5.4, this may be written as

$$S_\ell(\Omega) = a_\ell \int_0^\infty \exp\{-\left(\mathrm{i}(\Omega - \Omega_\ell) + \lambda_\ell\right) t\}\, \mathrm{d}t$$

Direct integration gives

$$S_\ell(\Omega) = -\frac{a_\ell}{\lambda_\ell + \mathrm{i}(\Omega - \Omega_\ell)} \Big[\exp\{-\left(\mathrm{i}(\Omega - \Omega_\ell) + \lambda_\ell\right) t\} \Big]_{t=0}^{t\to\infty} \tag{5.9}$$

Evaluation of the lower limit ($t = 0$) is easy, since $e^0 = 1$, but it is less obvious how to handle the upper limit ($t \to \infty$). The way forward is to use the following identity:

$$\exp\{(\mathrm{i}\omega - \lambda)t\} = (\cos \omega t + \mathrm{i} \sin \omega t)\ \exp\{-\lambda t\} \tag{5.10}$$

Since the decay rate constant λ is positive by definition, the factor $\exp\{-\lambda t\}$ goes to zero at $t \to \infty$. Furthermore, the cosine and sine functions are bounded by -1 and $+1$. Hence, the complete expression in Equation 5.10 must vanish at $t \to \infty$. The upper limit of Equation 5.9, therefore, evaluates to zero, and we get

$$S_\ell(\Omega) = -\frac{a_\ell}{\lambda_\ell + \mathrm{i}(\Omega - \Omega_\ell)} \times (0 - 1) = a_\ell \times \left(\frac{1}{\lambda_\ell + \mathrm{i}(\Omega - \Omega_\ell)} \right) \tag{5.11}$$

The bracketed function is very important in the theory of NMR, and is called the *complex Lorentzian*. It is defined as

$$\mathcal{L}(\Omega; \Omega_\ell, \lambda) = \frac{1}{\lambda + \mathrm{i}(\Omega - \Omega_\ell)} \tag{5.12}$$

In the notation $\mathcal{L}(\Omega; \Omega_\ell, \lambda)$, the 'argument' of the function is the frequency coordinate Ω, and Ω_ℓ and λ are 'parameters'. The parameter Ω_ℓ indicates the centre frequency of the peak and λ is a peakwidth parameter.

The spectral component in Equation 5.11 may be written in terms of the complex Lorentzian as follows:

$$S_\ell(\Omega) = a_\ell\, \mathcal{L}(\Omega; \Omega_\ell, \lambda_\ell) \tag{5.13}$$

Just as the time-domain signal is a superposition of oscillating components s_ℓ, the spectrum is a superposition of Lorentzian spectral components S_ℓ.

The *real part* of the complex Lorentzian is called the *absorption Lorentzian*:

$$\mathcal{A}(\Omega; \Omega_\ell, \lambda) = \mathrm{Re}\{\mathcal{L}(\Omega; \Omega_\ell, \lambda)\} = \frac{\lambda}{\lambda^2 + (\Omega - \Omega_\ell)^2} \tag{5.14}$$

The *imaginary part* of the complex Lorentzian is called the *dispersion Lorentzian*:[6]

$$\mathcal{D}(\Omega; \Omega_\ell, \lambda) = \mathrm{Im}\{\mathcal{L}(\Omega; \Omega_\ell, \lambda)\} = -\frac{\Omega - \Omega_\ell}{\lambda^2 + (\Omega - \Omega_\ell)^2} \tag{5.15}$$

The three peakshapes are related through the following equation:

$$\mathcal{L} = \mathcal{A} + i\mathcal{D} \tag{5.16}$$

Whereas $\mathcal{L}$ is a complex function, $\mathcal{A}$ and $\mathcal{D}$ are both real functions.

The absorption and dispersion Lorentzian peakshapes are sketched below:[7]

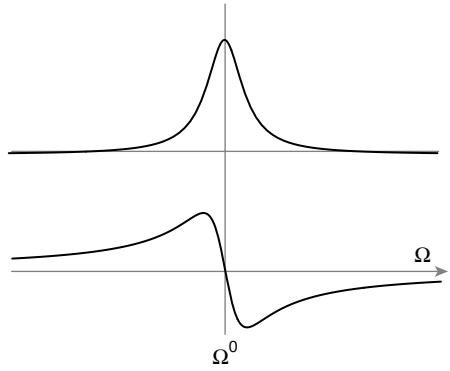


Figure 5.19
Absorption and dispersion Lorentzians.

The absorption Lorentzian has an integral of π and a maximum height equal to λ^{-1}. The *full-width-at-half-height* (FWHH) of the absorption Lorentzian is given by 2λ in units of radians per second. This is easily seen from

$$\mathcal{A}(\Omega; \Omega_\ell, \lambda) = \frac{1}{2}\lambda^{-1}$$

which has solutions at

$$\Omega = \Omega_\ell \pm \lambda$$

The FWHH of the absorption Lorentzian in units of hertz is $2\lambda/2\pi = \lambda/\pi$.

The dispersion Lorentzian has zero integral. The function passes through zero at the frequency coordinate Ω_ℓ. The frequency separation between the maximum and minimum points is 2λ in units of radians per second, which is the same as the FWHH of the absorption Lorentzian.

The absorption Lorentzian is proportional to $(\Omega - \Omega_\ell)^{-2}$ at large offsets $\Omega - \Omega_\ell$ from the centre frequency. This inverse square dependence causes the absorption Lorentzian to die off rather quickly away from the peak centre. The dispersion Lorentzian, on the other hand, has a $(\Omega - \Omega_\ell)^{-1}$ dependence at large offsets and dies away much more slowly. As a result, dispersion Lorentzians have broader wings than absorption Lorentzians. This fact is important in crowded spectra, since the wings from one peak may distort a neighbouring, partially overlapping, peak.

Equation 5.8 indicates that the spectrum $S(\Omega)$ is a superposition of complex Lorentzian peakshapes, each deriving from a single oscillating time-domain component $s_\ell(t)$. The effect of the Fourier transform is therefore to make visible the frequency components contained in the complicated time-domain signal $s(t)$. Each frequency component is marked by a Lorentzian peak.

The plots shown in Figure 5.20 present the real and imaginary parts of the spectrum for four different cases, each of which has only one spectral component.

Note how the parameters a_ℓ, Ω_ℓ and λ_ℓ affect the appearance of the real and imaginary parts of the spectrum. In the first three cases, the amplitudes a_ℓ are real, so the real part of the spectrum contains an absorption signal and the imaginary part of the spectrum contains a dispersion signal. In the last case, the amplitude a_ℓ is imaginary, so the positions of the absorption and dispersion signals are interchanged. The third case has a larger value of λ_ℓ, which gives rise to a broader peak.

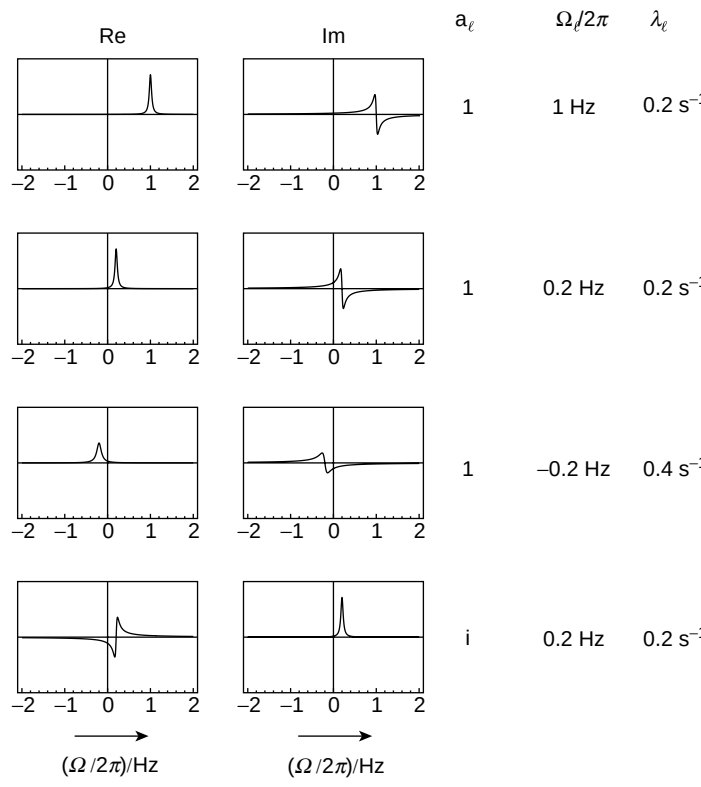


Figure 5.20
Four NMR spectra, each with a single component. These are the Fourier transforms of the signals in Figure 5.14.

This is an example of an important general principle:

$$\begin{aligned} \text{Fast signal decay} &\Longleftrightarrow \text{Broad peak} \\ \text{Slow signal decay} &\Longleftrightarrow \text{Narrow peak} \end{aligned} \tag{5.17}$$

A spectrum with two components:

$$S(\omega) = 0.5\mathcal{L}(\Omega; 2\pi, 0.2) + 0.5\mathcal{L}(\Omega; 0.4\pi, 0.2) \tag{5.18}$$

has the following appearance:

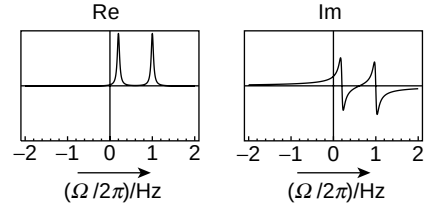


Figure 5.21
Four NMR spectra, each with two components. These are the Fourier transforms of the signals in Figure 5.15.

which is easy to interpret. Note how the position of each peak in the spectrum corresponds to the value of the parameter Ω_ℓ in the Lorentzian function $\mathcal{L}(\Omega; \Omega_\ell, \lambda)$.

Interpretation is still easy for a case with four signal components:

$$S(\Omega) = 0.25\mathcal{L}(\Omega; 2\pi, 0.2) + 0.25\mathcal{L}(\Omega; 0.4\pi, 0.2)$$
$$+0.25\mathcal{L}(\Omega; -0.4\pi, 0.4) + 0.25\mathrm{i}\mathcal{L}(\Omega; -3.6\pi, 0.2)$$

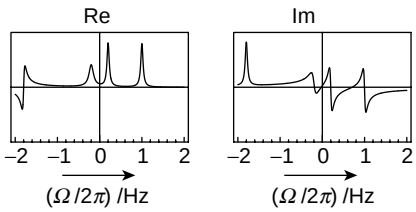


Figure 5.22
Four NMR spectra, each with four components. These are the Fourier transforms of the signals in Figure 5.16.

Note the appearance of the peak on the far left of the spectrum, which has an imaginary complex amplitude, and that one of the peaks is broader than the other three.

The Fourier transform of the experimental NMR signal shown in Figure 5.17a is shown in Figure 5.17b. Compared with the FID, the NMR spectrum is readily interpretable in terms of signals from protons in different types of molecular environment.

5.8.3 Explanation of fourier transformation

How does FT actually 'work'?

The following explanation considers only the real parts of signals, for the sake of simplicity:

1. Suppose that I have a signal $s(t)$ that looks like this:

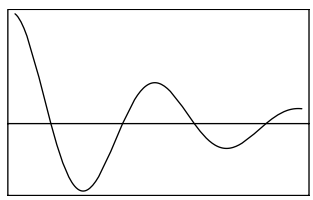


Figure 5.23
An oscillating signal.

I want to know its frequency.

2. First guess a frequency Ω_{guess}. Create a signal function s_{guess} that oscillates at this frequency:

$$s_{\text{guess}} = \exp\{\mathrm{i}\Omega_{\text{guess}}t\}$$

The real part of this function might look like this:

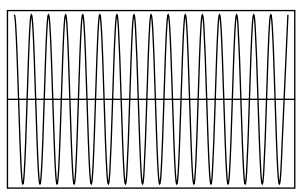


Figure 5.24
An oscillating function based on the first guess at the frequency.

In this case the guess was wrong. The guessed frequency is too high.

3. Multiply $s(t)$ by $s_{\text{guess}}(t)^*$. The real part of the result looks like this:

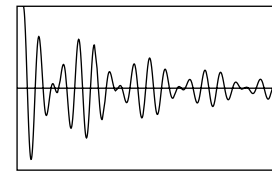


Figure 5.25
The product of the signal and the first oscillating function.

4. Integrate over time. In this case, the guess was not very good. The functions $s(t)$ and $s_{\text{guess}}(t)$ do not match, so the product $s(t)s_{\text{guess}}(t)^*$ has as many positive as negative excursions, and the integral is close to zero. The low value of this integral corresponds to the low value of the spectral function $S(\Omega)$ when Ω is well away from the centre of the peak.

5. Try again with a frequency Ω'_{guess}. The appropriate signal function is

$$s'_{\text{guess}} = \exp\{i\Omega'_{\text{guess}}t\}$$

Suppose it looks like this:

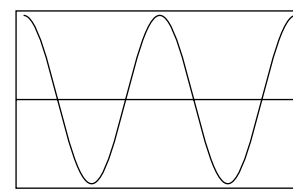


Figure 5.26
An oscillating function based on the second guess at the frequency.

This time the guess is good. The guessed function s'_{guess} matches the oscillations of the signal $s(t)$ rather well.

6. This is seen by multiplying $s(t)$ by $s_{\text{guess}}(t)^*$:

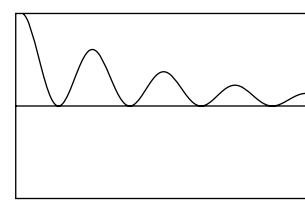


Figure 5.27
The product of the signal and the second oscillating function.

This time the positions of the positive and negative excursions match in $s(t)$ and $s'_{\text{guess}}(t)$. When $s(t)$ is positive, so is $s'_{\text{guess}}(t)$, and the product of the two is positive. When $s(t)$ is negative, so is $s'_{\text{guess}}(t)$, and the product of the two is again positive. As a result, the integral of $s(t)s'_{\text{guess}}(t)^*$ is large. The high value of this integral corresponds to the large value of the spectral function $S(\Omega)$ when Ω is close to the centre of the peak.

7. One can visualize the Fourier transform as an automation of this procedure, stepping through all possible values of Ω_{guess}. When Ω_{guess} hits the true frequency of a signal component, the integral goes up. When Ω_{guess} is far away from any of the frequency components in the signal, the integral goes down:

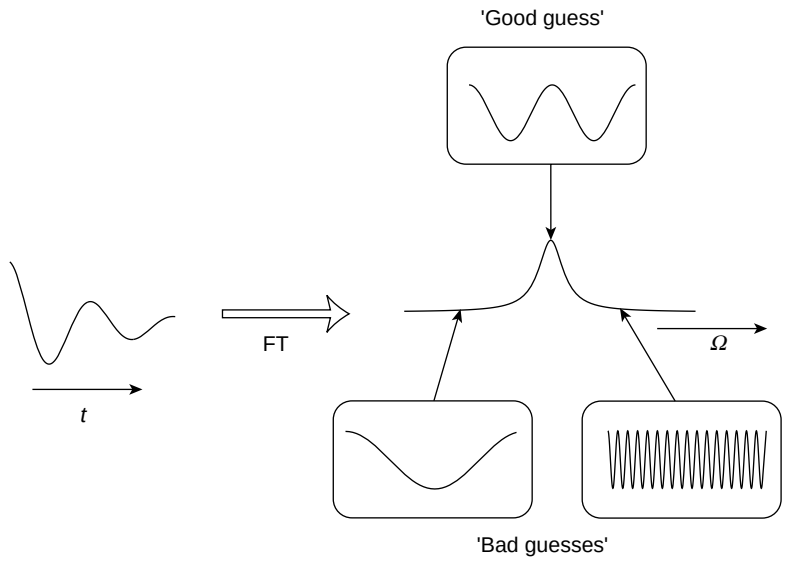


Figure 5.28
How FT works.

In this way, the Fourier transform 'detects' the presence of oscillating components in the time-domain signal and presents their frequencies in a visually accessible form.

Fourier transforms are not actually computed this way. In reality, a special numerical algorithm called the *fast Fourier transform* (FFT) is used.[8] This algorithm is very fast, but has the restriction that the number of Fourier-transformed points must be an integer power of 2. This is the reason for the peculiar numbers given in Table 4.2.

FT is not the only method for extracting the frequency components of a data set. There are other schemes in common use, such as the *maximum entropy method*, *linear prediction*, etc. None of these methods is as robust as the Fourier transform, although they may have advantages in certain cases, such as when the signal-to-noise ratio is good but when there are only a few sampling points available. See *Further Reading*.

5.8.4 Spectral phase shifts

If the amplitudes a_ℓ of the individual signal components are all real numbers, then the real part of the spectrum $\mathrm{Re}\{S(\Omega)\}$ consists of absorption Lorentzians and the imaginary part of the spectrum $\mathrm{Im}\{S(\Omega)\}$ consists of dispersion Lorentzians. Since absorption Lorentzians have better resolution than dispersion Lorentzians, only the real part of the spectrum is usually displayed.

In general, the amplitudes a_ℓ of the spectral components S_ℓ are complex. The real and imaginary components of the spectrum are then mixtures of absorption and dispersion Lorentzians. This is seen by applying the following identities:

$$\begin{aligned}\mathrm{Re}\{ab\} &= \mathrm{Re}\{a\}\,\mathrm{Re}\{b\} - \mathrm{Im}\{a\}\,\mathrm{Im}\{b\}\\ \mathrm{Im}\{ab\} &= \mathrm{Re}\{a\}\,\mathrm{Im}\{b\} + \mathrm{Im}\{a\}\,\mathrm{Re}\{b\}\end{aligned}$$

The result for one spectral component is

$$\begin{aligned}\mathrm{Re}\{a_\ell \mathcal{L}(\Omega;\Omega_\ell,\lambda_\ell)\} &= \mathrm{Re}\{a_\ell\}\mathcal{A}(\Omega;\Omega_\ell,\lambda_\ell) - \mathrm{Im}\{a_\ell\}\mathcal{D}(\Omega;\Omega_\ell,\lambda_\ell)\\ \mathrm{Im}\{a_\ell \mathcal{L}(\Omega;\Omega_\ell,\lambda_\ell)\} &= \mathrm{Re}\{a_\ell\}\mathcal{D}(\Omega;\Omega_\ell,\lambda_\ell) + \mathrm{Im}\{a_\ell\}\mathcal{A}(\Omega;\Omega_\ell,\lambda_\ell)\end{aligned}$$

or in terms of the phase ϕ_ℓ of the complex amplitude (Equation 5.5):[9]

$$\mathrm{Re}\{a_\ell \mathcal{L}(\Omega;\Omega_\ell,\lambda_\ell)\} = |a_\ell|\mathcal{A}(\Omega;\Omega_\ell,\lambda_\ell)\cos\phi_\ell - |a_\ell|\mathcal{D}(\Omega;\Omega_\ell,\lambda_\ell)\sin\phi_\ell$$
$$\mathrm{Im}\{a_\ell \mathcal{L}(\Omega;\Omega_\ell,\lambda_\ell)\} = |a_\ell|\mathcal{D}(\Omega;\Omega_\ell,\lambda_\ell)\cos\phi_\ell + |a_\ell|\mathcal{A}(\Omega;\Omega_\ell,\lambda_\ell)\sin\phi_\ell$$

The mixture of absorption and dispersion modes gives a 'skewed' appearance to the peakshapes: the spectrum is said to be 'out of phase'. The spectral appearance for different values of ϕ_ℓ is shown in Figure 5.29.

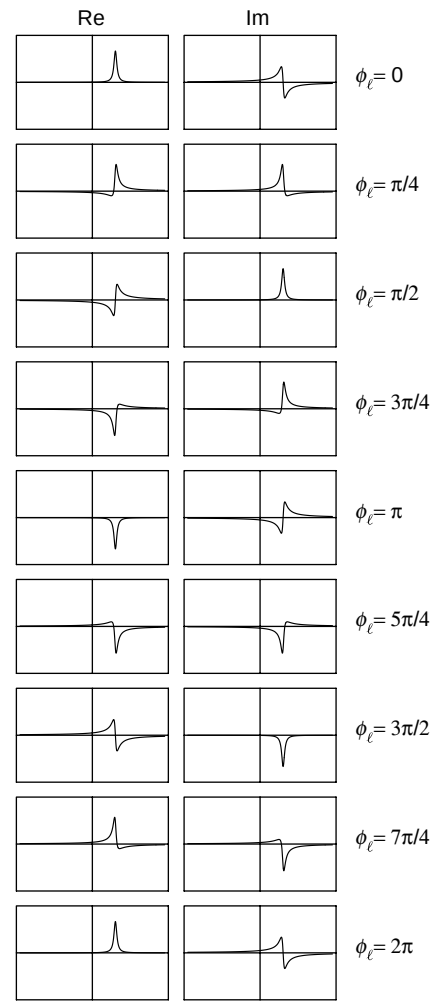


Figure 5.29
The phase of the complex amplitude and the real and imaginary parts of the spectrum.

5.8.5 Frequency-dependent phase correction

In general, the amplitudes a_ℓ of the different spectral components depend on the details of the pulse sequence before the acquisition interval and may have any possible phase. However, in some common

situations, the phases of the peaks depend in a simple, linear, fashion on the centre frequency of the spectral component:

$$\phi_\ell \cong \phi^{(0)} + \phi^{(1)}\Omega_\ell \tag{5.19}$$

where the terms $\phi^{(0)}$ and $\phi^{(1)}$ are the same for all peaks in the spectrum.

A simple, one-pulse, experiment often leads to peak phases of this form. The spectrometer electronics give rise to unavoidable $\phi^{(0)}$ and $\phi^{(1)}$ terms. In addition, $\phi^{(1)}$ is associated with experimental time delays, such as the delay that must be left after the r.f. pulse before the NMR signal can be sampled, in order to allow the pulse energy to dissipate. These phase shifts are often sample dependent and cannot be eliminated in a universal way by spectrometer adjustments.

A four-peak spectrum showing a linear frequency-dependent phase shift is shown in Figure 5.30a.

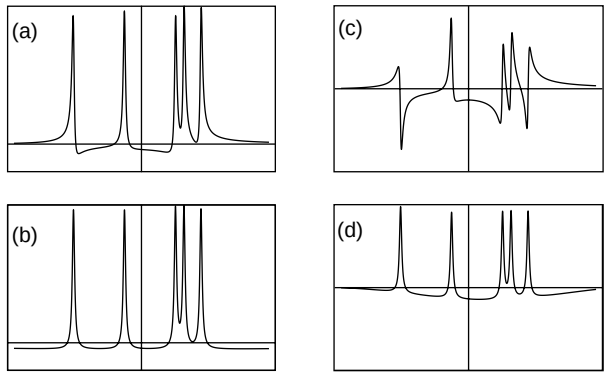


Figure 5.30 (a) A spectrum showing a mild frequency-dependent phase shift. (b) Phase-correction of the spectrum shown in (a). (c) A spectrum showing a strong frequency-dependent phase shift. (d) Phase-correction of the spectrum shown in (c), displaying a rolling baseline.

If $\phi^{(1)}$ is not too large, then the phase of the spectrum may be corrected after the FT, to a good approximation. Instead of displaying the real part of the spectrum $S(\Omega)$, one displays the real part of the *phase-corrected spectrum* $S_{\text{corr}}(\Omega)$:

$$S_{\text{corr}}(\Omega) = S(\Omega)\exp\{-\mathrm{i}\left(\phi^{(0)}_{\text{corr}} + \phi^{(1)}_{\text{corr}}\Omega\right)\} \tag{5.20}$$

In practice, the parameters $\phi^{(0)}_{\text{corr}}$ and $\phi^{(1)}_{\text{corr}}$ are varied manually in an interactive fashion, until the spectrum 'comes into phase'. The spectrum achieves its best appearance when $\phi^{(0)}_{\text{corr}} = \phi^{(0)}$ and $\phi^{(1)}_{\text{corr}} = \phi^{(1)}$. The result of phase-correcting Figure 5.30a is shown in Figure 5.30b.

Spectral phase correction is not a mathematically exact procedure for eliminating dispersion-mode contributions from the spectrum. This is because Equation 5.19 indicates the *centre frequency* Ω_ℓ of the peaks, whereas Equation 5.20 employs the *spectral frequency coordinate* Ω. As a result, frequency-dependent spectral phase correction does not work precisely in the wings of the peaks or in regions where peaks overlap. The distortions can be serious when the linear term $\phi^{(1)}$ is large, as illustrated in Figure 5.30c and d. Experimental NMR spectra often display such 'rolling baselines', which are frequently removed by a further stage of data correction (not dealt with here).

To summarize, a linear frequency dependence of the phase of the spectral peaks is expected in many simple experiments. If the frequency dependence is small, then the spectrum may be corrected, to a good approximation. In the case of an arbitrary excitation sequence, on the other hand, there is no simple relation-

ship between the phases of the peaks and their centre frequencies. No general correction of the spectrum is possible in this case.

5.9 Two-Dimensional Spectroscopy

5.9.1 Two-dimensional signal surface

Suppose that an *arrayed* experiment is conducted, as described in Section 5.6. The result is a two-dimensional data matrix as a function of two time variables, which can be visualized as a three-dimensional surface:

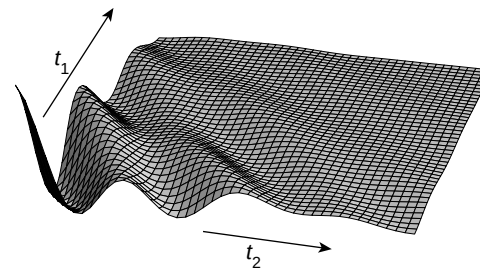


Figure 5.31
A two-dimensional signal surface.

As this book progresses, we will see that in many important cases the signal surface has the following general form:

$$s(t_1, t_2) = \sum_{\ell} s_\ell(t_1, t_2) \tag{5.21}$$

where

$$s_\ell(t_1, t_2) = a_\ell \exp\left\{\left(\mathrm{i}\Omega_\ell^{(1)} - \lambda_\ell^{(1)}\right) t_1 + \left(\mathrm{i}\Omega_\ell^{(2)} - \lambda_\ell^{(2)}\right) t_2\right\}$$

This is a straightforward extension of Equation 5.3: the two-dimensional signal is a sum of individual contributions ℓ; each contribution has a complex amplitude a_ℓ, with frequency $\Omega_\ell^{(1)}$ in the t_1 dimension and frequency $\Omega_\ell^{(2)}$ in the t_2 dimension, and peakwidth parameters $\lambda_\ell^{(1)}$ and $\lambda_\ell^{(2)}$ in the two dimensions.

Many examples of two-dimensional signal surfaces with the form of Equation 5.21 are given in later chapters, where the origin of the signal $s(t_1, t_2)$ is described in detail for several cases.

The two-dimensional signal surface expressed by Equation 5.21 is complex; it has a real and an imaginary part, just as the ordinary one-dimensional signal $s(t)$ has a real and an imaginary part.

5.9.2 Two-dimensional fourier transformation

The two-dimensional signal surface may be subjected to two-dimensional FT, which is defined through the double integral

$$S(\Omega_1, \Omega_2) = \int_0^\infty \mathrm{d}t_1 \int_0^\infty \mathrm{d}t_2 \; s(t_1, t_2) \exp\{-\mathrm{i}\,(\Omega_1 t_1 + \Omega_2 t_2)\} \tag{5.22}$$

in direct analogy to Equation 5.7. The function $S(\Omega_1, \Omega_2)$ is called the *two-dimensional spectrum*. It is a function of two frequency variables and may also be visualized as a surface:

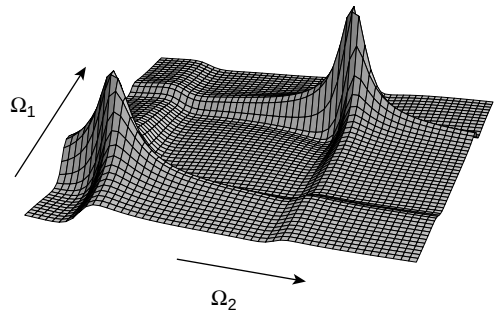


Figure 5.32
A two-dimensional spectrum.

The two-dimensional spectrum has peaks at the coordinates $(\Omega_\ell^{(1)}, \Omega_\ell^{(2)})$, corresponding to the frequencies of the underlying oscillating signal contributions s_ℓ.

Just as in the one-dimensional case, the real and imaginary parts of the two-dimensional spectrum are both derived from the real and imaginary parts of the time-domain signal:

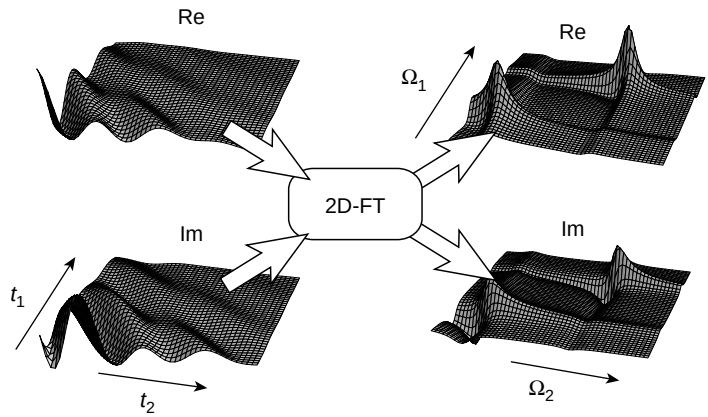


Figure 5.33
The relationship between the real and imaginary parts of $s(t_1, t_2)$ and $S(\Omega_1, \Omega_2)$.

Two-dimensional spectra may be depicted either as surface plots (as in the above figures) or as contour plots (as in a topographic map). The same two-dimensional NMR spectrum is shown in both representations in Figure 5.34.

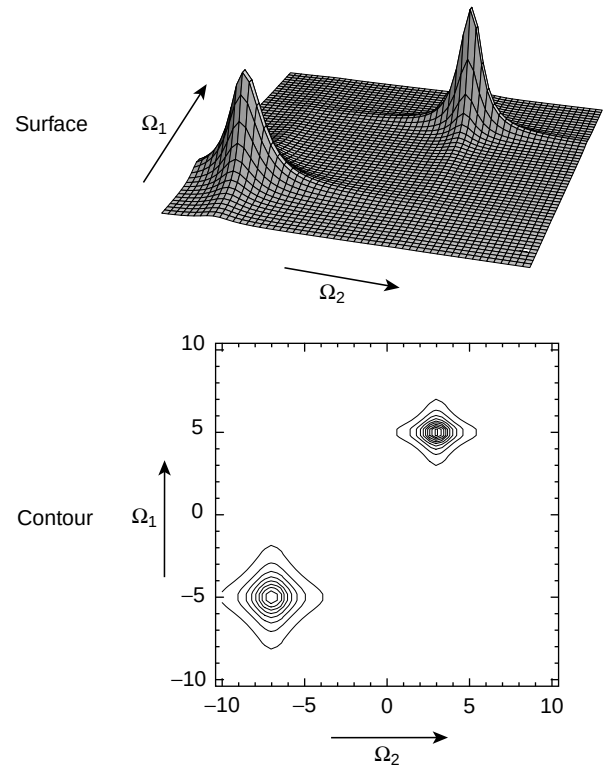


Figure 5.34
Two different representations of a two-dimensional spectrum.

5.9.3 Phase twist peaks

Evaluation of Equation 5.22 for the signal in Equation 5.21 gives the following form of the two-dimensional spectrum:

$$S(\Omega_1, \Omega_2) = \sum_{\ell} S_\ell(\Omega_1, \Omega_2) \tag{5.23}$$

where

$$S_\ell(\Omega_1, \Omega_2) = a_\ell \, \mathcal{L}(\Omega_1, \Omega_2; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}, \Omega_\ell^{(2)}, \lambda_\ell^{(2)})$$

The function $\mathcal{L}(\Omega_1, \Omega_2; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}, \Omega_\ell^{(2)}, \lambda_\ell^{(2)})$ is called the *two-dimensional complex Lorentzian*. It is equal to a product of two one-dimensional complex Lorentzians:

$$\mathcal{L}(\Omega_1, \Omega_2; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}, \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) = \mathcal{L}(\Omega_1; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}) \mathcal{L}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) \tag{5.24}$$

The one-dimensional complex Lorentzian peakshape is specified in Equation 5.12.

How does the complex Lorentzian look? The *real* part of the two-dimensional complex Lorentzian may be evaluated as follows:

$$\mathrm{Re}\{\mathcal{L}_1 \mathcal{L}_2\} = \mathrm{Re}\{\mathcal{L}_1\} \, \mathrm{Re}\{\mathcal{L}_2\} - \mathrm{Im}\{\mathcal{L}_1\} \, \mathrm{Im}\{\mathcal{L}_2\}$$

using obvious abbreviations. This leads to

$$\mathrm{Re}\{\mathcal{L}(\Omega_1, \Omega_2; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}, \Omega_\ell^{(2)}, \lambda_\ell^{(2)})\} =$$

$$\mathcal{A}(\Omega_1; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}) \mathcal{A}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) - \mathcal{D}(\Omega_1; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}) \mathcal{D}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) \tag{5.25}$$

This two-dimensional function is shown below:

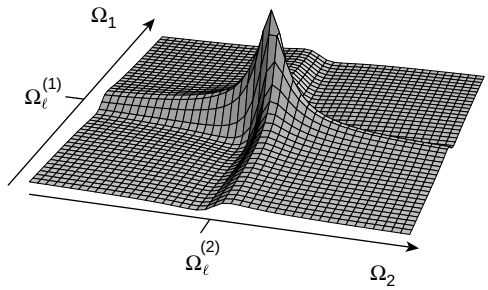


Figure 5.35
The real part of a complex two-dimensional Lorentzian.

As can be seen, it is a peculiar mixture of absorption and dispersion-mode one-dimensional Lorentzians. Near the centre of the peak, slices appear as absorption Lorentzians, but slices appear as dispersion Lorentzians far from the centre.

The *imaginary* part of the two-dimensional complex Lorentzian may be evaluated through the following:

$$\mathrm{Im}\{\mathcal{L}_1\mathcal{L}_2\} = \mathrm{Re}\{\mathcal{L}_1\}\,\mathrm{Im}\{\mathcal{L}_2\} + \mathrm{Im}\{\mathcal{L}_1\}\,\mathrm{Re}\{\mathcal{L}_2\}$$

leading to

$$\mathrm{Im}\{\mathcal{L}(\Omega_1, \Omega_2; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}, \Omega_\ell^{(2)}, \lambda_\ell^{(2)})\} =$$
$$\mathcal{A}(\Omega_1; \Omega_\ell^{(1)}, \lambda_\ell^{(1)})\mathcal{D}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) + \mathcal{D}(\Omega_1; \Omega_\ell^{(1)}, \lambda_\ell^{(1)})\mathcal{A}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)})$$

This function is plotted below:

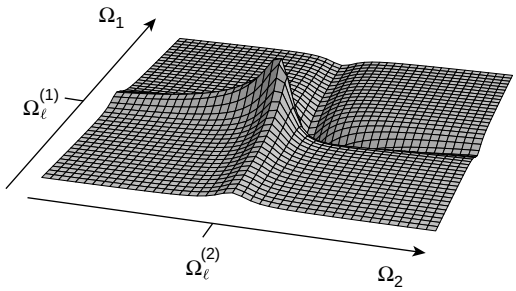


Figure 5.36
The imaginary part of a complex two-dimensional Lorentzian.

This is also a peculiar mixture of absorption and dispersion-mode one-dimensional Lorentzians, but this time with dispersion in the middle and absorption at the edges.

The strange appearance of these functions has led to the name *phase twist peaks*. Just as the absorption and dispersion Lorentzians are the 'fundamental' peakshapes in one-dimensional spectroscopy, so these weird phase twists are the 'fundamental' peakshapes in two-dimensional spectroscopy.

These phase twist peakshapes are undesirable because of their long dispersion tails. In contrast to one-dimensional FT NMR, the dispersion components in two-dimensional spectra may *not* be removed by phase correction of the spectrum, at least not in any simple way. As shown in Equation 5.35, the real part of the two-dimensional complex Lorentzian contains both absorption and dispersion mode shapes, inextricably entangled.

5.9.4 Pure absorption two-dimensional spectra

In *some* forms of two-dimensional spectroscopy, it is possible to suppress dispersion mode contributions to the two-dimensional spectrum, by combining together *two* different data sets in a careful way. The result is a spectrum containing *two-dimensional absorption Lorentzian* peaks, of the following form:

$$\mathcal{A}(\Omega_1, \Omega_2; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}, \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) = \mathcal{A}(\Omega_1; \Omega_\ell^{(1)}, \lambda_\ell^{(1)})\mathcal{A}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) \tag{5.26}$$

This has a much more attractive appearance, as shown in Figure 5.37.

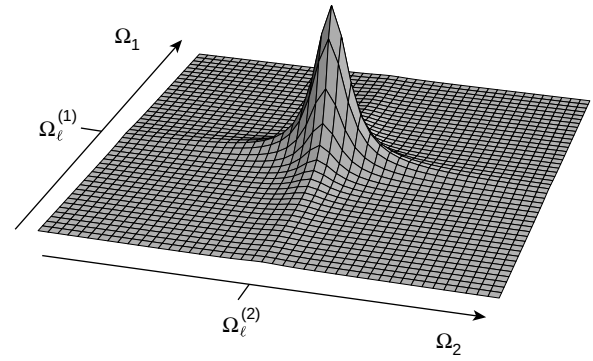


Figure 5.37
A pure absorption two-dimensional peak.

The removal of the dispersion contributions greatly improves the spectral resolution.

The engineering of absorption mode peakshapes in multidimensional spectra is a very important component of modern NMR technology. I now discuss the so-called *States method* for accomplishing this task[10] ('States' is the surname of one of the inventors of the technique).

The key to the States method is the concept of *amplitude modulation*. Consider the general form of the two-dimensional signal matrix, given in Equation 5.21:

$$s(t_1, t_2) = \sum_\ell a_\ell \exp\left\{\left(\mathrm{i}\Omega_\ell^{(1)} - \lambda_\ell^{(1)}\right) t_1 + \left(\mathrm{i}\Omega_\ell^{(2)} - \lambda_\ell^{(2)}\right) t_2\right\}$$

This represents a superposition of two-dimensional signals, with different complex amplitudes a_ℓ, frequencies $(\Omega_\ell^{(1)}, \Omega_\ell^{(2)})$, and decay rate constants $(\lambda_\ell^{(1)}, \lambda_\ell^{(2)})$. In general, there is no particular relationship between the amplitudes, frequencies, and decay rate constants of the different peaks.

However, in some special cases, the two-dimensional signal fulfils the following conditions:

1. The amplitudes a_ℓ are all *real numbers* ($a_\ell = a_\ell^*$).

2. For every component with frequency coordinates $(\Omega_\ell^{(1)}, \Omega_\ell^{(2)})$, there is a 'mirror image' component with identical amplitude a_ℓ and decay parameters $(\lambda_\ell^{(1)}, \lambda_\ell^{(2)})$, but with frequency coordinates $(-\Omega_\ell^{(1)}, \Omega_\ell^{(2)})$.

This means that the two-dimensional spectrum contains peaks that are arranged symmetrically about the $\Omega_1 = 0$ axis, as shown schematically below:

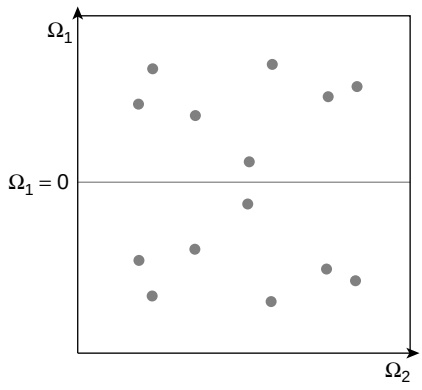


Figure 5.38
A two-dimensional spectrum with symmetry about the $\Omega_1 = 0$ axis.

In the time domain, this implies that the two-dimensional signal may be written as

$$s^{\cos}(t_1, t_2) = \sum_{\ell} a_\ell \cos(\Omega_\ell^{(1)} t_1)\ \exp\{-\lambda_\ell^{(1)} t_1\} \exp\{\left(\mathrm{i}\Omega_\ell^{(2)} - \lambda_\ell^{(2)}\right) t_2\} \tag{5.27}$$

This equation means that the NMR signal is *amplitude modulated* with respect to the evolution interval t_1. As t_1 increases, the amplitude of the NMR signal oscillates up and down, but its phase never changes. The form of the modulation is a cosine wave, hence the superscript 'cos' in Equation 5.27.

The effect of amplitude modulation is seen most clearly if a one-dimensional Fourier transform is calculated with respect to the time variable t_2, according to

$$S(t_1, \Omega_2) = \int_0^\infty \mathrm{d}t_2\ s(t_1, t_2) \exp\{-\mathrm{i}\Omega_2 t_2\}$$

Applied to the cosine-modulated signal, we get

$$S^{\cos}(t_1, \Omega_2) = \sum_{\ell} a_\ell \cos(\Omega_\ell^{(1)} t_1) \exp\{-\lambda_\ell^{(1)} t_1\}\ \mathcal{L}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) \tag{5.28}$$

The real part of $S^{\cos}(t_1, \Omega_2)$ for a typical cosine-modulated signal is shown below:

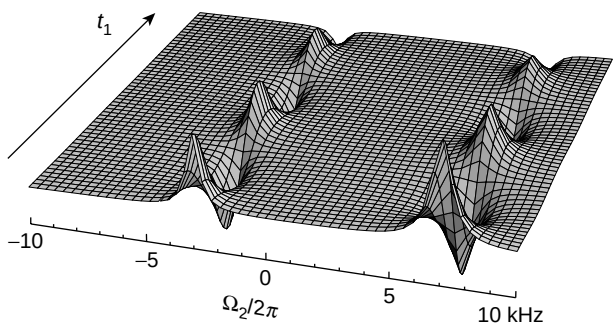


Figure 5.39
The real part of the cosine-modulated signal $S^{\cos}(t_1, \Omega_2)$.

In this example, the two-dimensional signal contains two signal components, with parameters $\{\Omega_\ell^{(1)}/2\pi, \Omega_\ell^{(2)}/2\pi\} = \{5\,\text{kHz}, -3\,\text{kHz}\}$ (on the left) and $\{-5\,\text{kHz}, +7\,\text{kHz}\}$ (on the right), as well as their mirror images at $\{\Omega_\ell^{(1)}/2\pi, \Omega_\ell^{(2)}/2\pi\} = \{-5\,\text{kHz}, -3\,\text{kHz}\}$ and $\{+5\,\text{kHz}, +7\,\text{kHz}\}$. Notice how the peak amplitudes are at a maximum for $t_1 = 0$, and oscillate as t_1 is increased, whereas the peak phases remain constant.

Now suppose that a *second* two-dimensional experiment is performed that generates an amplitude modulated two-dimensional signal, where the modulation function is a *sine wave*:

$$s^{\sin}(t_1, t_2) = \sum_\ell a_\ell \sin(\Omega_\ell^{(1)} t_1)\ \exp\{-\lambda_\ell^{(1)} t_1\} \exp\left\{\left(\mathrm{i}\Omega_\ell^{(2)} - \lambda_\ell^{(2)}\right) t_2\right\} \tag{5.29}$$

FT with respect to t_2 provides a signal of the form.

$$S^{\sin}(t_1, \Omega_2) = \sum_\ell a_\ell \sin(\Omega_\ell^{(1)} t_1) \exp\{-\lambda_\ell^{(1)} t_1\}\ \mathcal{L}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) \tag{5.30}$$

The real part of $S^{\sin}(t_1, \Omega_2)$ is shown below:

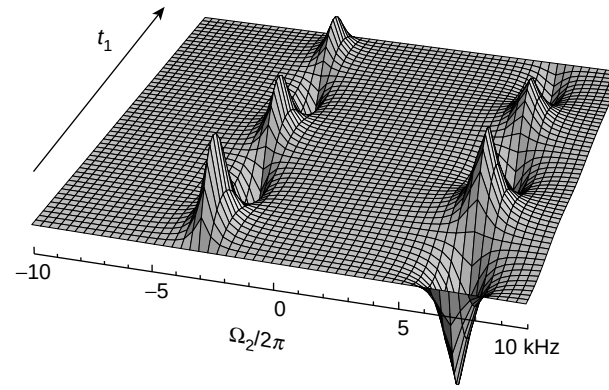


Figure 5.40
The real part of the sine-modulated signal $S^{\sin}(t_1, \Omega_2)$.

Note that the signal amplitude is zero for $t_1 = 0$. As t_1 is increased, the signal amplitude increases in the positive sense for the component with $\{\Omega_\ell^{(1)}/2\pi, \Omega_\ell^{(2)}/2\pi\} = \{5\,\text{kHz}, -3\,\text{kHz}\}$ (since $\Omega_\ell^{(1)}$ is positive), and it increases in the negative sense for the component with $\{\Omega_\ell^{(1)}/2\pi, \Omega_\ell^{(2)}/2\pi\} = \{-5\,\text{kHz}, +7\,\text{kHz}\}$ (since $\Omega_\ell^{(1)}$ is negative in this case).

Typically, the pulse sequences for generating $s^{\cos}$ and $s^{\sin}$ are related by changing the phase of one or more r.f. pulses. Some concrete examples are given in Sections 16.1, 19.7 and 20.6.

If two-dimensional Fourier transforms of the signals $s^{\cos}$ and $s^{\sin}$ are taken in the usual way, then one obtains an unattractive mess of overlapping phase twist peaks. The plots below show the real parts of the two-dimensional spectra $S^{\cos}(\Omega_1, \Omega_2)$ and $S^{\sin}(\Omega_1, \Omega_2)$, generated by two-dimensional FT of $s^{\cos}(t_1, t_2)$ and $s^{\sin}(t_1, t_2)$:

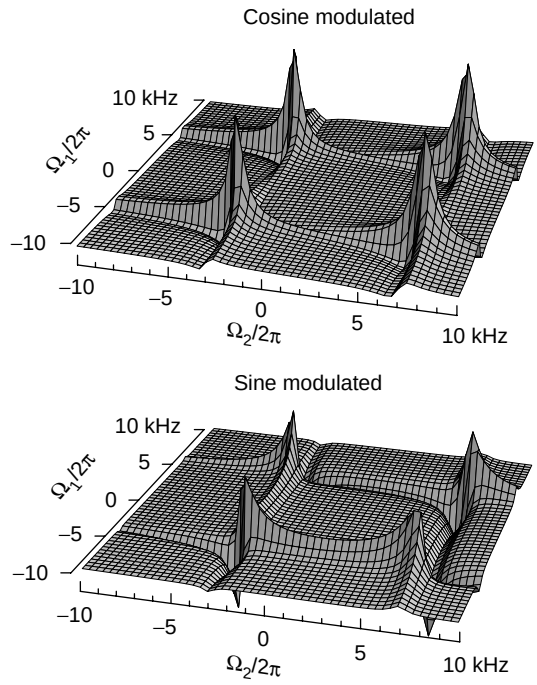


Figure 5.41
Two-dimensional spectra of the cosine- and sine-modulated signals.

These spectra are full of dispersion modes and spectral overlap and would not be of much use.

In the States method, the results of the two different two-dimensional experiments are combined in a special way. The real parts of the data matrices $S^{\cos}(t_1, \Omega_2)$ and $S^{\sin}(t_1, \Omega_2)$ are put together to make a new complex signal,[11] called here $S_{\text{States}}(t_1, \Omega_2)$:

$$S_{\text{States}}(t_1, \Omega_2) = \text{Re}\{S^{\cos}(t_1, \Omega_2)\} + \text{i Re}\{S^{\sin}(t_1, \Omega_2)\} \tag{5.31}$$

This 'hybrid' signal has the following mathematical form:

$$S_{\text{States}}(t_1, \Omega_2) = \sum_{\ell} a_\ell \exp\{(\text{i}\Omega_\ell^{(1)} - \lambda_\ell^{(1)})t_1\}\, \mathcal{A}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) \tag{5.32}$$

FT of the States signal with respect to t_1 gives a two-dimensional spectrum:

$$\begin{aligned} S_{\text{States}}(\Omega_1, \Omega_2) &= \int_0^\infty \text{d}t_1 \;\; S_{\text{States}}(t_1, \Omega_2) \exp\{-\text{i}\Omega_1 t_1\} \\ &= \sum_{\ell} a_\ell \mathcal{L}(\Omega_1; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}) \mathcal{A}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)}) \end{aligned} \tag{5.33}$$

The real part of this spectrum is in pure absorption:

$$\text{Re}\{S_{\text{States}}(\Omega_1, \Omega_2)\} = \sum_{\ell} a_\ell \mathcal{A}(\Omega_1; \Omega_\ell^{(1)}, \lambda_\ell^{(1)}) \mathcal{A}(\Omega_2; \Omega_\ell^{(2)}, \lambda_\ell^{(2)})$$

as shown below:

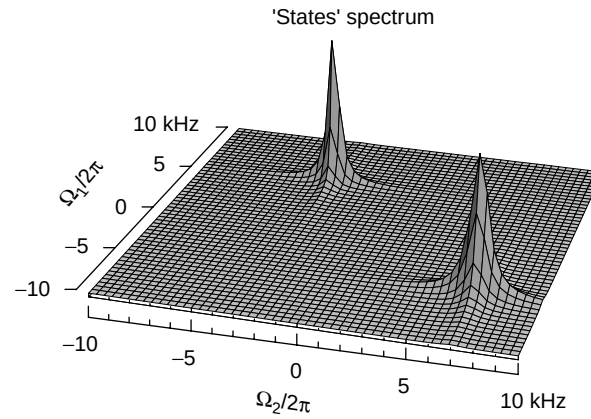


Figure 5.42
The real part of the States two-dimensional spectrum.

The 'image' peaks are eliminated and the peaks are in pure absorption. Note that the signs of the Ω_1 coordinates appear correctly for the two spectral components $\{\Omega_\ell^{(1)}/2\pi, \Omega_\ell^{(2)}/2\pi\} = \{5\,\text{kHz}, -3\,\text{kHz}\}$ and $\{-5\,\text{kHz}, +7\,\text{kHz}\}$.

To summarize, the States data acquisition/data processing scheme involves the following steps:

1. Design the two-dimensional NMR experiment so as to provide a cosine-modulated two-dimensional signal, of the form given in Equation 5.27.

2. Perform this experiment so as to acquire a data matrix $s^{\cos}(t_1, t_2)$.

3. Compute the Fourier transform with respect to t_2 to obtain the data matrix $S^{\cos}(t_1, \Omega_2)$ (Equation 5.28).

4. Construct a different two-dimensional NMR experiment so as to provide a sine-modulated two-dimensional signal, of the form given in Equation 5.29.

5. Perform this second experiment so as to acquire a data matrix $s^{\sin}(t_1, t_2)$.

6. Compute the Fourier transform with respect to t_2 to obtain the data matrix $S^{\sin}(t_1, \Omega_2)$ (Equation 5.30).

7. Combine the real parts of $S^{\cos}(t_1, \Omega_2)$ and $S^{\sin}(t_1, \Omega_2)$ as a complex pair, in order to obtain the data matrix $S_{\text{States}}(t_1, \Omega_2)$ (Equation 5.31).

8. Compute the Fourier transform with respect to t_1 to obtain the two-dimensional spectrum $S_{\text{States}}(\Omega_1, \Omega_2)$ (Equation 5.33).

9. The real part of this two-dimensional spectrum is in pure absorption mode.

A flow chart for the States procedure is shown in Figure 5.43.

Not all two-dimensional experiments are amenable to States data acquisition and processing. The conditions for cosine and sine modulation are quite stringent and are not satsified in all two-dimensional experiments. However, the practical advantages of pure absorption two-dimensional spectra are so great that essentially all *popular* two-dimensional experiments employ the States method (or an equivalent procedure; see Note 10).

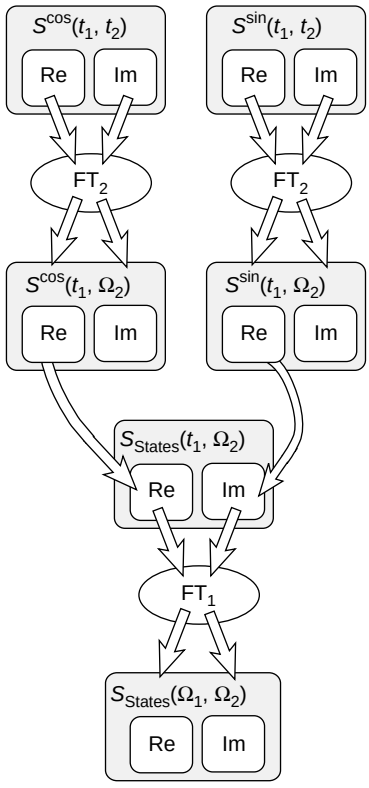


Figure 5.43
A flow chart for the 'States' procedure (see Note 10).

The schemes described in this section are known under a variety of names: *quadrature detection in the second dimension*, *pure phase two-dimensional spectroscopy*, and *phase-sensitive two-dimensional spectroscopy*. Some of these terms do not seem to be too meaningful (why 'phase sensitive' for instance?), but they are used prolifically and more or less interchangeably in the NMR literature.

5.10 Three-Dimensional Spectroscopy

A three-dimensional 'data cube' $s(t_1, t_2, t_3)$ may be acquired by incrementing the values of two pulse sequence delays t_1 and t_2 independently, while acquiring the signal as a function of t_3. The three-dimensional spectrum $S(\Omega_1, \Omega_2, \Omega_3)$ is computed by applying a three-dimensional Fourier transform to the signal $s(t_1, t_2, t_3)$:

$$S(\Omega_1, \Omega_2, \Omega_3) = \int_0^\infty \mathrm{d}t_1 \int_0^\infty \mathrm{d}t_2 \int_0^\infty \mathrm{d}t_3 \; s(t_1, t_2, t_3) \exp\{-\mathrm{i}\,(\Omega_1 t_1 + \Omega_2 t_2 + \Omega_3 t_3)\} \tag{5.34}$$

The three-dimensional spectrum may be visualized as a cube, with peaks suspended at particular frequency coordinates. An experimental example is shown in Figure 5.44.

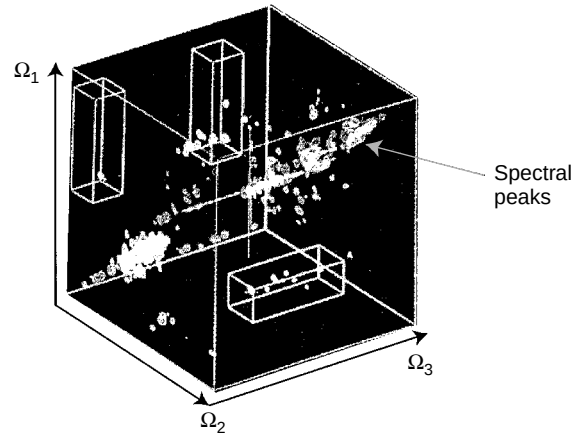


Figure 5.44
A three-dimensional spectrum. Adapted from C. Griesinger *et al.*, *J. Magn. Reson.* **84**, 14 (1989). (Copyright, Academic Press.)

In practice, such fully three-dimensional displays are rather difficult to use. It is more common to examine a set of sections of the three-dimensional spectrum, each of which is a two-dimensional spectrum.

Three- and higher-dimensional spectroscopies are an important feature of modern biomolecular NMR. However, I will not treat them further in this book.

Notes

1. Since the main source of r.f. noise is often the motion of electrons in the coil, the signal-to-noise ratio of NMR spectrometers may be considerably increased by cooling the coil and associated electronics to cryogenic temperatures (usually in the vicinity of 20 K). Probes equipped with such devices are called *cryoprobes*.

2. The most common procedure is to multiply the NMR signal by a *matched weighting function* before applying the Fourier transform. To obtain the optimal signal-to-noise ratio in the final spectrum, the weighting function should match the envelope of the NMR signal exactly: where the NMR signal is large, the weighting function should also be large, and where the NMR signal is small, the weighting function should also be small. Matched weighting accentuates the NMR signal at those times when it has the best chance of dominating the noise, and it reduces the noise at those times when the signal is weak. However, the gain in signal-to-noise does not come for free: it is always accompanied by a loss in spectral resolution. See *Further Reading*.

3. 'Ultrafast' forms of two-dimensional spectroscopy have been invented that do not require arrayed signal acquisition: instead of performing a set of consecutive experiments with different timings, the 'ultrafast' experiments use pulsed magnetic field gradients to perform many experiments in parallel in different parts of the sample tube. See for example L. Frydman *et al.*, *Proc. Natl. Acad. Sci. USA* **99**, 15 858–15 862 (2002) and M. Gal *et al.*, *J. Am. Chem. Soc.* **128**, 951–956 (2006).

4. There are a number of subtle effects that cause the peak frequencies and damping constants to depend on what happens *before* the signal is detected. All of these effects involve some form of feedback from the macroscopic nuclear spin magnetization to the magnetic fields acting on the spins themselves. The effects called *radiation damping* and the *nuclear demagnetizing field* fall into this category. For a review of this subject, see M. H. Levitt, *Concepts in Magn. Reson.* **8,** 77 (1996) and references therein. I ignore these effects in this book.

5. In practice, the FT is computed from the set of complex points sampled by the two ADCs. This numerical computation differs from Equation 5.7 in a number of respects. First, the value of the signal is available only up to a maximum time τ_{acq}, not to infinity. Second, the signal is available as a set of discrete samples, evenly spaced along the time axis, rather than as a continuous mathematical function. The discrepancy between the 'discrete FT' performed by the computer and the 'continuous FT' indicated by Equation 5.7 leads to certain spectral effects, such as *truncation distortions* and the *folding* of peaks. Fortunately, most of the dangerous effects of the discrete FT may be avoided by simple experimental precautions. See *Further Reading* for a discussion of these technical issues.

6. A common definition of the dispersion Lorentzian has the opposite sign.

7. The terms 'absorption' and 'dispersion' are purely historical in origin and have no physical significance in the ordinary NMR context. They are *not* related to the absorption of energy or the dispersion of electromagnetic waves. The term 'emission', which is sometimes used to indicate a negative absorption peak, is not recommended.

8. The FFT algorithm was invented by Cooley and Tukey in the 1950s, and has had a crucial influence on the development of NMR. Tukey made a further memorable contribution: He invented the term 'software'!

9. The word 'phase' has a number of different uses in NMR:
   - The term *phase of matter* indicates whether the substance is a solid, isotropic liquid, anisotropic liquid, or a gas.
   - The *phase of an r.f. signal* indicates the time origin of the oscillation; this is the meaning of phase in, for example, Section 4.2.1.
   - The *phase of a complex number* indicates the value of ϕ in expressions of the form $a = |a| \exp\{i\phi\}$.
   - The *phase of a spectral peak* indicates the proportions of absorption and dispersion peakshape contributions, as discussed in Section 5.8.4.

   The last three meanings are closely related, as explained in this book.

10. The States method was originally described in D. J. States, R. A. Haberkorn, D. J. Ruben *J. Magn. Reson.* **48,** 286(1982). There is a slightly different data acquisition/processing scheme called *time-proportional phase incrementation* (TPPI) that also suppresses image peaks and leads to pure absorption two-dimensional spectra (D. Marion, K. Wüthrich, *Biochem. Biophys. Res. Commun.* **113,** 967 (1983)). These methods are not totally equivalent, and in some circumstances the TPPI method is slightly superior. As described in the text, the States method combines two data sets that differ only in the phase of the excited coherences. The evolution frequency of the excited coherences is unchanged. In the TPPI method, on the other hand, the phase of the excited coherences is linked to incrementation of the evolution interval, generating an effective frequency change of the coherences. This turns out to be useful. For example, consider the case where longitudinal magnetization is present during the evolution interval t_1. This magnetization does not evolve (neglecting relaxation), and so generates 'axial peaks' on the $\Omega_1 = 0$ axis. In the States method, these axial peaks sit in the middle of the single-quantum spectrum and must usually be removed by a further stage of phase cycling (see Appendix A.11). In the TPPI method, on the other hand, the effective frequency shift of the coherences displaces the interesting two-dimensional peaks from the vicinity of the $\Omega_1 = 0$ axis to the edges of the two-dimensional spectrum, so that, in many circumstances, a further level of phase cycling is unnecessary. In such cases, the TPPI method is more economical of instrument time than the States method. Despite this subtle advantage of TPPI, I have chosen to emphasize the States method in this book, since it is somewhat easier to explain.

11. For historical reasons, most current practical implementations of the States method generate two data sets, one of which is cosine modulated and the second of which is sine modulated, but *with an inverted sign*. This requires that the sign of the imaginary part must also be inverted when combining the data

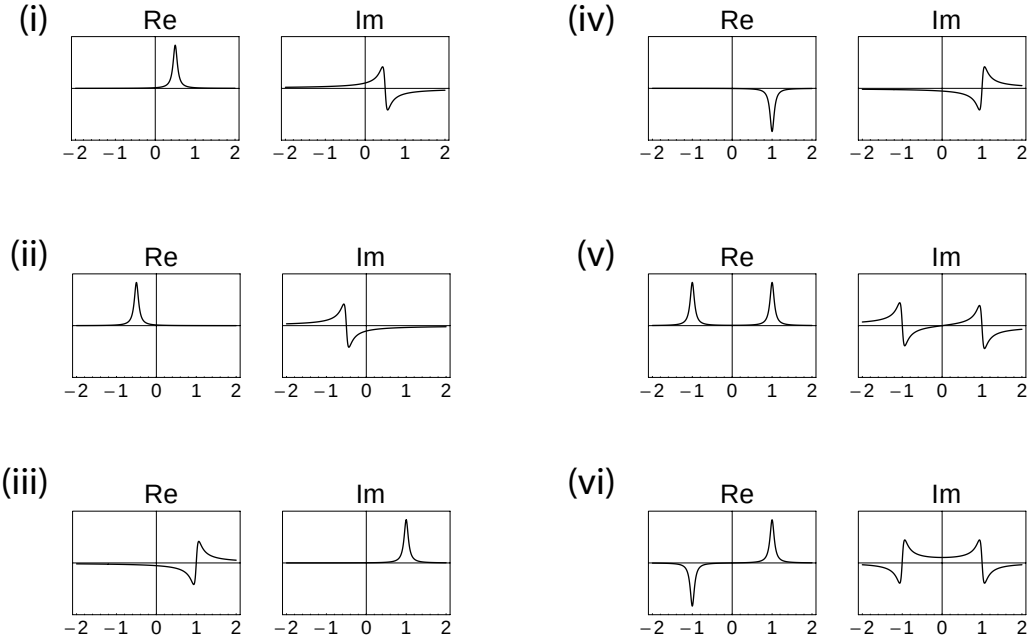
sets according to Equation 5.31. In this book, I have imposed a more natural phase convention for the States procedure. All the phase cycles given in this book are compatible with Equation 5.31, as written.

Further Reading

- For a discussion of post-processing of the NMR data to optimize sensitivity or resolution, see J. Keeler, *Understanding NMR Spectroscopy*, Wiley, Chichester, 2005, and A. E. Derome, *Modern NMR Techniques in Chemistry Research*, Pergamon Press, Oxford, 1990.
- For a discussion of the equivalence between the States method and TPPI, and a more thorough treatment of two-dimensional data processing, see R. R. Ernst, G. Bodenhausen and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford, 1987.
- For more information on FT and some other data processing methods, see: A. G. Marshall and F. R. Verdun, *Fourier Transforms in NMR, Optical, and Mass Spectrometry*, Elsevier, Amsterdam, 1990 ; R. N. Bracewell, *The Fourier Transform and its Applications*, McGraw-Hill, New York, 1986; E. O. Brigham, *The Fast Fourier Transform and its Applications*, Prentice Hall, Upper Saddle River, NJ, USA, 1988.
- For the mathematical principles underlying the FFT computer algorithm, see G. Strang, *Linear Algebra and its Applications*, 3rd edition, Harcourt Brace Jovanovich, San Diego, CA, USA, 1988.
- For a review of fast data acquisition and processing schemes, see R. Freeman and E. Kupče, *Concepts Magn. Reson.* **23**, 63–75 (2004).

Exercises

5.1 Sketch roughly the Fourier transforms of the following time-domain functions (the horizontal axes show time in units of seconds):



5.2 The T_1 of a certain sample decreases with increasing temperature. At 20 °C it is necessary to leave 10 s between transients, whereas at 40 °C only 5 s is required. Suppose that 3 h of NMR spectrometer time is available and that the instrument is already set up for operation at 20 °C. It takes 1 h to warm the sample to 40 °C and to stabilize the temperature. Assume that the NMR signals are identical at the two temperatures. What is the best strategy for acquiring the signals: running for 3 h at 20 °C, or warming the sample and running for 2 h at 40 °C?