Chapter 2. The Physics of Magnetic Resonance Imaging

2.1. Introduction

The origins of the Nuclear Magnetic Resonance (NMR) signal and how it is manipulated to form images are the subjects of this chapter. MRI is a very flexible and complex technology, and can produce many different kinds of images. This chapter will aid you in understanding the terminology used in the literature and the scanner room, and in understanding the technological tradeoffs and limitations. The scanning methods used in MRI and fMRI are still evolving rapidly, so this knowledge is essential background that will enable you to keep current as new techniques are developed.

Capsule Summary

Protons by themselves, as in a hydrogen nucleus, are slightly magnetic. When immersed in a large magnetic field, protons tend to line up with this field. As a result, water (H_2O) in a magnetic field becomes slightly magnetized. When undisturbed, the magnetization of the water protons is lined up with the externally applied field. Microwave radiofrequency radiation (RF) applied to the water can disturb this alignment of the water proton magnetization.. When the applied RF is turned off, the magnetization relaxes back to alignment with the external field. During the realignment time, the protons re-emit RF, which can be detected by a sensitive receiver placed around the object or subject. The frequency of the RF emitted from a given location depends on the strength of the large external magnetic field at that location. By making the externally applied field have different strengths at different locations, and by detecting the emitted RF signal over a range of frequencies, the strength of the RF signal originating from different locations can be reconstructed. The result is an image of the emitted RF signal intensity across the object in the field. Different kinds of manipulations of the applied RF and of the external magnetic field result in different tissue properties being emphasized in the emitted RF signal strength and so in the image.

Magnetic Fields

Most of the discussion in this chapter concerns the *magnetic fields* inside tissue, some of which are intrinsic to the tissue and some of which are applied by the MRI scanner hardware. A *field* is simply some quantity that varies over a spatial region, and may also vary in time. The air temperature over North America is an example of a field. A *vector* is a quantity that has both magnitude and direction. The wind velocity over North America is an example of a vector field—knowing the wind direction is as important as knowing the wind speed for weather prediction. A magnetic field is a vector field that is defined by its effects on magnets: the field pushes and pulls on a magnet so as to make the magnet's North and South poles line up with the direction of the magnetic field at the location of the magnet (see Fig. 2.1).

In magnetic resonance imaging, the externally applied magnetic fields are at least as important as illumination is to optical imaging. The magnetic fields create the substance being imaged (magnetization: §2.2), make it emit detectable signals (excitation: §2.3), and manipulate the signals so that an image can be formed (slice selection and gradient scanning: §2.5). In addition, the weak intrinsic magnetic fields of the tissue being scanned strongly affect the emitted signals and the resulting image (relaxation and contrast: §2.4 and §2.6).

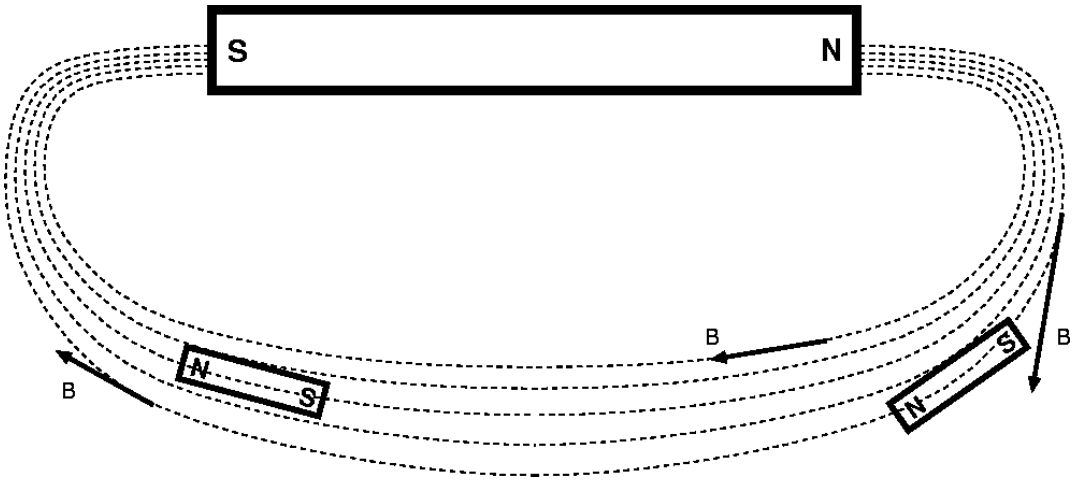

Figure 2.1. The large magnetic field applies a force that makes the smaller magnets line up with the field. At each point, the magnetic field **B** is a vector tangent to the "field lines" (only some of which are shown). For simplicity, the magnetic fields from the smaller magnets are not shown; in practice, these fields would add to **B**.

2.2. Creation of Magnetization M by the Magnetic Field B

Some atomic nuclei are magnetic. Each such nucleus is like a tiny weak bar magnet, with North and South poles. The most magnetic nucleus is a single proton—the hydrogen nucleus (^{1}H), which is ubiquitous in tissue, mostly in the form of H_2O. (Other magnetic nuclei that have been used in MRI include ^{13}C, ^{19}F, ^{23}Na, and ^{31}P.) Magnetic nuclei are often called *spins* in the MRI and NMR literature. This nomenclature is due to the connection between the quantum mechanical property called spin—analogous to the classical mechanical property of rotational angular momentum—and the magnetic strength of the nucleus.

Water is not normally magnetic, since the hydrogen protons are not lined up. The net magnetic field from protons pointing randomly in all directions is zero. The reason that the protons are not lined up is that there is no internal or external force that tends to align them. The energy of random thermal motions of the water molecules keeps the protons pointing in random directions.

Putting a water sample (or water-containing sample, such as a human subject) into a large magnetic field will make the protons tend to line up with the magnetic field, just as the magnetic field from a large bar magnet can be used to align a small bar magnet. This tendency is weak compared to the randomizing effect of thermal motions, so the amount of alignment at any given moment is very small. Applying a larger magnetic field will overcome the thermal agitation more, resulting in more protons being aligned. The net effect is that the water becomes slightly magnetized itself, and the amount of magnetization is proportional to the strength of the applied field (see Fig. 2.2).

The symbol for magnetic field is **B**; the unit of magnetic field strength is Tesla (T). Another unit that is sometimes used is the Gauss (G); 1 Gauss=10^{-4} Tesla. The strength of the Earth's magnetic field is about 5×10^{-5} Tesla (0.5 Gauss). The symbol for the strength of the large magnetic field of the scanner—in which the subject is immersed—is B_0 (non-boldface). Another term for B_0 is the *main field*. A typical magnetic field strength used for fMRI is B_0=1.5 Tesla, 30,000 times stronger than the Earth's field. The field **B** is sometimes called the *static field* since it does not change in time, or only changes slowly. It is important to realize that the static field includes not only the main field, but small additions and subtractions to it induced by the properties of the sample—these perturbations are very important, and are discussed in §2.4.

At 1.5 T, about 0.0005% of the protons in water are aligned with **B** at any given moment; the rest of the protons are pointing in random directions. At 3.0 T, 0.0010% (twice as many) of the protons would be aligned. Although these numbers are small, their net result is measurable, since the magnetic fields from the remaining randomly aligned protons add up to zero. An analogous effect is the wind: the average thermal speed of an air molecule is about 480 m/s (1080 miles/hour), but even a 5 mile per hour breeze is quite perceptible. A breeze is also a small collective result of a small net alignment superimposed on a larger field of randomness; in this case, alignment of the direction of motion of the air molecules.

The amount per unit volume that the object inside **B** is magnetized at any given place is called the *magnetization density* (usually just shortened to "magnetization"). The symbol for magnetization is **M**. Magnetic fields and magnetization are vectors (which is why their symbols are in boldface). When undisturbed, the **M** that results from **B** will be aligned with the direction of **B**, and the magnitude of **M** will be proportional to the magnitude of **B**. This situation is called *fully relaxed magnetization*. It typically takes 3-6 seconds for **M** to become fully relaxed (§2.4). The magnitude of fully relaxed magnetization is denoted by M_0. In tissue, M_0 is not spatially uniform (i.e., it depends on location), since different amounts of water are present in different types of tissue. Since the NMR signal is proportional to M_0 (§2.3), this is one way of distinguishing tissue types in NMR images.

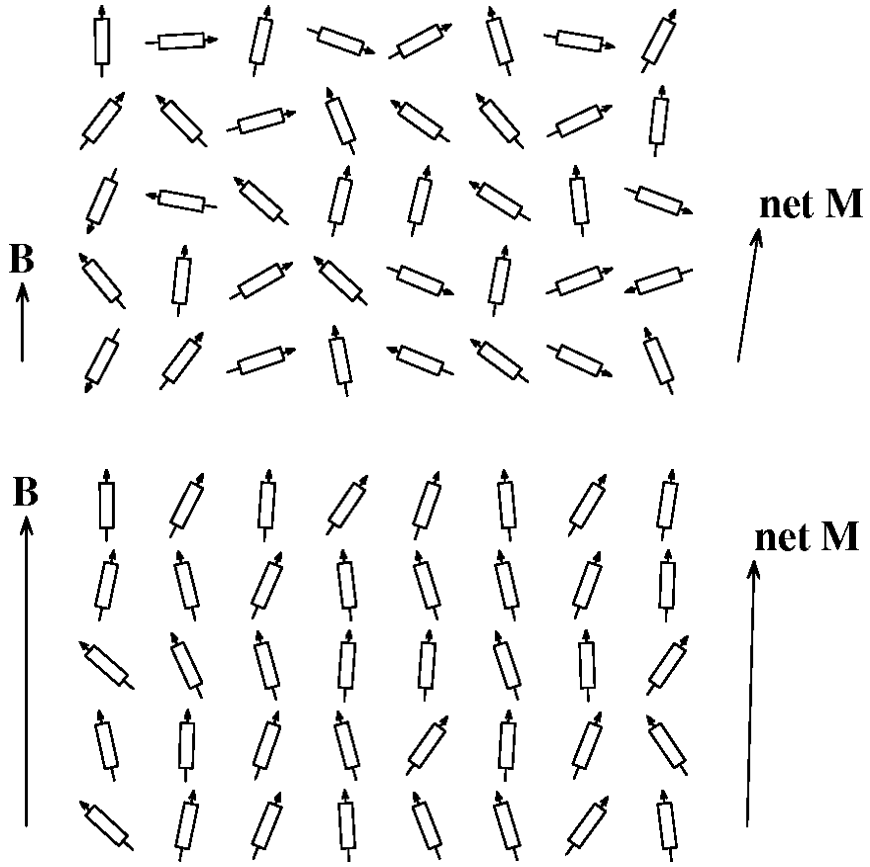

Figure 2.2. Nuclei being aligned by an external magnetic field **B** and also being misaligned by thermal agitation. Each box with an arrow represents one nuclear magnet; the arrow represents the strength and orientation of the magnetization of each nucleus. To the right is shown the net magnetization vector **M**, proportional to the sum of the individual nuclear magnetization vectors. (above) With a weak **B**, the amount of alignment is minimal—although the spins are not completely randomized here—and the net **M** is small. (below) With a larger **B**, the amount of alignment is greater and the net **M** increases. In reality, the number of water protons is vast (3×10^{16} per mm^3), and the actual amount of alignment is less than is shown here. The small component of **M** perpendicular to **B** shown is due to the tiny number of nuclei (40) in this numerical simulation. With a realistic number of nuclei, the component **M** perpendicular to **B** averages to zero.

2.3. Precession of M

What happens when M is not parallel to B

If the vector **M** is not parallel to **B**, then over an interval of a few seconds it will realign itself to point in the **B**-direction. It does not follow a simple path along the way. The behavior of **M** on its way back to the equilibrium situation is the subject of this and the next section.

The largest force on **M** causes it to rotate (or *precess*) clockwise around the **B**-direction, as shown in Fig. 2.3. The frequency with which **M** precesses is proportional to the strength of the magnetic field:

$$f = \gamma \cdot B \qquad [2.1]$$

Here, f is the frequency of precession (called the *Larmor* or *resonant frequency*), and is measured in Hertz: 1 Hertz (Hz) is one full revolution (360°) per second. The constant γ equals 42.54 MegaHertz/Tesla (MHz/T), or 4254 Hz/Gauss. (For magnetic nuclei other than protons, γ is smaller.)

At B_0=1.5 Tesla, f=63.81 MHz, which means that the direction of **M** spins around the **B**-direction 63,810,000 times in one second. In other words, **M** spins through 360° in 15.67 nanoseconds.

Precession of **M** is similar to the precession of a spinning gyroscope whose axis is not vertical. If the gyroscope were not spinning, it would simply fall over (i.e., try to become aligned with the gravitational field). The effect of its angular momentum is that the gravitational force downwards causes the rotational axis of the gyroscope itself to rotate sideways (see Fig. 2.4). Similarly, the effect of the magnetic force of **B** on the nuclei is to make them align with the magnetic field, but the spin angular momentum of the nuclei converts the effect into the precession of **M**.

During precession, **M** changes its direction rapidly and cyclically, but its length changes only very slowly. The forces that change the length of **M** and the forces that tend to realign the direction of **M** back with **B** are much smaller than the precessional force, and so operate over much longer time scales (milliseconds to seconds: thousands to millions of times slower than the precessional force). The effects of these forces are called *relaxation*, and are discussed in §2.4.

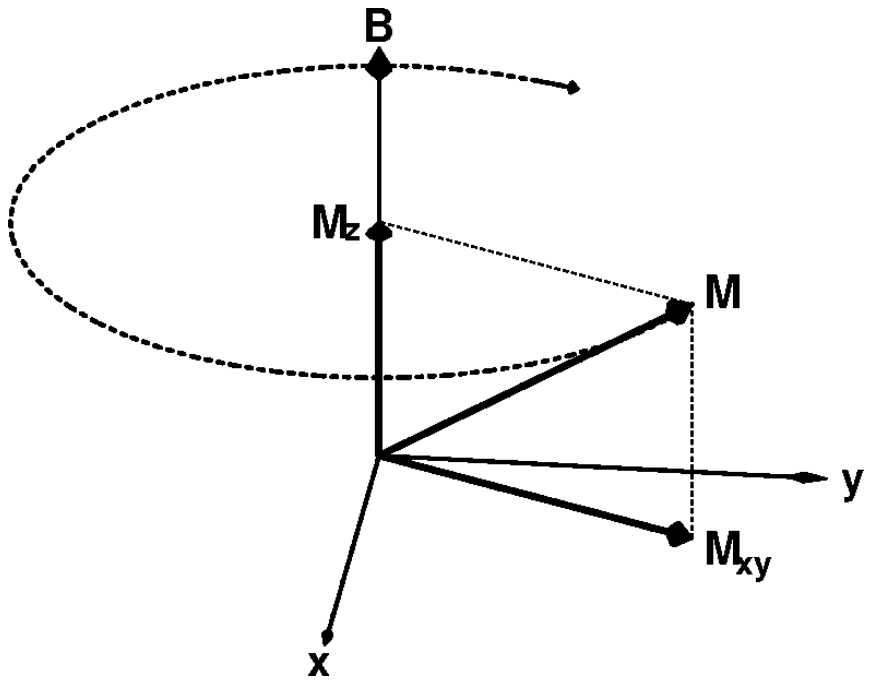

Figure 2.3. When the magnetization **M** is not aligned with the direction of the magnetic field **B**, the largest force on **M** makes the magnetization precess clockwise about the direction of **B**. The speed of this precession at each location is proportional to the size of **B** at that location. Not that as **M** precesses, M_z is unchanging but M_{xy} is oscillating. The length of **M** changes very slowly compared to the precession rate; in addition, the direction of **M** will very slowly alter towards the direction of **B** (§2.4).

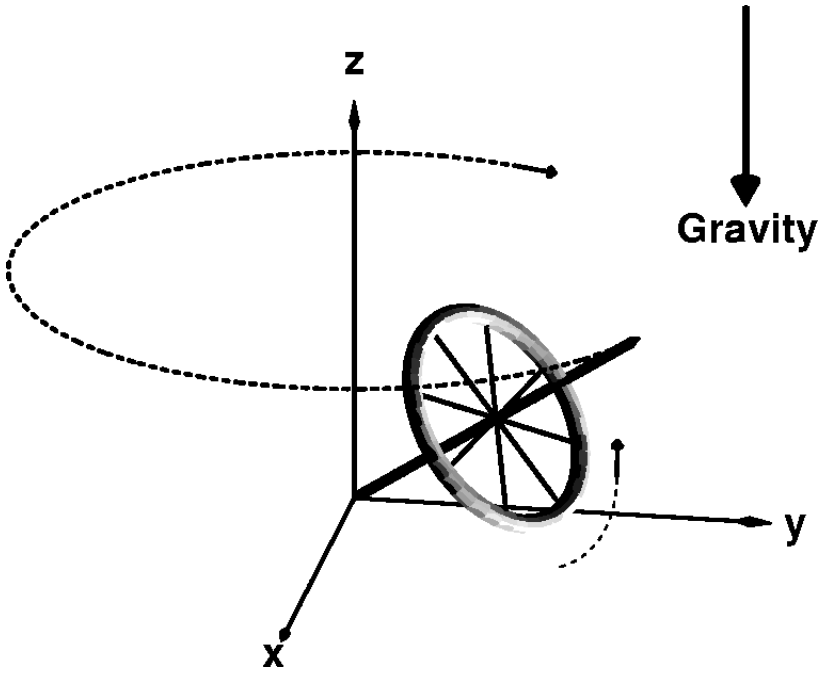

Figure 2.4. A spinning gyroscope whose axis is not aligned with the gravitational field is analogous to magnetization that is not aligned with the magnetic field. The gyroscope is pulled down by gravity, but its angular momentum causes this force to rotate the gyroscope's rotational axis about the vertical gravitational field. Friction causes the gyroscope's rotational axis to slowly alter towards the direction of the gravitational field.

Rotation of M by applied RF

When a subject is immersed in the static field, the water in his tissues becomes magnetized and **M** aligns with **B**. Precession only occurs when the direction of **M** is pushed away from the direction of **B**. This change of direction can be accomplished by adding an extra magnetic field to the main field. This new field is not static: its strength oscillates in time.

A tiny magnetic field that oscillates at the Larmor frequency and points perpendicularly to the main field **B** will have a dramatic effect on **M**, causing it to rotate away from the direction of the much larger static magnetic field at the same time it is precessing around the direction of **B**. This effect is called *resonance* or *resonant excitation*. It is analogous to the "pumping" effect on a playground swing. Gravity tends to align a swing to point downwards from its attachment point on the swingset frame. The pumping motion of the swing occupant's legs produces a sideways force much smaller than the gravitational force downward on the occupant. If the pumping force oscillates in synchrony with the natural pendulum frequency of the swing, even this small force can build up to displace the swing very far away from the natural downwards position. Similarly, the effects of a tiny time-varying magnetic field perpendicular to the large static **B** can build up over many cycles to have a large effect on **M**.

Magnetic fields that oscillate in time are always accompanied by oscillating electric fields. This combined type of field is usually called an electromagnetic field, or electromagnetic wave. The resonant frequencies typically encountered in MRI are in the same range as radio and television signals, and so the usual term for this type of electromagnetic radiation is radiofrequency radiation, abbreviated simply to *RF*.

The symbol for the strength of the time-varying magnetic field used to excite the magnetization **M** is B_1. A typical value of B_1 in MRI is 10^{-6} Tesla. The RF transmission time (T_{RF}) is usually just a few milliseconds—for this reason, the transmitted RF radiation is often called the *RF pulse*. The angle

through which **M** rotates away from **B** due to B_1 is $\gamma \cdot B_1 \cdot T_{RF}$; for example, with γ=42.54 MHz/Tesla, $B_1=10^{-6}$ Tesla, and T_{RF}=5.9 ms, this *flip angle* is 90° (¼ of a full rotation). During this 5.9 ms period, **M** also rotates through about 376,000 full rotations about the direction of **B**. The motion of **M** is really a spiraling outwards from the direction of **B**; in this example, moving 0.00024° away from **B** each time it spins through 360° around **B** (see Fig 2.5).

If the RF field B_1 does not oscillate precisely at the Larmor frequency, then its effect on **M** is weakened. The more "off resonant" that B_1 is (i.e., the farther away its frequency is from $\gamma \cdot B$), the smaller the flip angle will be. In the swingset analogy, this is like pumping one's legs at the wrong rate—the result is a smaller amplitude swinging motion.

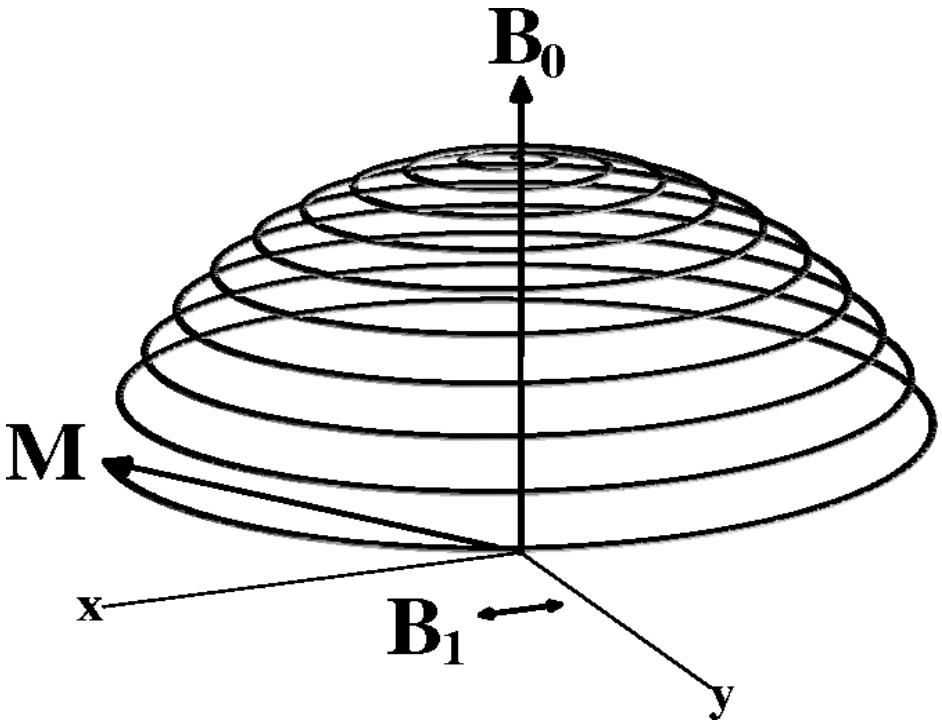

Figure 2.5. Path followed by **M** in the presence of a large field **B** and a small oscillating field B_1 which is perpendicular to **B** (in this case, B_1 is parallel to the x-axis) . **M** is initially aligned with **B**, and slowly moves away from **B** while precessing clockwise rapidly about **B**. In reality, many more precession cycles around **B** are needed to excite **M** to the flip angle shown (72°).

Emission of RF by M

The portion of the vector **M** that is perpendicular to the main static field vector **B** is also oscillating at the Larmor frequency. As an oscillating electromagnetic object, it emits RF at the Larmor frequency. This electromagnetic radiation is the fundamental NMR signal: it is a very weak radio frequency wave that is picked up by a receiver (i.e., a radio antenna) near the subject. A basic NMR experiment is thus to immerse the subject in the main field, transmit a strong RF pulse to excite **M**, turn off the transmitter, and turn on the receiver to collect the small RF signal that the subject's water protons re-transmit. The principal difficulty in NMR and MRI is that the RF radiation transmitted by **M** is very weak; for this reason, a larger B_0 is better, since it produces a larger M_0, which will produce a larger RF signal after **M** is excited away from the direction of **B**. The weak signal is also why the scanner room is heavily

shielded with metal in the walls, so that external RF signals can be kept from interfering with the reception (e.g., f=63.81 MHz at B_0=1.5 T is in the middle of the VHF television band).

The component of **M** that is perpendicular to **B** is called the *transverse magnetization* (denoted by M_{xy}); the component parallel to **B** is called the *longitudinal magnetization* (denoted by M_z)—see Fig. 2.3. The transverse magnetization is largest when the flip angle is 90º; this flip angle is often used to make the NMR signal as large as possible. The flip angle is controlled by varying the strength and/or duration of the applied RF field B_1.

It is most common to detect the NMR signal emitted by the subject with a single receiver, which is usually in the shape of a tube of wires surrounding the subject (see §2.8). All the RF signals emitted from all the water protons are implicitly added up by this arrangement. The signal is detected over a *readout interval* ranging from 5 to 100 ms, depending on the imaging method being used. The signal is converted to digital form, sampled rapidly (about 1 million times per second over the readout interval), and stored in a computer for later processing.

Frequency, Phase, and Interference

The frequency of the NMR signal emitted from location x is $f(x)=\gamma\cdot B(x)$, where $B(x)$ is the strength of the static magnetic field at x. If $B(x)$ is uniform, so that $f(x)$ is a constant, then the RF signal from the subject in the scanner will be proportional to $M_0(x)\cdot\sin(\alpha)\cdot\cos(2\pi\cdot f\cdot t)$, where α is the flip angle and t is time. Graphs of a few such cosine functions, with slightly different frequencies, are shown in Fig 2.6.

If the static magnetic field is not perfectly uniform, then the detected NMR signal will be the sum of functions oscillating at many different frequencies. Over the readout interval, this frequency dispersion will cause the detected signal to gradually decay away. At first, the signal is large, because the oscillations in the emitted RF start out lined up, so that the signals from each x add up with the same sign (positive or negative). As t gets larger, the signals with different frequencies will get different signs, and so *interfere* destructively: at any given instant, some will be positive and some will be negative, and their sum will be small. Figure 2.7 shows how the total signal strength diminishes due to interference when the frequencies $f(x)$ of the individual components are scattered randomly about the central frequency $f_0=\gamma\cdot B_0$ (see §2.4).

The *phase* of an individual cosine (or sine) wave of frequency f at any time t is defined to be $360°\cdot t\cdot f$. Phases can only be distinguished in the interval from 0° to 360°; values past 360° wrap back (e.g., 362° wraps to 2°). For example, if f=100 Hz and t=23.21793 s, then $360°\cdot t\cdot f$=835845.48°, which wraps back to 285.48°. When adding up cosine waves of different frequencies, strong destructive interference will occur at time t if the phases of the cosine waves are spread out over the whole range of 0°..360°. This is because the values of the cosines will be spread out over the range –1..1, and so when added will tend to cancel. If the phases of the cosine waves are all tightly clustered about their mean value, then destructive interference will be small, and the resulting sum of cosines will be large. In the NMR jargon, destructive interference is sometimes called *dephasing*.

Phase can be thought of as the accumulation of frequency over time. If the frequency of a signal changes, this will be reflected in the accumulated phase. For example, if f=100 Hz for 1.3712 s, and then f=50 Hz for 0.2272 s more, the phase at t= 1.5984 s is $360°\cdot(100\cdot 1.3712+50\cdot 0.2272)= 53452.80°$, which wraps back to 172.80°. If the phase of a signal is measured repeatedly and rapidly, it is possible to compute the history of its frequency changes. This is because another way to think of the relationship between phase and frequency is that frequency is the rate of change of phase with respect to time.

In NMR imaging, the phase of the received signal is measured as well as the amplitude. This allows the frequency history of the data to be computed. Combined with manipulations of the magnetic field (i.e., of the resonant frequency), the received signal can be broken down into spatially localized sources (§2.6).

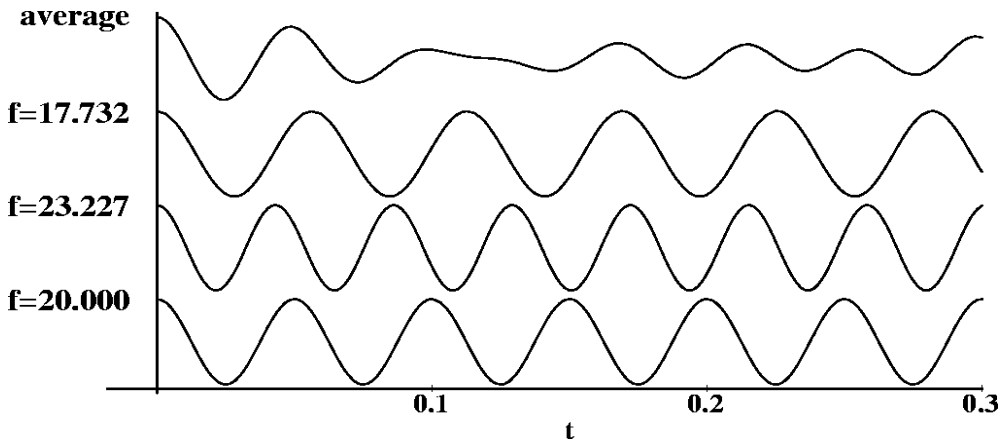

Figure 2.6. Cosine functions at 3 slightly different frequencies, and the average of these 3 functions. For small times, the cosine functions have similar phases and their average is similarly shaped. For larger times, the cosine functions have different phases, and their average is small.

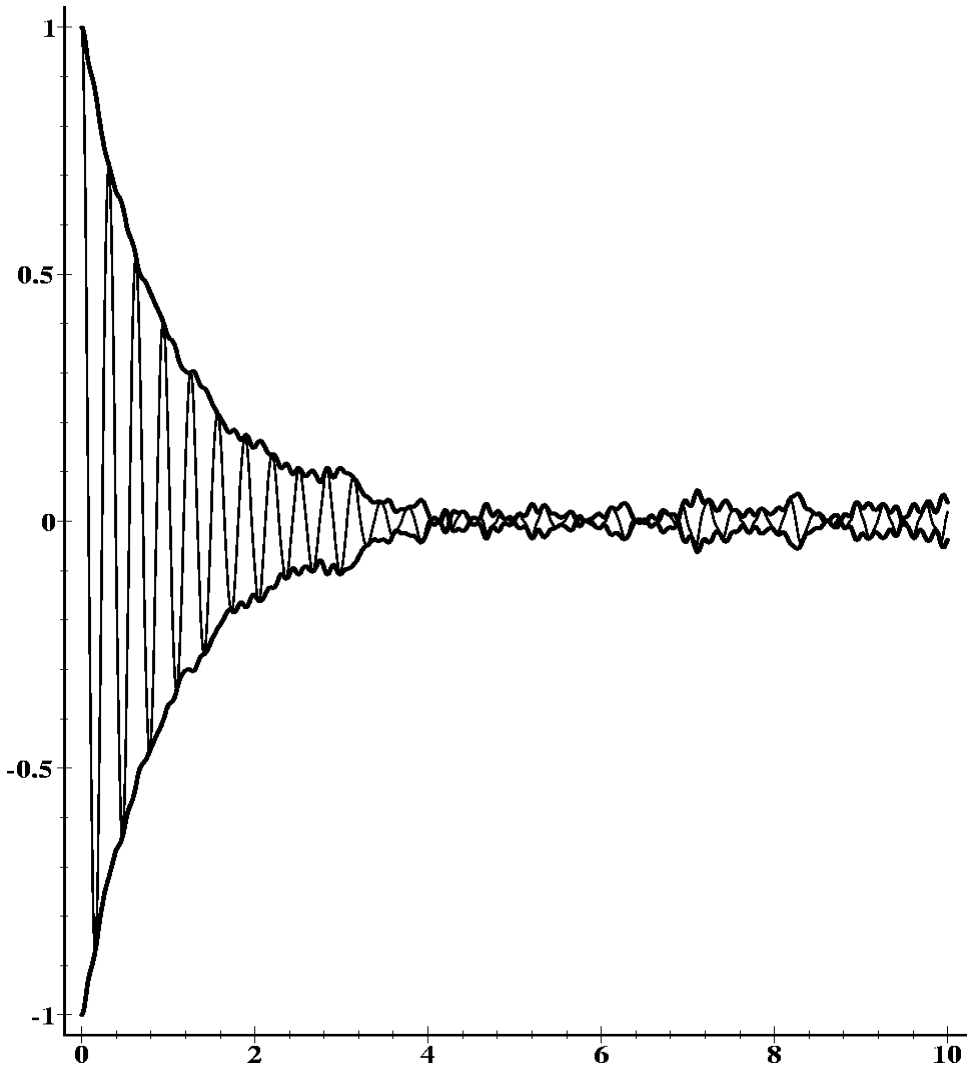

Figure 2.7. The sum of 500 cosine waves with random frequencies. The rapidly oscillating curve shows the RF oscillations at the central frequency $f = \gamma \cdot B_0$. Each of the 500 cosine waves oscillated at the central frequency plus a random offset. The enveloping dark lines show how the amplitude of the net RF falls off as time progresses, due to destructive interference from the differing frequencies. In NMR, there are many more RF emitters, and the amplitude envelope decays smoothly. The actual NMR signal decay takes millions of oscillations, not 10 or so shown here.

2.4. Relaxation of M

After the magnetization is excited by the applied RF, two things happen to make it go back into alignment with the main field **B**. The transverse magnetization M_{xy} decays away to zero as it continues to precess at the Larmor frequency, and the longitudinal magnetization M_z grows back towards its original strength M_0. These two processes that make **M** go back into alignment with **B** are called *relaxation*. Each relaxation process is characterized by the amount of time it takes to change the magnetization by a factor of $1/e \approx 0.37$ ($e \approx 2.718$ is the basis of natural logarithms). A longer relaxation time corresponds to a weaker relaxation process. The origins and definitions of the various relaxation time parameters T2*, T2, and T1 are given in the sections below; some typical values for different brain tissues are shown in Table 2.1 [*not yet created*].

Transverse Relaxation (T2*) and Tissue Structure

M_{xy} decays to zero faster than M_z is restored back to M_0. The main reason for this rapid decay is that the static magnetic field is not uniform in space. Even a perfectly uniform main field **B** will be rendered nonuniform when a nonuniform object (like a person) is put into the scanner bore. This is due to *susceptibility*: the generation of extra magnetic fields in materials that are immersed in an external magnetic field. Most tissue is *diamagnetic*: it produces a field that very slightly opposes **B**. Some tissue is *paramagnetic*: it produces a field that very slightly reinforces **B**. At the microscopic level, the magnetic field is a random looking jumble, with fluctuations in magnitude of about $\pm 10^{-7} \cdot B_0$ occurring over distances of microns. These fluctuations in **B** are caused by the intricate material structure and composition of tissue.

These small spatial variations in **B** cause neighboring protons to precess at slightly different frequencies. As a result, the transverse magnetization M_{xy}, when added up over any region larger than a cell, will be the sum of many components at different frequencies. As time passes, these components will start to interfere destructively, and the net M_{xy} will be reduced. At B_0=1.5 Tesla, fluctuations of $10^{-7} \cdot B_0$ have a frequency of about 6.4 Hz, which is a full cycle (360º) in about 160 ms. After about half this time, M_{xy} will have decayed to $1/e \approx 37\%$ of its original value. Since M_{xy} is the source of the emitted RF, the NMR signal will also decay at this rate.

The time over which the NMR signal decays by a factor of e (≈2.718) due to the spatial inhomogeneities in **B** is called *T2** in the imaging literature. T2* at a given location depends on the type of tissue present there, since each type of tissue will have a distinct microscopic structure and composition, producing a distinct level of microscopic magnetic field perturbations.

Transverse Relaxation (T2) and the Hahn Spin Echo

One of the cleverest concepts in NMR is the *spin echo*, invented by Erwin Hahn (Hahn 1950). The idea is to apply *two* separate RF excitation pulses before doing the signal readout. The first RF pulse is designed to produce a flip angle of 90º, converting M_z into transverse magnetization, which will immediately start to undergo T2* decay. After a few milliseconds, the second RF pulse is applied, this one designed to produce a flip angle of 180º (this is called an *inversion pulse*, or sometimes a *refocusing pulse*). The time between the two RF pulses is denoted by the symbol *TI* ("I" for inversion). The readout interval occurs after the second RF pulse.

Figure 2.8 shows sequence of events and their effect on the magnetization. Between the two RF pulses, spatial nonuniformities in **B** mean that the magnetization in some regions precess faster than average and in other regions slower than average. The result is T2* relaxation of M_{xy}. The inversion pulse does not change this—protons that happen to be in a region with large **B** will still be in the same place after the second RF is applied, so they will still precess faster than average. However, the effect of the 180º flip is to put the faster protons *behind* the slower protons by exactly the amount they were ahead just before the inversion. It now takes them the same amount of time to catch up to the slower protons as it took them get ahead: TI. At time $2 \cdot TI$ after the 90º RF pulse, all the protons are back into phase, and M_{xy} has been restored in strength.

This effect is analogous to an acoustic echo. Shouting in a canyon causes the sound to disperse away. The canyon walls cause the sound waves to be reflected back; when the sound has made the round trip from shouter to canyon walls and back to shouter, the sound waves are back together and an echo is heard. In the spin echo, the protons don't physically travel apart, but their phases travel apart due to differences in **B**. The 180º pulse acts like the canyon walls, and starts the reconvergence part of the echo. In the acoustic echo, when the sound reaches the reflector, half of the time needed for the echo formation has passed. With the spin echo, when the 180º pulse is applied, it is halfway to the echo time—the time when the emitted RF signal is largest. The echo time is denoted by the symbol *TE*,

which is measured from the center of the initial 90º RF pulse. After TE has passed, the signal decays away again, since the faster protons now pass the slower protons and get out of phase again.

T2* signal decay is conceptually separated into two further components. These components are defined by the spatial extent of the nonuniformities in **B** that cause them. The decay caused by inhomogeneities below about 10 μm in extent is called *T2* relaxation (without the asterisk). The decay caused by larger scale fluctuations in **B** is called *T2′* relaxation. Since each component of T2* relaxation by itself is weaker than their sum, the relaxation times T2 and T2' are both longer than T2*. (The relationship between these T2 times is 1/T2*=1/T2+1/T2'). In typical brain tissue, T2 is about twice T2*, reflecting the fact that the smaller scale nonuniformities in **B** account for about half of the signal decay.

The reason for this conceptual separation is *diffusion*: the random movement of water molecules through tissue due to thermal agitation. Water molecules diffuse about 10 μm during the typical readout interval of MRI. The spin echo is predicated on the assumption that the protons stay in one place between the initial RF pulse and the time TE, so that faster precessing protons remain faster and slower precessing protons remain slower. To the extent that this assumption is violated by physical displacement of the water molecules, the spin echo will not be perfect. Magnetic field inhomogeneities that are smaller in scale than the diffusion distance will cause irreversible loss of transverse magnetization, since the motion of the water molecules means that those hydrogen nuclei which are in regions of larger **B** before the 180º pulse may move into regions of smaller **B** after the 180º pulse. This change in the **B** experienced by diffusing protons means that the spin echo will not form perfectly. The resulting loss of signal is irreversible, since diffusion is a random process. At the spin echo time, the transverse magnetization will be restored to the value $M_{xy}=M_0 \cdot e^{-TE/T2}$ instead of all the way to $M_{xy}=M_0$.

Spatial nonuniformities in **B** larger than 10 μm will not be seen by the diffusing water molecules, and so the effects of these inhomogeneities can be refocused by the 180º pulse. The diffusion distance of 10 μm is larger than red blood cells (5 μm) and capillaries (6 μm), but smaller than most other blood vessels. This means that spin echoes are sensitive mostly to the microcirculatory structure, while NMR data taken without the refocusing pulse will be sensitive to all scales of vessels.

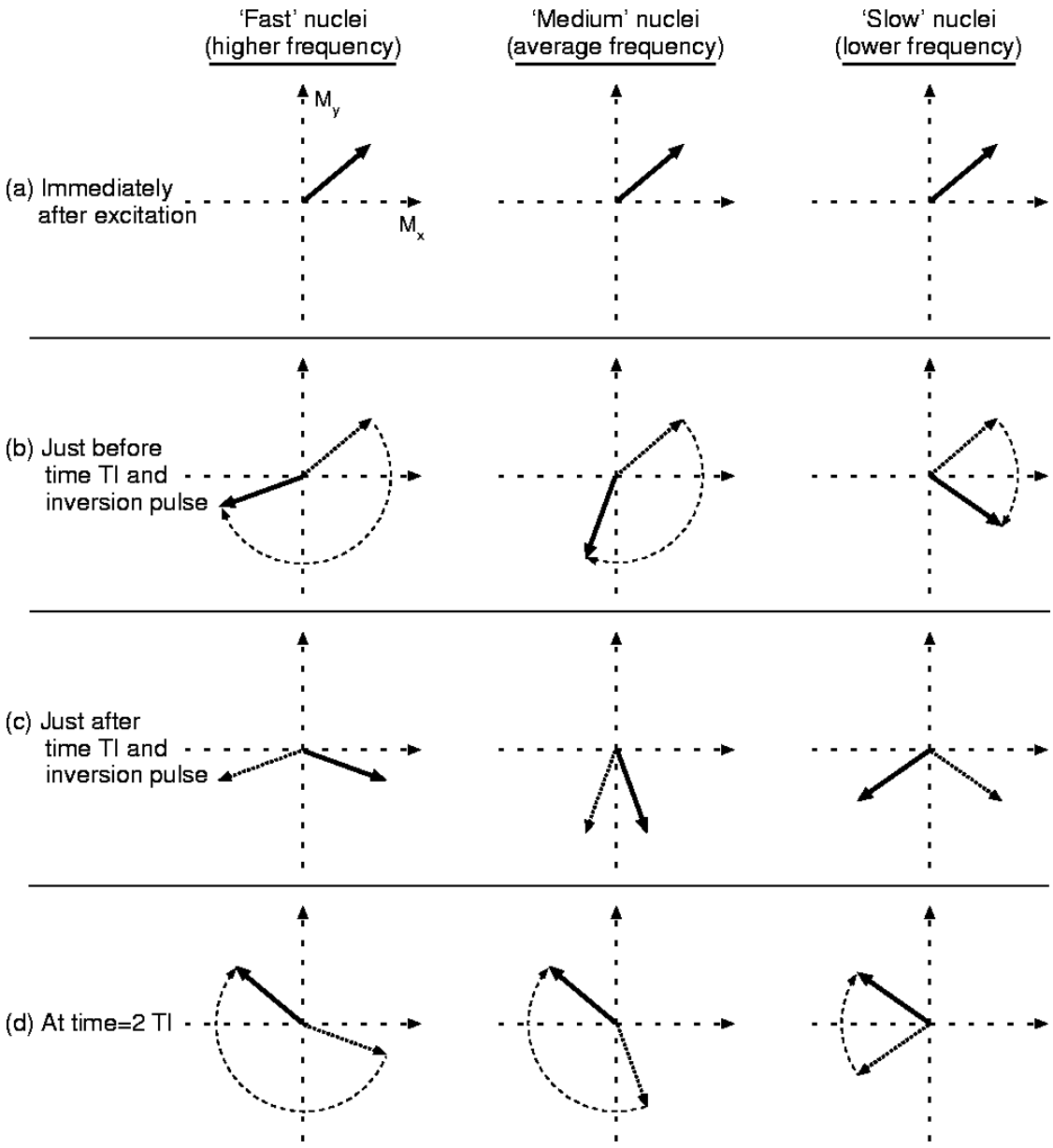

Figure 2.8. Sequence of events in a Hahn spin echo. The graphs show M_{xy} for various regions at selected times. The left column is a region with larger **B**, and so the nuclei there precess faster; the middle column is a region with average **B**; the right column is a region with smaller **B**. (a) Immediately after the initial RF excitation, the magnetization vectors are all aligned, in this case between the M_x- and M_y-axes. At this moment , the emitted RF signal would be large. (b) After some time has passed, the magnetization vectors point in different direction since they are precessing at different rates. The emitted RF signal would be small. (c) An RF pulse flips all the magnetization vectors through 180º, in this case about the M_y-axis. Each vector **M** rotates out of the M_x-M_y plane—through 3D (M_x,M_y,M_z) space—during this RF pulse, but ends up lying back in the M_x-M_y plane, reflected through the M_y-axis from its pre-pulse orientation. (d) When the same amount of time passes again, the magnetization vectors all come back into phase. The emitted RF signal would be large again—this is the "echo".

Longitudinal Relaxation (T1)

The growth of M_z back towards its equilibrium value M_0 is called longitudinal relaxation. An RF pulse at the resonant frequency causes **M** to rotate away from the **B**-direction, which reduces M_z. In principal,

an opposite RF pulse could restore **M** back to its equilibrium value, or at least move it towards being aligned with **B**.

Even without externally applied RF radiation, there are still electromagnetic oscillations inside water that affect the protons. Water molecules are continuously moving and rotating randomly, colliding with other molecules. Some of these molecules will have their own magnetic fields (due to unpaired electrons, rather than to nuclei). From the viewpoint of a water proton moving nearer and farther from such molecules, the magnetic field is changing very slightly in time in a random way. Some of these changes occur near the resonant frequency, and so will affect the nuclei, just as does an externally applied RF pulse at the resonant frequency. On average, their net effect is to drive **M** back towards its preferred value, which is M_0 and aligned with **B**.

The time scale over which M_z grows back to M_0 is called *T1*. Immediately after a 90º excitation pulse, $M_z=0$; at time t after the pulse, $M_z=M_0\cdot(1-e^{-t/T1})$. The T1 relaxation rate depends on how many impurity molecules are present to provide magnetic field fluctuations, which is why T1 varies strongly between different tissue types.

The magnetic field fluctuations at the molecular level that cause longitudinal (T1) relaxation also contribute to transverse (T2 and T2*) relaxation. Transverse relaxation is also caused by larger spatial scale (microns to millimeters) variations in **B** that have little effect on M_z. As a result, transverse relaxation rates are always larger than longitudinal relaxation rates, meaning that $T2^* < T2 < T1$.

2.5. Formation of Images

All the RF signals emitted by all the excited protons sum up together in the RF signal that is detected. At any given instant in time, only two numbers can be read out of the detector system into the scanner's computer: the amplitude of the signal, proportional to $(M_x^2+M_y^2)^{1/2}$, and its net phase, $\tan^{-1}(M_y/M_x)$. Here, M_x and M_y are the x- and y-components of the transverse magnetization M_{xy}, summed up over the entire region that was excited by the transmitted RF pulse.

Images are formed by reading the RF signal out over an extended interval. During this time, the magnetic field **B** is manipulated to alter the signal transmitted by the tissue so as to allow different locations to be distinguished (Lauterbur 1973).

Gradients in B: Frequency Depends on Spatial Location

On top of the unchanging large static and spatially uniform field $\mathbf{B_0}$, the scanner hardware is capable of adding smaller magnetic fields whose strength varies linearly in space. These fields are known as *gradient fields*, and are denoted by the symbols G_x, G_y, and G_z. The overall strength of the scanner's magnetic field at spatial location (x,y,z) and time t is $B_0+x\cdot G_x(t)+y\cdot G_y(t)+z\cdot G_z(t)$. Unlike the main field strength, the sizes of the gradient fields are under the control of the scanner's computer and can be programmed to fluctuate in time, and can be either positive or negative. A typical maximum strength of any gradient field is 10 milliTesla/meter (=1 Gauss/cm). This translates into an NMR signal frequency gradient of 425.4 Hz/mm. That is, the resonant frequency of two protons 1 cm apart can be made to differ by 4254 Hz. The left-to-right width of the human brain is typically 140 mm; at 1 G/cm, this corresponds to a frequency separation of 59.6 KHz. This is much larger than the susceptibility induced frequency shifts, which are on the order of 10 Hz.

Slice Selective Excitation Using Gradients

The first step in the imaging process is the creation of transverse magnetization M_{xy} by transmitting RF. If the magnetic field is uniform in space, then the resonant frequency of all the protons is the same, and all the protons will be equally affected by the RF pulse. This procedure is called *volume excitation*.

Although volume excitation has some uses, in MRI it is more common to excite only a thin slice of the tissue in the magnetic field. This is done by applying a gradient field while the RF transmitter

is on. Only those portions of the tissue that are resonant with the frequencies contained in the RF pulse will be excited. *Slice selective excitation* is illustrated in Fig. 2.9. A range of frequencies is transmitted; that is, the RF pulse is synthesized from a sum of sine and cosine waves. The size of this range of frequencies is called the *bandwidth* of the RF pulse. A typical bandwidth for slice selective excitation is 1000 Hz. The magnetic field gradient determines the thickness of the chosen slice; if the maximum gradient in frequency is 425.4 Hz/mm, then the thinnest slice a 1000 Hz RF pulse can excite is 1000/425.4=2.35 mm. Smaller gradients can be used to select thicker slices. The location of the slice is controlled by setting the central frequency of the transmitted RF radiation.

Slices can be selected in any plane, since the G_x, G_y, and G_z fields are independently controlled. Slices can even be selected in oblique planes, since combining more than one gradient field produces a frequency gradient in a slanted direction.

By doing slice selective excitation 3 times in a row, first with G_x, second with G_y, third with G_z, the excited volume will become the intersection of three slices, approximately a single cube. This technique could be used to acquire the NMR signal from a single volume element (*voxel*) in the brain. The main drawback to using this technique for image acquisition is the amount of time it would take to acquire data covering the entire brain—approximately 1 hour per acquisition. Single voxel data acquisition *is* used in NMR spectroscopy, but this has had little application to fMRI, and so is not discussed here.

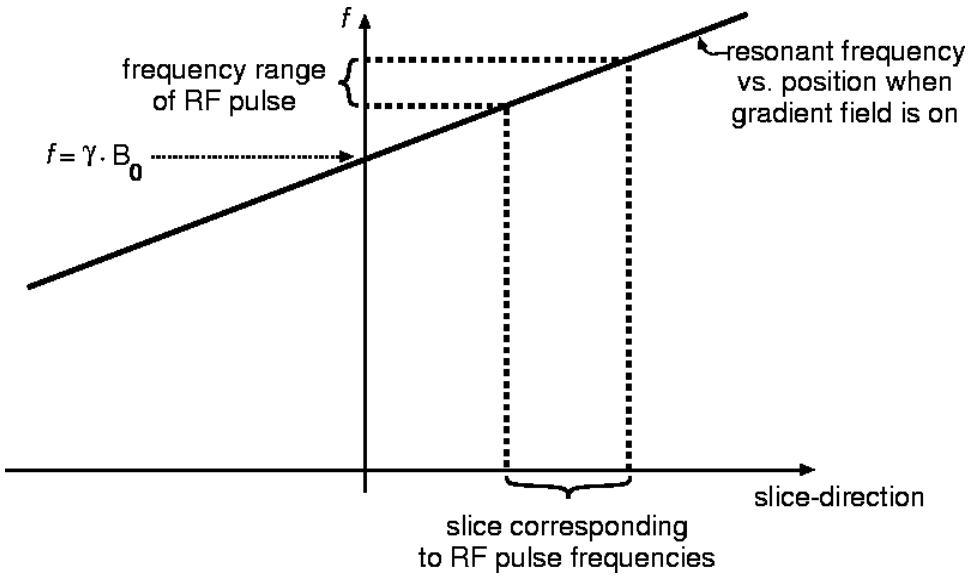

Figure 2.9. Slice selective excitation is accomplished by transmitting the excitation RF pulse at the same time a magnetic field gradient is turned on. The slice location depends on the central frequency of the RF pulse and the strength of the gradient field. The slice thickness depends on the mix of frequencies used in the RF pulse and the strength of the gradient field. The slice is oriented perpendicularly to the direction of the gradient field; for example, application of G_z produces transverse magnetization in a slice that is parallel to the xy-plane.

Distinguishing Signals from Different Regions Using Gradients

Magnetic field gradients are used to select which tissue slice is affected by the transmitted RF pulse. Gradients are also used during the signal readout step of the imaging process. The frequency of the RF signal emitted from each region depends on the magnetic field at that location. Imagine that a gradient field is applied from left-to-right, so that the left side of the subject is at a lower resonant frequency and the right at a higher resonant frequency. Although the RF signals from the entire excited slice are mixed

together in the receiver, signals from the subject's left can be distinguished from signals originating in the right by their different frequencies. This is referred to as *frequency encoding*; in this example, the left-to-right direction is called the frequency encoding direction.

While the frequency encoding gradient is on, the NMR signal is measured rapidly and repeatedly for a few milliseconds (usually about once per microsecond). This train of measurements is then resolved into its separate frequency components on a computer. Each frequency component corresponds to a different position along the frequency encoding direction. The spatial resolution that can be achieved depends on the strength of the gradient field and on how long the data acquisition lasts. A stronger gradient means that nearby positions are more well separated in frequency, and so can be resolved more easily. A longer acquisition interval means that closer frequencies can be distinguished in the data processing, since their phases will have had more time to drift apart. Since the gradient field means that frequency corresponds directly to position, measuring frequencies more accurately implies higher spatial resolution. The spatial resolution is $1/(\gamma \cdot G_{freq} \cdot T_{readout})$. For a 5 ms readout with a 1 G/cm gradient, this resolution is 0.47 mm.

The *field-of-view* (FOV) is the physical dimension of the whole image; it is equal to the voxel dimension (resolution) times the number of voxels along an image edge. The FOV is usually matched to the dimensions of the region being imaged; for the head, FOV=200-240 mm is a common range. There is no value in acquiring a larger FOV, since the voxels outside the head have no protons and hence produce no data. Acquiring a smaller FOV does not produce an image of just one portion of the head (like a cropped photograph). Since the NMR signal originates from the entire slice, after image reconstruction, the data from each voxel in the slice must end up somewhere. If the FOV is smaller than the object, the image wraps around from one edge to the opposite edge. For example, it is possible to end up with an image that has the subject's nose cut off the front of the image and appearing inside the back of his head. While amusing, this is not useful.

Slice selective excitation provides one dimension of the imaging, by restricting the source of the emitted signal to a thin slab through the subject. Frequency encoding provides the second dimension. A third dimension is needed to enable the NMR signals to be isolated down to roughly cubical individual voxels. This third dimension is provided by applying a gradient field in the third direction, perpendicular to both the slice selection direction and the frequency encoding direction. The third direction is called the *phase encoding* direction.

The purpose of the third gradient is to make the phase of the NMR signal depend on position along the third direction. The procedure is to do data readout repeatedly, with the frequency encoding gradient turned on the same way during each readout interval. If nothing were changed between readouts, this procedure would just result in duplicate data. Applying the phase encoding gradient briefly before each readout will change the phase of the NMR signal, since phase is frequency×duration and the phase encoding gradient field changes the NMR frequency for as long as the gradient is turned on. As a result, each readout will be different than all the others, depending on the amount of extra phase that is applied with the phase encoding gradient. With enough different readouts, the image reconstruction software can distinguish signals in the phase encoding direction. Each readout is separately resolved into frequency components (=position along the frequency encoding direction). At each given frequency, the different values obtained from the many readouts are further resolved into sub-components that fluctuate at different rates depending on the amount of phase encoding gradient applied. For example, the mean value from all the readouts does not depend on the phase encoding gradient at all. This mean value corresponds to the location where the phase encoding gradient field is zero. Those sub-components that fluctuate slowly relative to the amount of phase encoding come from positions where the phase encoding gradient is weak; that is, from where the total magnetic field does

not differ much from B_0 no matter if the phase encoding gradient is on or off. Those sub-components that fluctuate most rapidly relative to the amount of phase encoding come from positions where the phase encoding gradient is strongest, which are at the extreme edges of the slice being imaged.

The resolution of the readout data into frequency components, and the further resolution of each frequency component into phase sub-components, are the same mathematical operation—the *Fourier transform*. In each case, a set of data is transformed into a series of numbers that represents the magnitude of fluctuations in the data at each different frequency. In the frequency encoding direction, the independent variable along which the data fluctuates is elapsed time since the start of the readout. In the phase encoding direction, the independent variable is the total amount of phase encoding applied ($G_{phase} \cdot T_{phase}$, where G_{phase}=strength of phase encoding gradient field and T_{phase}=amount of time the phase encoding gradient was on between the last RF pulse and the beginning of the data readout).

MRI as "Fourier-Space" Scanning

In normal optical imaging (e.g., photography), the data for all the image elements are gathered simultaneously. In MRI, only one piece of information is acquired at each instant in time, since there is only one RF receiver. A somewhat similar situation in optical imaging would be if the illumination was provided by a narrow laser beam which scanned over the scene being photographed. At each instant, only a small part of the data needed for the whole image would be available. Only by systematically sweeping the laser over the scene and waiting for each element in the camera field-of-view to be illuminated at least once would the entire image be acquired.

In MRI, the systematic sweeping of the laser is replaced by systematic sweeping of the relationship between frequency and location, using the gradient fields. Unlike the optical analogy above, the scanning does not take place in physical coordinates—what MRI physicists call "real-space" or "image-space"—but in a coordinate system of NMR signal phase—what MRI physicists call "Fourier-space" or "k-space" ("k" is the symbol used for phase coordinates). At each instant, the datum that is acquired contains information from the entire slice that was excited. The sweeping of the imaging gradients means that each datum depends on the slice magnetization in a different fashion. When enough different types of data are gathered, the spatial distribution of magnetization can be reconstructed by transforming from Fourier-space to real-space.

The number computed for each voxel is proportional to the received RF signal averaged over the voxel volume. Since a voxel will most likely contain a mixture of different tissue types, this averaging can make it difficult to see effects that originate from a single tissue type—this is called the *partial volume effect*. For example, almost all cortical gray matter voxels will contain a significant amount of CSF unless the resolution is very fine ($\leq$ 1 mm). Comparing the functional MRI signals (Chapter 3) from different gray matter regions is then problematic, since it will be unknown what the proportion of CSF is in the different voxels. This is a general problem with using magnetic resonance images for quantitative purposes—it is hard to make comparisons between voxels even in the same subject, since each voxel will have an unknown mix of tissue types. One way around this is to attempt to estimate the amount of each major tissue type in each voxel. This can be done by acquiring several different images with different contrasts.

The Gradient Echo

In the spin echo (§2.4), dephasing of the NMR signal from regions with different resonant frequencies is corrected for by using a 180° inversion pulse. When the different frequencies are due to externally applied gradient fields (vs. the spatially nonuniform tissue-generated fields that cause T2* decay), a different method can be used to correct for the signal dephasing. This new method is simply to reverse the gradient field. Nuclei that were at frequencies larger than $\gamma \cdot B_0$ will now be at frequencies smaller than $\gamma \cdot B_0$, and vice-versa. The effect is like having a foot race with some fast runners and some slow

runners. After a certain amount of time, the runners will be spread far apart. Then suppose that the fast runners are suddenly turned into slow runners and the slow runners are suddenly turned into fast runners. When enough more time passes, the runners will be back together.

When a gradient field is applied, the nuclei in locations where $f > \gamma \cdot B_0$ are like the fast runners: their phase changes more quickly than the spins where $f < \gamma \cdot B_0$. The reversal of the gradient field will change the high frequency nuclei into low frequency nuclei and will change the low frequency nuclei into high frequency nuclei. After enough time passes, the individual nuclear **M**'s will have the same phase. When their phases are widely different, the total NMR signal will be small. When their phases are nearly the same, the total NMR signal will be large. By applying a gradient field for a time, and then by reversing it, the result is to force the signal to decay, and then to restore its magnitude (T2* decay will still occur during this process). This effect is called the *gradient echo*, in analogy with the older concept of the spin echo.

A Word About Fat

The emphasis has been on the hydrogen nuclei in water molecules. Hydrogen is present in many other larger biological molecules, but these protons do not contribute strongly to the observable magnetization. The reason for this is that the interior of these molecules have very irregular magnetic fields, which means that the signals from these nuclei have a very short T2—typically about 10 μs, so that these signals decay before readout begins. The principal exception to this rule is the protons present in the methyl groups of fat molecules. These groups stick out into the "sea" of water molecules, where on average **B** fluctuates less, and so these fat protons actually contribute to the NMR signal. However, due to the *chemical shift*—local magnetic field changes resulting from electron rearrangements in the methyl groups in the presence of the large external **B**—the fat signal is at a slightly different frequency than the water frequency (at B_0=1.5 T, the fat signal is shifted by about 200 Hz). Since images are formed by manipulating the precession frequency, the fat component of a voxel will appear in a different location than the water component This undesirable effect has led to the invention of *fat suppression* techniques, in which a separate RF pulse is used to excite only the fat protons at their distinct frequency band, and then their resulting transverse magnetization is *crushed* using a large gradient pulse. This all takes place prior to the water excitation, and is usually omitted from the discussion of various imaging methods.

2.6. Image Contrast

Magnetic resonance images are measurements of the NMR signal received from each voxel. The magnitude of the NMR signal depends on many factors, including water content, magnetic molecular impurities (affecting T1), and the microscopic structure of the tissue (affecting T2 and T2*). The way an image is acquired affects which of these factors are most important.

Usually the absolute level of the NMR signal is unimportant. What matters most is the ability to detect changes in the signal (i.e., in the values stored in each image voxel) between two regions that have different tissue properties. For any given image acquisition method, the property that most strongly affects the changes in image intensity is called the *weighting* or *contrast* of that method.

One of the strengths of MRI is the variety of image acquisition methods that have been devised to produce different image contrasts. Imaging techniques generally differ in the type of RF pulses applied, in the order and amount of gradients applied before and during data readout, and in the timing of the data readout interval.

Perhaps the most obvious type of image contrast is the local water content. Since water hydrogen protons are the source of the signal, a voxel that has more water (vs. other tissue components) will have a larger signal. However, this turns out to be relatively useless for biomedical purposes, since most

tissue contains about the same quantity of free water. As a result, although the NMR signal is large, it shows little variation across the image, and so such *proton density* weighted images are rarely acquired. (When imaging other nuclei, their density is usually the desired quantity.) This section discusses imaging methods that provide contrast based on the three relaxation mechanisms discussed in §2.4.

Imaging for T2* Contrast

T2* signal decay starts to occur immediately after the RF pulse creates transverse magnetization. This decay will happen more quickly in voxels that are have more heterogeneous magnetic fields; that is, T2* will be smaller in such voxels, and will be larger in homogeneous voxel (e.g., in CSF). The imaging tool for acquiring T2*-weighted data is simple: *time*. That is, the data readout does not occur immediately after the RF pulse, but is delayed for a while in order to let the transverse magnetization M_{xy} decay at its own rate in each voxel.

If the readout occurred quickly just after the RF pulse, there would be little time for T2* signal decay to happen, and the resulting image would be mostly proton density weighted. If the readout were delayed too long, the NMR signal would all decay away, and there would be nothing to measure (the image would be black). To get the maximum image contrast between voxels with varying values of T2*, it is best to time the data readout so that it occurs at time T2* after the RF pulse. Here, an average value of T2* is chosen to define the readout time interval. At 1.5 Tesla and in the brain, this would be about 40 ms. In T2*-weighted images, tissue with large T2* (slow signal decay) will show up more brightly than tissue with small T2* (quick signal decay).

The gradient echo technique (§2.5) is used to measure a T2*-weighted image. After the RF pulse, the frequency encoding gradient is applied with a negative value for an interval, and then the frequency encoding gradient is reversed (turned positive) and the data readout begins. The method is timed so that the actual gradient echo (rephasing of the magnetization) occurs at the center of the readout interval, which is chosen to be the average T2* of the tissue being scanned. The time at which the gradient echo occurs is called *TE*.

The sequence of events is illustrated in Fig. 2.10. There are five scanner components that are involved: the RF transmitter, the three gradient field generators (one for each spatial coordinate axis), and the RF receiver. Correspondingly, five graphs vs. time are shown, indicating the sequence of operations of each scanner component. This type of display is called a *pulse sequence* diagram. MRI physicists are fond of displaying and publishing such diagrams, since they convey a great deal of technical information in a small space.

The negative frequency encoding gradient dephases the NMR signals, and the positive frequency encoding gradient rephases them, producing the gradient echo. It would be possible to acquire the image data without the negative frequency encoding—this procedure is called a "½ k-space" technique. For the best T2* contrast, it is still necessary to put the readout signal maximum at the average T2* of the tissue. Without the negative frequency encoding, the signal maximum occurs at the start of the readout interval, so it would be necessary to delay the data readout. In the imaging technique shown in Fig. 2.10, there is no advantage to ½ k-space imaging, since the total time required for the acquisition does not change much. The advantage of the gradient echo is that it acquires both the dephasing and rephasing parts of the data. These are in principle nearly identical, but in practice, measuring the same information twice improves the image quality by reducing the noise level. Since there is no significant time penalty for this strategy, it is the one that is usually used.

The pulse sequence illustrated in Fig. 2.10 does not acquire all the data needed for one image. With each RF pulse, one value of the phase encoding gradient is applied, and then the frequency encoding gradient is combined with the data readout. A large number of different phase encoding steps are needed to reconstruct an image (§2.5)—one for each voxel in the phase encoding direction. The

sequence of Fig. 2.10 must be repeated at least this many times to acquire the raw data needed for a single image. At each repetition, everything is identical except for the phase encoding step, which is varied in strength. (In the MRI jargon, each repetition of the basic sequence is called a *shot*. The use of several shots to acquire image data makes the technique described here a *multi-shot* imaging method. *Single-shot* imaging methods will be outlined in §2.7.)

The time between the RF pulses that excite the same tissue slice is called *TR*. When imaging methods are specified, the generic class of the pulse sequence is given (e.g., gradient echo), and the specific parameters of how it was used (e.g., TE, TR, flip angle). A list of the parameters that are used to specify completely an imaging method is given in Appendix A.

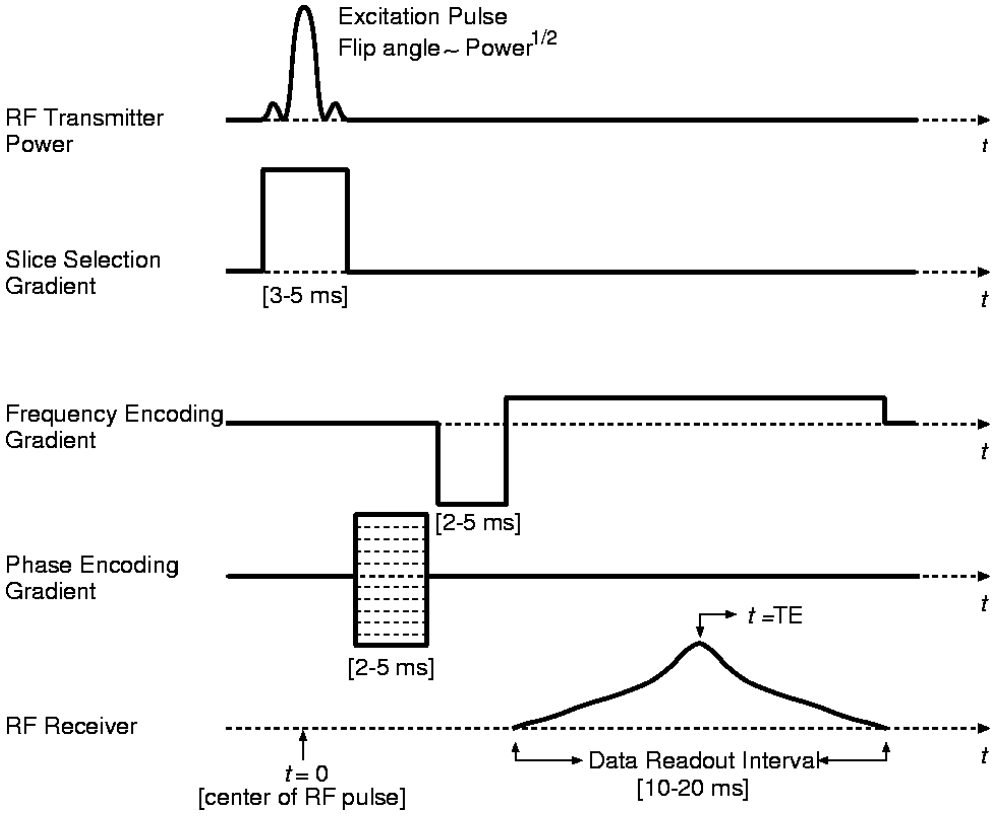

Figure 2.10. The sequence of activity in a gradient echo (GE) imaging sequence. Each line shows the utilization of one of the components of the scanner. Typical values for the duration of each action are shown in brackets. The phase encoding action is shown as taking many different values, because the illustrated pulse sequence must be repeated many times with different phase encodings. The data magnitude is small at the beginning of readout because the NMR signals have been dephased by application of the negative frequency encoding gradient. As the positive frequency encoding gradient brings the NMR signals across the slice back into phase, the data increases until time t=TE. At that time, the signal in each voxel has decayed from its original strength by the factor exp[−TE/T2*].

Imaging for T2 Contrast

Images with T2* contrast are very useful for functional MRI of the brain, as will be seen in Chapter 3. For clinical purposes, images with T2 contrast are more useful. T2 differs considerably between normal brain tissue and tumor tissue or infarcted tissue.

T2-weighted images are acquired using spin echoes (§2.4) rather than gradient echoes. The sequence of events is shown in Fig. 2.11. The principal difference between Figs. 2.10 and 2.11 is the application of the 180° inversion pulse before the frequency encoding gradient is used. To achieve the

best contrast between tissues with different values of T2, it is best to set TE to the average value of T2 that is expected. In T2-weighted images, tissue with large T2 (slow signal decay) will show up more brightly than tissue with small T2 (quick signal decay). This means that CSF shows up most brightly, and such images often look like photographic negatives of what one expects to see.

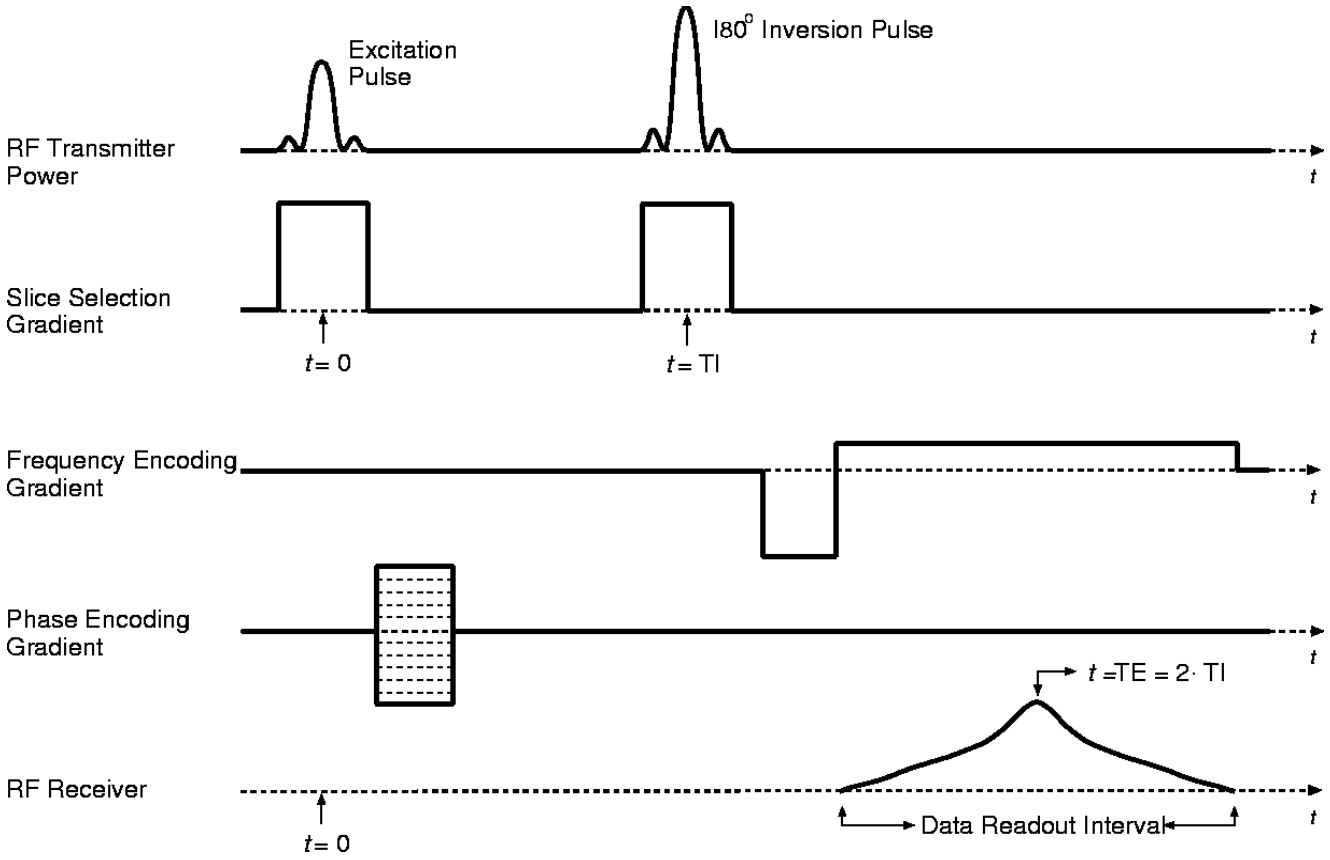

Figure 2.11. The sequence of events in a spin echo (SE) imaging sequence. The principal difference with the gradient echo sequence of Fig. 2.10 is the 180º RF pulse between the initial RF excitation and the data readout. In a "pure" spin echo, the center of the data readout interval is designed to coincide with the spin echo time 2·TI; that is, the gradient echo (resulting from the negative then positive frequency encoding gradients) and the spin echo (resulting from the 180º RF pulse) occur together. In an "asymmetric" spin echo (ASE), these two times do not coincide.

Imaging for T1 Contrast

The relaxation parameter T1 is a measure of the time that it takes for the longitudinal magnetization M_z (parallel to the main field **B**) to be restored after it has been perturbed by RF pulses. Since M_z is not directly measurable, the effects of T1 must be made to appear in M_{xy} in order to produce T1-weighted images. Since M_{xy} is "manufactured" from M_z by the RF excitation, this goal is easily achieved.

The tool is again time; in this case, TR—the time between RF excitations. It takes more than one RF excitation to get the all the data needed for image reconstruction. This interval determines how much the longitudinal magnetization in each voxel recovers before the next RF pulse. If TR is very long ($\geq 3 \cdot T1$), then M_z in each voxel will be restored all the way back to M_0. Images acquired with long TR thus do not depend on T1, but will tend to be proton density weighted. If TR is short, then the voxels with small T1 (rapid recovery) will have M_z grow back toward M_0 much more than voxels with long T1 (slow recovery).

Another factor that affects the T1 contrast is the flip angle used during the RF excitation. If the flip angle is small, much of the longitudinal magnetization will not be converted into transverse magnetization. In this case, M_z immediately after the RF pulse remains relatively close to M_0. If the flip angle is large (near 90°), then M_z immediately after the RF pulse is small. This latter case will tend to accentuate differences in longitudinal magnetization recovery rates, since even small changes from a small basis can appear large in the final image.

To avoid having the images also have T2 or T2* contrast, it is necessary to use a relatively small TE, so that transverse relaxation of the NMR signal has little time to take effect. Thus, the prescription for T1-weighted images is a short TR, a large flip angle, and a short TE. In such images, tissue with shorter T1 will show up as brighter than tissue with longer T1. In the brain, the T1 of white matter is shorter than that of gray matter, which is shorter that the T1 of CSF (see Table 2.1). As a result, T1-weighted images of the brain look something like one expects to see: the white matter is brightest, the gray matter is grayish, and the CSF is dark.

In T2- and T2*-weighted images, it is desirable to have a relatively long TR, so that the effects of T1 variations will be small. This puts a limit on the speed of T2- and T2*-weighted imaging. The actual time needed for a single shot is 20-100 ms, depending on the scanner hardware. An image with 256 different phase encoding steps would then take a 5.12-25.6 s to acquire. But with such a short TR, the image would be strongly T1-weighted. This effect can be reduced by lowering the flip angle, but this also lowers the total NMR signal and so makes the image noisier. Stretching TR out to 1 s would reduce the effect of T1, but then it would take 256 s (4.3 min) to acquire a single image. One way to overcome this difficulty is described in §2.7.

Contrast Agents

In many clinical applications, the difference between the intrinsic magnetic properties of two different tissue types is too small to show up clearly on an image. This is a big problem when the diagnostic goal is to find small tumors or other small regions of diseased tissue. A small region that has a small change in signal intensity relative to the rest of the image is hard to see and is hard to interpret.

For this reason, it has become common to use external *contrast agents* in clinical MRI. An external contrast agent is a pharmaceutical that will change T2, T2*, and/or T1 in the tissue where it is present. Such drugs are designed to have as little chemical reactivity as possible; their purpose is not to interfere with the biochemical functioning of tissue, but to interfere with the magnetic properties of tissue. At this writing, the most commonly used such agent in the USA is gadolinium diethylene-triaminepentaacetic acid (Gd-DTPA, trade name Magnevist), which decreases T1 and T2. The gadolinium atom has an unpaired electron with a large magnetic moment, which means that around each Gd-DTPA molecule is a magnetic field that perturbs the value of **B**. All the water molecules that come near a Gd-DTPA molecule will be influenced by this perturbing field, producing the relaxation enhancing effects described in §2.4. The changes in T1 and T2 are proportional to the local concentration of Gd-DTPA.

The blood-brain barrier normally prevents Gd-DTPA from entering the brain tissue directly. As a result, the T1 effect in the brain is usually small, since only about 5% of the brain volume is occupied by blood. However, if the blood-brain barrier is disrupted, the Gd-DTPA can leak into the brain tissue, reducing T1 significantly. This effect provides a way of detecting even small tumors.

2.7. Imaging Methods Used in fMRI

There are at least 50 different imaging methods used in MRI, each one designed to provide a different kind of contrast or to overcome various limitations of the scanner hardware. This section discusses only some of the methods most commonly used in functional brain imaging.

Multi-Slice and 3D Imaging Techniques

The scanning methods described previously produce images of a single slice. This is inadequate for most anatomical and functional imaging purposes. The obvious way to image the entire brain is to acquire all the data from one slice, then move on to another slice and scan it, etc. This is exactly what is done in the *single-shot* imaging methods described later. In multi-shot imaging, it is possible to acquire the data from different slices in an interleaved fashion. Instead of exciting the same slice with each shot in succession, the RF pulses are programmed to cycle through the slices systematically. This is how the TR can be stretched out without wasting time. For example, if it takes 40 ms to carry out a single shot, and the desired TR is 1 s, then 25 slices can be acquired in this cyclic fashion. Instead of waiting 960 ms after the end of data readout to begin the next RF pulse, the scanner simply moves to another slice and excites it. Since the RF pulses are slice selective, the magnetization in the first slice isn't affected by the excitation applied to the other 24 slices. As a result, the magnetization in the first slice will have time to recover (T1 relaxation) the desired amount—which is the purpose of making TR longer.

It is possible to acquire the slices in any order. In practice, slices are almost always gathered in an interleaved order: for 7 slices, the order would be 1-3-5-7-2-4-6. The reason for this is that slice selection is not perfect. Most of the RF pulse affects the magnetization in the desired slice, but a little of it leaks over to the neighboring slices. If slices were gathered in sequential order (1-2-3-…), then the effects of the RF excitation for slice 3 would be visible in slice 4. By delaying the imaging of slice 4 as long as possible, the small changes in magnetization in slice 4 due to slice selection leakage from slice 3 will have had time to relax away. With the interleaved slice order, each slice is scanned approximately ½·TR away from its neighbors.

Another way to minimize slice leakage that is sometimes used is to specify a small gap between the slices; for example, to use 5 mm thick slices on a 6 mm spacing, leaving a 1 mm gap. The disadvantage of this is obvious; less than full brain coverage. It is normally desirable only to do this if TR is very short and experimentation shows that the inter-slice leakage effect is unacceptable.

If a gradient field is not applied during the RF excitation, all the tissue in the range of the RF transmitter will have its magnetization excited. It is possible to do true 3D (or volume) imaging using such a non-selective RF pulse (sometimes called a *hard pulse*). The slice selection gradient is replaced by a second axis of phase encoding after the excitation. The advantage of volume imaging is that the slices can be made thinner than is convenient with 2D multi-slice techniques (thinner slices require applying more RF excitation energy). The disadvantage of volume imaging is that the TR must be small if all the data needed for 3D reconstruction is to be gathered in a reasonable time. As a result, 3D images tend to be T1-weighted. For anatomical reference purposes, this is very useful, since gray and white matter have significantly different values of T1. For other purpose, this T1-weighting is often undesirable.

FLASH—Fast Low Angle SHot Imaging

The multi-shot imaging methods shown in Figs. 2.10-11 are called *FLASH* techniques. For the gradient echo sequence, it is practical to get a single data acquisition time (RF pulse, phase encoding, and readout) down to about 20 ms. Acquisition of a 64×64 image then takes 64 shots, or about 1.2 s. In the MRI literature, this speed range is referred to as "fast imaging".

FLASH techniques have been used for functional MRI experiments. The principal advantage of FLASH is that it does not require very strong gradient fields. The principal disadvantage is that coverage of the whole brain, with 25 slices (say), takes about 30 seconds. As will be discussed later, it takes 60-100 measurements of each voxel to get good fMRI data; this means that single experiment would take 30-50 minutes. The longest a subject can usually tolerate lying in the scanner is 1-2 hours.

As a result, FLASH is usually used to study only a few slices (e.g., to cover primary visual cortex). For whole brain fMRI, "ultra-fast" imaging methods are needed.

EPI—Echo Planar Imaging

With stronger gradient fields, it is possible to complete the data readout portion of an imaging sequence in just 1 or 2 ms. Since the RF signal lasts for about 2·T2* (80 ms at 1.5 Tesla), such rapidity seems wasteful—the signal is read out very quickly, but after the readout is over, the signal persists for much longer.

In echo planar imaging, the transverse magnetization M_{xy} that still exists after the data readout is recycled and reacquired. This is done by "rewinding" the frequency encoding gradient, applying some more phase encoding gradient, and then doing the frequency encoding readout all over again. And again, and again. The only thing that limits how long this can go on is the intrinsic signal decay. With gradients that are both strong enough and agile enough to switch back and forth rapidly, an entire image can be read out with only one RF excitation, in a period of 40 ms (or even less). At this speed, 25 slices can be acquired in 1 second, and an entire fMRI experiment only takes a few minutes.

Fat suppression is important in EPI. Brain tissue itself has little fat, but the scalp and skull bone marrow have a great deal of fat. The very long readout times of echo planar imaging mean that the images are very sensitive to frequency errors. As a result, the fat image is displaced about ½ the size of the image in the phase encoding direction. Instead of being an easily ignored halo surrounding the brain, the fat image in EPI will lie in circular arcs curving through the center of the image—unless the NMR signals from fat molecules are properly suppressed.

Spiral Imaging

The imaging methods presented thus far have treated phase and frequency encoding as separate steps, with only one of the gradient fields active during data readout. Frequency and phase are intimately related: phase is the accumulation of past frequency, or frequency is the rate of change of phase. There is no particular reason why both gradient fields can't be on during readout. To prevent the overall NMR signal from becoming too dephased, each gradient must cycle between positive and negative values.

This class of imaging techniques is called spiral scanning, since in Fourier-space, the trajectory of phase coordinates describes a spiral. (In EPI, the trajectory of phase coordinates is like a back-and-forth raster scan.) As in EPI, spiral imaging can be carried out using one excitation or using several. Single shot spiral imaging requires strong gradients, like EPI; however, since the gradient strengths are changing continuously rather than intermittently, the gradient field system does not have to be able to switch on and off as rapidly as for echo planar imaging. The principal disadvantage of spiral imaging comes from the effects of imperfections in the scanner system. In EPI, large scale (≥5 cm) inhomogeneities in the static magnetic field result in distorted but clear images. In spiral imaging, the same phenomenon results in undistorted but highly blurred images. Usually the former situation is preferable. Both problems can be corrected for by measuring the magnetic field and performing compensations in the reconstruction software. At present, EPI is more common than spiral scanning perhaps simply because it is there is more experience with EPI than with spiral, and because the reconstruction software is simpler.

2.8. Hardware Issues

The Main Field

The large static magnetic field is generated by an electromagnet: currents flowing in cylindrically wound loops of wire—essentially a large solenoid. The inside of the cylinder is where the magnetic field is the strongest; this location is called the *bore* of the magnet. (Up to B_0=0.3 T, a permanent magnet can be used to generate the main field. This small field strength is useless for functional brain imaging, as far

as anyone knows.) The currents required are quite large; in some systems, over 1000 Amperes. In normal wires, the resistive heating would be very large. For this reason, high field scanners ($B_0 \geq 1.5$ T) use superconducting wire. Once established, the currents producing the main field will flow forever with essentially no losses. It is important to realize that the field is *always on*. Although it is insensible, the main field contains a great deal of energy; it is very dangerous to bring ferrous tools into the scanner room, since they might fly into the bore at a high speed.

To keep the wire in the superconducting state, it must be cooled to extremely low temperatures. This is usually done with liquid helium, which is contained in a huge Dewar surrounding the superconducting coils of wire. If the temperature of these coils ever rises above the "critical temperature" for their material, the wires convert back into normal resistive conductors. The current is still present, but it will now generate a huge amount of heat. This sudden influx of heat would destroy the magnet, so above the Dewar is a large heat exchange unit designed to remove this energy. Such an event is called a "quench" of the magnet. Quenching is very rare in properly installed and maintained systems.

Effects of Main Field Strength

NMR Signal Strength: When B_0 increases, the relaxed magnetization M_0 increases proportionately. Since M_0 is the ultimate source of the signal, a larger main field will produce a stronger image. The noise in the measurement of the signal also increases with B_0, but not as fast as the signal increases. At a smaller main field, the lower *signal-to-noise ratio* means that it takes longer to acquire high quality images.

Relaxation Rates: When B_0 increases, the small fields induced by tissue susceptibility also increase. This means that the microscopic fluctuations in the magnetic field are larger, which in turns means that the nuclear precession frequency varies more. The result is that the transverse relaxation rates increase; in other words, T2 and T2* both decrease as B_0 increases. In contrast, T1 increases as B_0 gets larger. The internal electromagnetic oscillations that cause longitudinal relaxation are weaker at higher frequencies, which means that it takes longer for them to have an effect on **M**.

Shim Coils

One reason that MRI is so expensive is the main magnet. It is not easy to build a system to generate a gigantic magnetic field that is also quite uniform. The field generated by the superconducting current is usually not quite uniform enough for imaging purposes. Even if it were, the introduction of a nonuniform object into the field would make the field nonuniform. The *shim coils* are an extra normally conductive electromagnet system, inside the superconducting coils. The currents in the shim coils are adjustable to add and subtract magnetic field strength in various locations. Some scanners have relatively few adjustable shim coils. Scanners with many shim coils offer the opportunity to adjust the shim currents for each individual subject to make the magnetic field more uniform, which will improve the image quality . The shimming process can be carried out automatically by special pulse sequences and software, or can be carried out manually by the scanner operator.

Gradient Coils

The wire loops that carry the currents which generate the spatial varying gradient magnetic fields are called gradient coils. There is one wire path for each field component (G_x, G_y, and G_z), but these coils are usually physically mounted together in the same device. The portion of the subject that is being imaged must be inside the gradient coil. In most scanners, there is a set of gradient coils integrated into the bore of the magnet—these are called the *whole body gradient coils*. It is also possible to have detached local gradient coils that are smaller and only enclose a part of the subject's body. One advantage of a local gradient coil is that the gradient strength (in Gauss/cm) is larger for the same input current. Another advantage is that it is possible to switch the currents on and off much faster in a small

coil than in a large coil. Both of these advantages make EPI of the brain much easier with local gradient coils than with whole body coils. The amount of current required for strong gradient fields is about 100 Amperes. For ultra-fast imaging such as EPI, these currents must be switched on and off in less than 100 microseconds.

The power supplies that provide the current for the gradient coils are called the *gradient amplifiers*. They are programmed by the scanner computer to provide the desired amount of current at the desired times. They must provide large currents on command, within a few microseconds, and do this quite reliably, so that the many repetitions of the pulse sequence required for image acquisition can be carried out exactly as programmed.

A electrical current in a magnetic field feels a force that pushes the current sideways. The current cannot leave the wire, so the force ends up pushing the wire itself sideways. The result is that the structure to which the wire is attached will bend a little under the influence of the force. When the current is switched off (or reversed), the force will be removed (or be reversed). When the current is switched rapidly, the result is to cause the whole gradient coil structure to vibrate. In EPI, the currents are switched about 1000 times per second. This rate of vibration is in the acoustic range, so the coil makes a audible noise while the gradient currents are switching rapidly—the entire readout interval. Since the gradients are used less in the other stages of an imaging sequence, the effect is an intermittent beeping sound, one beep per image. This noise can be very loud, especially in a head gradient coil, which is very close to the subject's ears. At the Medical College of Wisconsin, we have measured sounds of up to 95 dB at 1.5 Tesla, and up to 120 dB at 3.0 Tesla. It is crucial that subject wear earplugs or other hearing protection when echo planar imaging is being used. Other imaging methods are somewhat less noisy.

RF Coils

Inside the gradient coil is the RF coil. This is the device that transmits the radiofrequency energy for excitation. The shape of the RF transmit pulse is controlled by the RF synthesizer; the magnitude is controlled by the transmit amplifier. The shape of the RF pulse determines the frequency content, which is part of what determines the thickness of the excited slice (Fig. 2.9). The magnitude and duration of the RF pulse determines the flip angle of the excited magnetization. The RF coil also detects the emitted radiofrequency energy. A small voltage is induced in the coil, which is amplified and digitized into the reconstruction computer.

Whole body RF coils are often built into the scanner bore, so that the subject isn't aware of lying inside this elaborate assembly of nested coils. If a local gradient coil is used, a local RF coil must be used as well, since it is impractical to transmit the RF pulse through the metal in the gradient coil. (The situation is about like putting metal into a microwave oven.)

It is possible to use separate RF coils for transmission and reception. This is usually done when it is desired to acquire images from only a small region with the high signal level that results from placing the receive coil close to the RF emitting protons of interest. In such a case, the receive coil may be a small loop of wire placed over the desired area, and the transmit coil may be a larger device (perhaps the whole body coil). This type of imaging has a number of clinical applications, but has not been used extensively in fMRI research—most of the examples have involved a local coil placed over the occiput in order to acquire data from visual cortex. It is also possible to use multiple receive coils; the main purposes of this strategy are to increase the signal level or to extend the region being imaged. This technique is most frequently used in spinal imaging, and is not common in human brain fMRI.

Penetration Panel

The RF emitted by the subject is very weak. To prevent interference from other sources (e.g., TV stations), the entire scanner room is heavily shielded with metal. Nevertheless, electronic signals must

be able to get into and out of the scanner room—to control the electronics and to get the RF signal into the amplifier system. The single place where cables go into the scanner room is the penetration panel. At this point, all electrical paths that lead into the room are filtered to remove undesired RF.

In fMRI, it is often necessary to bring new signals into and out of the scanner room, to control the stimulus presentation equipment and to receive the subject monitoring data. These signals must also go through the penetration panel and have filters attached to their lines. All electronic equipment in the scanner room must be tested to see if it will work in the magnetic field environment and to see if it will interfere with the image quality. Some main magnets are shielded: they have a weaker electromagnet surrounding them with currents flowing in the opposite direction. The effect is to partly cancel the main field outside the bore. Since large magnetic fields will interfere with most electronic equipment, a shielded system can be installed in smaller quarters than an unshielded system, and may be easier to use with various peripheral devices.