4

Diffusion magnetic resonance imaging

The magnetic resonance application that the rest of this thesis will be concerned with is that of diffusion magnetic resonance imaging (dmri). This chapter provides a brief description of diffusion and how it can be examined in the brain with dmri. Mathematical models for diffusion in the brain are also presented, along with their theoretical and practical benefits and limitations. Finally, we look at some of the clinical uses of dmri.

4.1 The Einstein picture

Diffusion is a spontaneous phenomenon in any fluid whose temperature is greater than absolute zero (0 Kelvin). The molecules making up the fluid possess kinetic energy and are therefore constantly moving—the greater the energy, the faster the movement. The direction of this movement is random, and will typically change regularly as molecules collide with one another. Diffusion is often thought of as the process by which concentration gradients are flattened out, and we will initially describe it in these terms; but the principle is equally applicable to the movement of molecules within a fluid composed of a single type of molecule—in the latter case, the process is known as self-diffusion. Diffusion is well described by classical mechanics, so we will not need to make another foray into the quantum domain.

Consider first a one-dimensional example. We denote the concentration of some molecule at location x and time t with $C(x,t)$. The *flux*, or rate of movement of the molecules normal to the concentration gradient, is then given by

$$F = -D\frac{\partial C}{\partial x}\,, \tag{4.1}$$

where D is a constant known as the **diffusivity** of the fluid. As a result of this flux, however, the local concentration gradient will decrease, and so a time-dependent aspect needs to be introduced to describe the picture more fully. The equation

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} \tag{4.2}$$

was first arrived at by Adolf Fick, and so Eqs (4.1) and (4.2) are called Fick's Laws of diffusion (Fick, 1855; reprinted in translation in Fick, 1995).

If we assume that there are n molecules in total, all of which are at the location $x = 0$ at time $t = 0$—so that $C(x,0)$ is a Dirac δ-function—then diffusion will proceed such that

$$C(x,t) = \frac{n}{\sqrt{4\pi D t}}\exp\left(-\frac{x^2}{4Dt}\right), \tag{4.3}$$

as described by Einstein (1905). If we divide Eq. (4.3) by n, we obtain a properly normalised p.d.f. that describes the distribution of the molecules in the x dimension—or rather, since the distribution is dependent on t, a continuous-time stochastic process. In particular, we can see by inspection that the distribution $P(x|t)$ is a Gaussian distribution with $\mu = 0$ and $\sigma^2 = 2Dt$.

If the diffusion process is *isotropic*, or homogeneous across all orientations, then the generalisation to three dimensions is straightforward. Diffusion collinear with each of the vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$—the orthonormal unit vectors in the x, y and z directions respectively—is independent, and so the joint distribution is given by

$$P(\mathbf{r}|\mathbf{r}_0,t) = \frac{1}{(4\pi Dt)^{3/2}} \exp\left(-\frac{(\mathbf{r}-\mathbf{r}_0)^2}{4Dt}\right), \tag{4.4}$$

where $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, and $\mathbf{r}_0$ is the initial location of the molecules, which is not assumed to be zero in this general case. The mean of the distribution is now a vector, $\boldsymbol{\mu} = \mathbf{r}_0 = (x_0, y_0, z_0)$, while the variance is just as it was before: $\sigma^2 = 2Dt$.

The dependence of the description above on a concentration gradient does not present a problem for the case of self-diffusion. The fluid molecules may all be of a single species under these circumstances, but we can mentally label a molecule with initial position $\mathbf{r}_0$ as being (uniquely) of interest; and thereafter treat it as distinct from the rest of the fluid. The mean-squared distance that this molecule will travel during a diffusion time t is given by

$$\begin{aligned}\left\langle |\mathbf{r}-\mathbf{r}_0|^2 \right\rangle &= \left\langle (x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2 \right\rangle \\ &= \left\langle (x-x_0)^2 \right\rangle + \left\langle (y-y_0)^2 \right\rangle + \left\langle (z-z_0)^2 \right\rangle ,\end{aligned}$$

which is equivalent to the sum of the variances along each dimension, since $\mu_x = x_0$ and so on. We therefore easily arrive at

$$\left\langle |\mathbf{r}-\mathbf{r}_0|^2 \right\rangle = \sigma_x^2 + \sigma_y^2 + \sigma_z^2 = 6Dt\,. \tag{4.5}$$

Note that this equation for the mean-squared diffusion distance has no dependence on $\mathbf{r}_0$ since the fluid is assumed to be homogeneous, so that diffusion from all starting locations is statistically identical. Wherever a particular molecule starts, its diffusion distance from that point will be the same on average.

In general, diffusion is not isotropic. In a bowl of water it will be very close to isotropic, but in brain tissues—which contain large amounts of water but also various impermeable or semipermeable structures—diffusivity will vary from one direction to another. The Gaussian displacement distribution at time t therefore has in general the covariance matrix

$$\Sigma = 2\mathsf{D}t = \begin{bmatrix} 2D_{xx}t & 2D_{xy}t & 2D_{xz}t \\ 2D_{xy}t & 2D_{yy}t & 2D_{yz}t \\ 2D_{xz}t & 2D_{yz}t & 2D_{zz}t \end{bmatrix}, \tag{4.6}$$

which is symmetric, like any covariance matrix. The diffusivity values making up the matrix D are the components of a three-dimensional **diffusion tensor**, relative to the particular orthonormal basis set, $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$.[a] The special case of isotropic diffusion is then equivalent to the conditions

$$D_{xx} = D_{yy} = D_{zz} = D \qquad\qquad D_{xy} = D_{xz} = D_{yz} = 0\,.$$

In the brain, the main diffusing molecular species is water; and since a molecule of water contains two hydrogen nuclei it is visible to MRI. Anisotropic—that is, directionally inhomogeneous—diffusion is associated primarily with white matter, due to the highly linearised structure of this type of tissue (see Fig. 4.1), which is such that the local self-diffusion of water molecules is restricted to a far greater degree across a white matter tract than it is along it. Grey matter, by contrast, lacks any coherent linear structure, and so diffusion around that kind of tissue is much closer to isotropic.

[a] A tensor is an abstract mathematical construction which is independent of the coordinate frame being used. However, relative to any given set of basis vectors, it can simply be represented as a matrix of numbers. Further details would be superfluous here, but can be found in Riley *et al.* (2002).

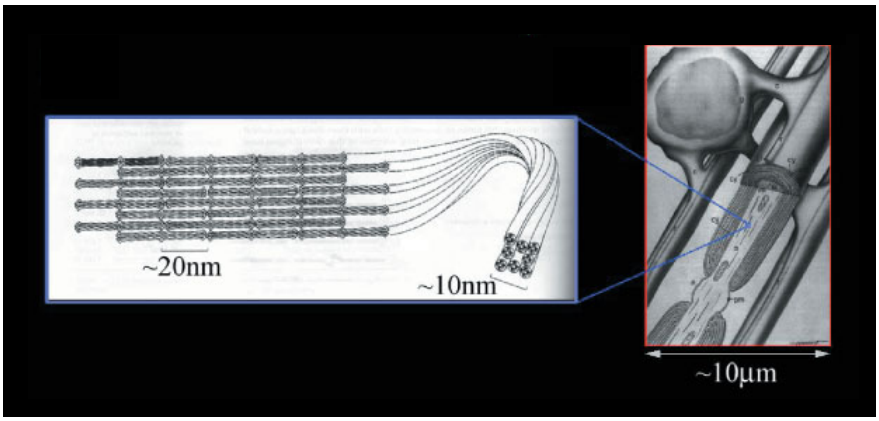

Figure 4.1: The linear microstructure of neural white matter. The axons and glia which make up connective tissue at the micron scale, and the neuronal filaments which are present at the nanometre scale, are mainly collinear, producing a consist bias in the local self-diffusion of water. Adapted from Mori & van Zijl (2002).

4.2 Diffusion tensor imaging

Diffusion sensitisation can be added to the standard spin-echo pulse sequence described in §3.4 by adding a symmetric pair of diffusion weighting gradients either side of the refocussing (180°) pulse, as shown in Fig. 4.2. The first of these gradients will offset the phase of the spins by an amount that depends on their location, and the second will provide equal and opposite rephasing *if* the spins have not moved. Since in practice the spins do move, and randomly, an isochromat will become dephased as the component spins spread out. The further the water molecules have diffused during the time, Δ, between applications of the gradient, the less perfect this rephasing will be, resulting in a smaller magnitude of final signal. Greater diffusivity is therefore indicated by a more greatly *attenuated* signal. It should be noted that this effect differs from coherent motion or flow, which will produce a phase shift in the spin isochromats, but will not attenuate the signal as random motion does.

It was shown by Stejskal & Tanner (1965), who first proposed the sequence, that for a diffusion weighting gradient of maximal magnitude G, applied for a time δ, the log-ratio between the signal, A, after the full echo time, TE, and that produced by the initial 90° RF pulse is given by

$$\ln\left(\frac{A(b)}{A_0}\right) = -\gamma^2\delta^2\left(\Delta - \frac{\delta}{3}\right)G^2D^{\text{eff}} = -bD^{\text{eff}}\,, \tag{4.7}$$

where b, which incorporates the relevant characteristics of the diffusion gradients, is known as

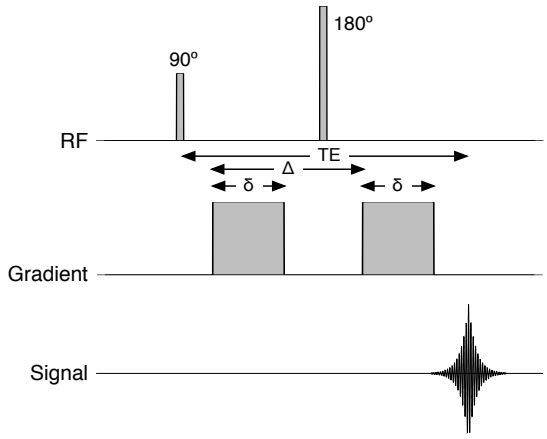

Figure 4.2: Pulse sequence timing diagram for a diffusion-weighted spin-echo experiment. Two diffusion sensitisation gradients are applied either side of the 180° pulse. They are switched on for a time δ in each case, and separated by a time Δ.

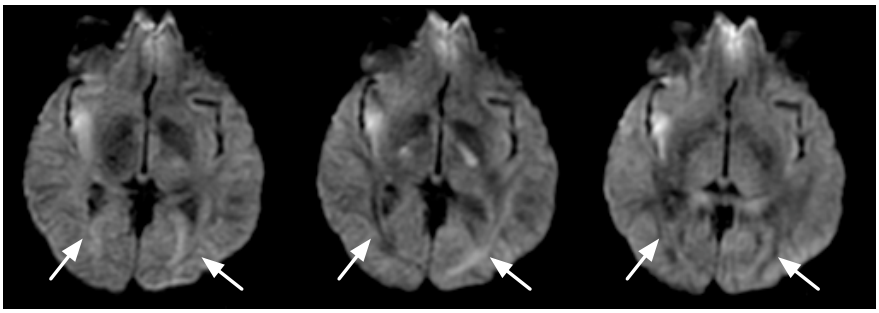

Figure 4.3: Diffusion-weighted MR images acquired with diffusion sensitisation along three orthogonal axes. The level of signal attenuation in some areas (such as those indicated with arrows) is evidently dependent on this direction. Images courtesy of Dr Susana Muñoz Maniega.

the diffusion weighting factor—a notation introduced by Le Bihan *et al.* (1986).

It is generally the case that diffusivity appears to vary with time rather than being constant, so Eq. (4.7) describes an *effective* diffusivity, D^{eff}, averaged over the diffusion time of the experiment. In tissue with an anisotropic diffusion profile, this "constant" will also vary with the orientation of the diffusion gradient applied to the sample (Moseley *et al.*, 1991; see also Fig. 4.3); and so we need to measure the whole diffusion tensor if we wish to characterise this situation more accurately. The extension of the principles described above to **diffusion tensor imaging** (DTI) was described by Basser *et al.* (1994a). In this case, Eq. (4.7) generalises to

$$\ln\left(\frac{A(\mathsf{b})}{A_0}\right) = -\gamma^2\delta^2\left(\Delta - \frac{\delta}{3}\right)G^2\mathsf{R}^{\mathsf{T}}\mathsf{D}^{\text{eff}}\mathsf{R} = -\sum_i\sum_j b_{ij}D_{ij}^{\text{eff}}\,, \tag{4.8}$$

where R is a normalised column vector describing the direction of the applied gradient, and b_{ij} are the elements of a symmetric matrix, b, which is analogous to the scalar weighting factor in Eq. (4.7). The elements of this weighting matrix encode various interactions between diffusion and imaging gradients, which can be quite complex and which vary from one type of sequence to another (the EPI case is described in Mattiello *et al.*, 1997). The equivalent scalar diffusion weighting factor to b is given by the trace of the matrix.

Since knowledge of the pulse sequence design is sufficient to establish the b matrix for any given acquisition, Eq. (4.8) represents a system of linear equations that can be solved for the six independent components of the tensor given values of A for six noncollinear diffusion gradient directions, plus the T_2-weighted signal, A_0.[b] However, in practice it is usual to apply more than six different gradient directions, since MR signal measurements are noisy, and then to fit the tensor statistically using multivariate linear regression.

There has been some debate in the literature over the particularities of optimising the choice of gradient scheme for various purposes (Hasan *et al.*, 2001; Papadakis *et al.*, 1999; Skare *et al.*, 2000), particularly the calculation of tensor-derived scalar metrics, which are described below. Broadly speaking, it is as well to acquire data for as many gradient directions as possible (Jones, 2004); and these are commonly arranged to coincide with the vertices of an icosahedron (Batchelor *et al.*, 2003), or to minimise the electrostatic repulsion force when the gradients are treated as point charges[c] (Conturo *et al.*, 1996; Jones, 2004).

Once the effective diffusion tensor has been estimated, it can be used to characterise local diffusion at each voxel in the brain in various ways. Since the matrix representing the tensor is symmetric in any coordinate frame, its eigenvectors are orthogonal and its eigenvalues real. We can therefore construct a local coordinate system from the eigenvectors, $\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3\}$, which are arranged by convention such that the largest eigenvalue is λ_1—corresponding to $\boldsymbol{\varepsilon}_1$—and the smallest is λ_3 (Basser *et al.*, 1994b).

The general shape of the diffusion tensor is commonly visualised using ellipsoids whose radii along each eigenvector direction are given by the square root of the corresponding

[b] Recall from chapter 3 that the basic spin-echo sequence is T_2-weighted, and this is the only factor in a sequence

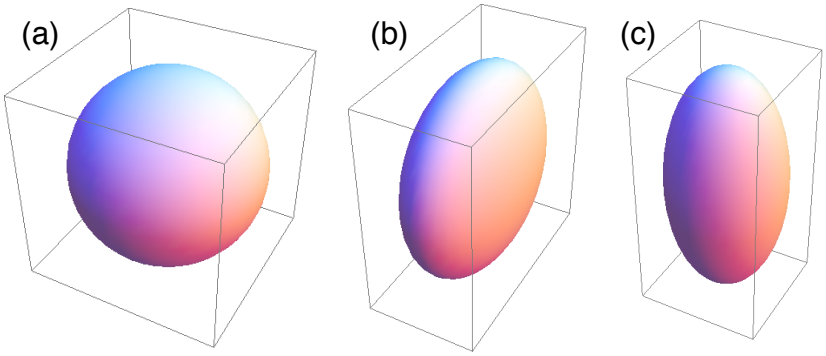

Figure 4.4: Ellipsoids representing isotropic **(a)**, oblate **(b)** and prolate **(c)** diffusion profiles.

eigenvalue (see Fig. 4.4). Thus, the case of isotropic diffusion ($\lambda_1 = \lambda_2 = \lambda_3$) is represented by a sphere, while oblate diffusion ($\lambda_1 = \lambda_2 > \lambda_3$) appears disc-shaped, and prolate diffusion ($\lambda_1 > \lambda_2 = \lambda_3$) appears cigar-shaped.

The average magnitude of the diffusion along the three eigenvectors can be calculated in a rotation-invariant way by taking the trace of the tensor matrix; or equivalently, the sum of the eigenvalues. This quantity is known as the *mean diffusivity* (MD):

$$\text{MD} = \langle \mathsf{D} \rangle = \frac{\mathrm{Tr}(\mathsf{D})}{3} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \,. \tag{4.9}$$

This quantity gives no indication of the anisotropy of the tensor, since it takes into account only the mean of the eigenvalues. There is, moreover, no single obvious way to index anisotropy. Three scalar valued measures that have been proposed are fractional anisotropy (FA), relative anisotropy (RA) and the volume ratio (VR), which are defined as follows (cf. Basser & Pierpaoli, 1996; Pierpaoli & Basser, 1996).

$$\text{FA} = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1 - \langle \mathsf{D} \rangle)^2 + (\lambda_2 - \langle \mathsf{D} \rangle)^2 + (\lambda_3 - \langle \mathsf{D} \rangle)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \tag{4.10}$$

$$\text{RA} = \frac{1}{\sqrt{3}} \frac{\sqrt{(\lambda_1 - \langle \mathsf{D} \rangle)^2 + (\lambda_2 - \langle \mathsf{D} \rangle)^2 + (\lambda_3 - \langle \mathsf{D} \rangle)^2}}{\langle \mathsf{D} \rangle} \tag{4.11}$$

$$\text{VR} = \frac{\lambda_1 \lambda_2 \lambda_3}{\langle \mathsf{D} \rangle^3} \tag{4.12}$$

A VR of unity represents isotropic diffusion, whereas FA and RA are zero when all three eigenvalues are equal. At the other end of the scale, FA and RA are maximal when $\lambda_2 = \lambda_3 = 0$, whereas VR is zero when any of the eigenvalues is zero. Of the three, FA gives the highest signal to noise ratio (Papadakis *et al.*, 1999), and is by far the most commonly used in the literature.

4.3 A more general displacement distribution

The tensor model makes the assumption that diffusion at the scale of a voxel is essentially Gaussian, which allows us to use the generalised Einstein equation—with covariance matrix given by Eq. (4.6)—as an appropriate model of the underlying process. However this assumption, as we will see later, is not always appropriate; and it is particularly prone to fail in regions where white matter tracts cross one another. Alternative models of diffusion which have been

with diffusion gradients of zero magnitude.

[c]The inspiration here is the behaviour of electrons in atomic orbitals, which are equally charged and therefore repel one another. They spontaneously space themselves out as a result.

developed with the particular application of fibre tracking in mind will be discussed in chapter 5, but we will describe here an alternative which predates such applications substantially.

The origins of **q**-space MRI can be traced back to the work of Edward Stejskal, who described how a special case of the Stejskal–Tanner pulse sequence (Fig. 4.2) could be used to infer an arbitrary local displacement distribution. If the time during which the diffusion gradient is applied, δ, is made to be very short—in particular so that $\delta \ll \Delta$—then the signal attenuation ratio of the experiment is given in general by

$$\frac{A(\mathbf{G},\Delta)}{A_0} = \int P(\mathbf{r}_0) \int P(\mathbf{r}|\mathbf{r}_0,\Delta) \exp(-\mathrm{i}\gamma\delta(\mathbf{r}-\mathbf{r}_0)\cdot\mathbf{G})\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}_0\,, \tag{4.13}$$

where the vector **G** embodies the direction and magnitude of the diffusion gradient (Stejskal, 1965). Under the assumption of local homogeneity $P(\mathbf{r}_0)$, which represents the initial distribution of diffusing molecules within the volume of interest, is uniform; and so the outer integral can be ignored. Stejskal also showed that if the Gaussian displacement distribution given by Eq. (4.4) were used for $P(\mathbf{r}|\mathbf{r}_0,\Delta)$, then Eq. (4.13) becomes equivalent to Eq. (4.8), albeit with $\Delta - \delta/3$ replaced by Δ due to the narrow gradient pulse assumption.

Callaghan *et al.* (1988) later proposed that the direction, magnitude and duration properties of the diffusion gradient be parameterised as

$$\mathbf{q} = \frac{1}{2\pi}\gamma\delta\mathbf{G}\,,$$

by direct analogy with the **k** vector that is so central to magnetic resonance imaging theory (cf. §3.4). Using this notation, and taking $\mathbf{r}_0 = \mathbf{0}$, which gives no loss of generality if we assume local homogeneity, the attenuation ratio becomes

$$\frac{A(\mathbf{q},\Delta)}{A_0} = \int P(\mathbf{r}|\Delta) \exp(-\mathrm{i}2\pi\mathbf{q}\cdot\mathbf{r})\,\mathrm{d}\mathbf{r}\,, \tag{4.14}$$

which represents a Fourier transform of the displacement distribution. By the Fourier inversion theorem (Riley *et al.*, 2002), we can therefore recover the distribution by means of the inverse transform

$$P(\mathbf{r}|\Delta) = \frac{1}{A_0} \int A(\mathbf{q},\Delta) \exp(\mathrm{i}2\pi\mathbf{q}\cdot\mathbf{r})\,\mathrm{d}\mathbf{q}\,. \tag{4.15}$$

By sampling signal values from a series of locations in **q**-space—typically achieved by incrementally stepping up the gradient strength and changing its direction—one can therefore capture the diffusive behaviour of water molecules in the brain at different length scales and over various diffusion times.

The appeal in acquiring a model-free estimate of the diffusion displacement distribution is clear—modelling assumptions are avoided, and so one need not worry about their validity. However, the narrow gradient pulse assumption made by **q**-space theory is itself problematic. Whilst δ must be small enough so that the pulse can be approximated by a Dirac delta function, the time integral of the pulse given by $\delta\mathbf{G}$ must be finite, otherwise **q** will be zero and there will be no signal attenuation at all. As a result, the magnitude of the gradient pulse needs to be very large. Such gradient strengths are attainable using modern hardware—although they are out of the reach of most clinical MR scanners—but they are very demanding to generate and may have adverse effects on the subject. Hence, studies that have closely approximated the narrow pulse assumption (e.g. Biton *et al.*, 2006, who used the parameters $\delta = 2$ ms, $\Delta = 50$ ms, $G_{\max} = 500$ mT m^{-1}) have worked with excised (*ex vivo*) rather than living (*in vivo*) tissue.

4.4 The role of registration

Since all but the simplest of DMRI experiments require multiple image acquisitions with different gradient directions, the basic data from which information will be derived is a series of brain volumes. Although motion within volumes will be minimised by using an EPI-based pulse sequence, one cannot rule out the possibility that the subject will move during the whole

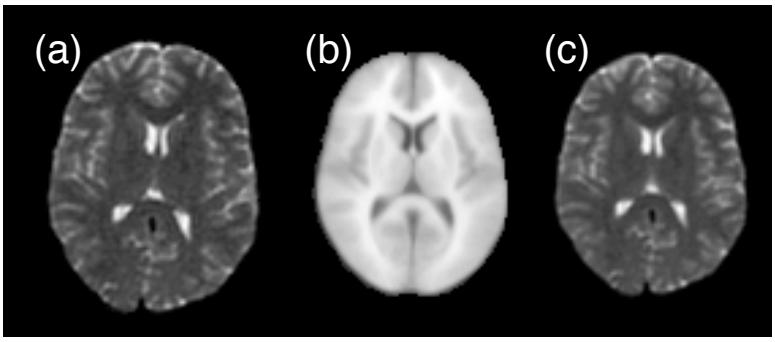

Figure 4.5: Registration of a T_2-weighted EPI image **(a)** to a T_1-weighted standard brain volume **(b)** produces a version of the original in standard space **(c)**. Note that the general orientation and scale of subfigure (c) correspond to those of (b), but the details of the image do not match perfectly. The different contrast types of subfigures (a) and (b) is not a problem if the cost function is chosen appropriately.

experiment, particularly if the number of gradient directions is large. Moreover, the varying orientations of the diffusion gradients will result in differing eddy current induced distortion effects from one volume to another. It is therefore unwise to assume that the subject's brain is positioned consistently in the field of view throughout a scanning session.

The process of realigning the three-dimensional images is **registration**. Image registration is usually framed as an optimisation problem in which an algorithm attempts to find a global transformation which minimises some cost function indicating the "distance" between two images. A number of cost functions have been used for this purpose, typically based on the correlation or mutual information between image intensity data; but a more divisive issue is the scope of the transformations allowed by the algorithm. The number of degrees of freedom varies from six for a rigid-body transformation—translation by a vector $\mathbf{L} = (L_x, L_y, L_z)$ and rotation by angles ϕ, θ and ψ about the x, y and z axes—up to hundreds or thousands for a complex nonlinear approach, which may involve local as well as global optimisation. Nonlinear methods have the advantage of providing a better match between the original image and the target image, but are slower due to having to optimise over a much larger parameter space, and pose a risk of overfitting.

General purpose linear registration algorithms optimise over affine transformations (Friston *et al.*, 1995; Jenkinson & Smith, 2001; Woods *et al.*, 1998), which incorporate the rigid-body parameters as well as a scaling vector, $\mathbf{S} = (S_x, S_y, S_z)$, and three shear terms: H_{xy}, H_{xz} and H_{yz}. The resulting affine transformation matrix is therefore composed of the product

$$\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 & L_x \\ 0 & 1 & 0 & L_y \\ 0 & 0 & 1 & L_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & \sin\phi & 0 \\ 0 & -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & -\sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ \sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$\times \begin{bmatrix} \cos\psi & \sin\psi & 0 & 0 \\ -\sin\psi & \cos\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & H_{xy} & H_{xz} & 0 \\ 0 & 1 & H_{yz} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

This composite matrix may then be used to transform directly the grid of voxel locations, making up the **native space** in which the original image is acquired, to their equivalent points in the target space. The image data must then be interpolated onto this new grid. The interpolation scheme for this final step may need to be chosen to suit the particular application, but a trilinear scheme is often adequate.

Affine registration of diffusion-weighted images to a T_2-weighted reference image from the same scanning session is an effective way to correct for eddy current induced distortions in the former, and it simultaneously transforms all of the individual scans into a common space

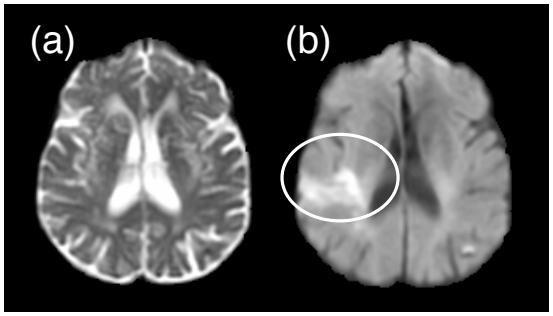

Figure 4.6: Images taken from a patient 11 hours after stroke onset. The T_2-weighted image **(a)** is normal in the lesion region, but the averaged diffusion-weighted image **(b)** shows significantly reduced diffusion compared with the equivalent region in the contralateral hemisphere. Images courtesy of Dr Susana Muñoz Maniega.

so that the correspondence between voxels in each volume is improved. Registration is never perfect, however, and it should be borne in mind that the data used to fit a diffusion tensor (for example) at each voxel cannot truly be said to be taken from a single fixed location in the brain. Some inaccuracy is inevitable.

For comparative studies involving multiple subjects, a popular strategy is to register each subject's reference image to an established standard image such as that described by Evans *et al.* (1993), thus transforming them all into a common **standard space** (see Fig. 4.5). In this case, since no two brains are merely stretched and sheared versions of one another, linear registration is strictly inadequate. The approximation suffices, however, for some purposes.

4.5 Diffusion MRI in the clinic

Le Bihan *et al.* (1986) were the first to demonstrate the clinical potential of DMRI. They showed that the presence of astrocytomas (a type of tumour originating in astrocytes) or oedema (swelling due to the accumulation of excess fluid) produced measurable differences in effective diffusivity, when compared with normal tissue. They also demonstrated reduced diffusivity in normal white matter compared to grey matter, which is now established as a standard finding.

Diffusion imaging has been widely used to study acute ischaemic stroke (damage to the brain resulting from a blockage in its blood supply), and has been shown to provide useful information beyond that which is available to structural T_1- or T_2-weighted MRI (Baird & Warach, 1998). In particular, reduced diffusivity can be observed in ischaemic tissue very soon after the stroke onset, while T_2 relaxation times are largely unaffected until oedema develops, which takes place much later (Knight *et al.*, 1991; see also Fig. 4.6).

The advent of DTI has made it possible to examine the effects of disease on the diffusion properties of anisotropic tissues—i.e. white matter. With mean diffusivity acting as a proxy for overall water content, and anisotropy indices—in practice, almost invariably FA—indicating the degree of "coherence" or "integrity" of the linear structure intrinsic to white matter, various low level pathological processes such as oedema, neurotoxicity or Wallerian degeneration might plausibly be expected to have some MR-visible impact. These DMRI-derived measures have therefore been applied to investigate the effects of a diverse array of diseases such as multiple sclerosis, amyotrophic lateral sclerosis and Alzheimer's disease (Horsfield & Jones, 2002); as well as psychiatric disorders like schizophrenia, alcoholism and geriatric depression (Lim & Helpern, 2002). More pathologies are being studied year on year.

There has also been significant interest in the effects of normal ageing on white matter (Moseley, 2002; Sullivan & Pfefferbaum, 2006). Anisotropy has been found to be higher in young adults than children (Klingberg *et al.*, 1999), but it then tends to reduce with time through adulthood and into old age (Pfefferbaum *et al.*, 2000), presumably representing the processes of maturation and then degradation of connective tissue. The gradual decline in white matter integrity is accompanied by a tendency for subjects' performance on mental tasks, particularly those using working memory, to decrease with time; and may represent its cause. Moreover, it has been shown that statistically compensating for mental ability in childhood—as measured with an IQ test at age 11—substantially attenuates the relationship, at age 83, between FA and cognitive test performance (Deary *et al.*, 2006), suggesting that childhood IQ may have a bearing on white matter integrity later in life.

Another interesting aspect of normal ageing which has been investigated with DMRI is the phenomenon of leukoaraiosis—also known, rather less concisely, as periventricular white matter hyperintensity—which manifests itself as regions of abnormally high signal on T_2-weighted images, and which occurs in many healthy older subjects as well as some stroke patients. Jones *et al.* (1999) demonstrated higher diffusivity and lower anisotropy in areas of leukoaraiosis than in normal tissue, and showed that a map of MD highlights the distinction between leukoaraiosis and the ventricles better than a T_2-weighted image. More recently, Bastin *et al.* (2007) further demonstrated that FA correlates strongly with magnetisation transfer ratio[d] in regions of leukoaraiosis, but not in comparable normal-appearing white matter, indicating that the loss of white matter integrity in such regions may be tied to a breakdown in myelination.

Due to the demands of **q**-space imaging on MRI hardware, it has been used far less than other forms of DMRI in the clinical domain. Those studies that have employed the technique have been required to essentially abandon the narrow gradient pulse requirement—Assaf *et al.* (2002, 2005) used the parameters $\delta = 65$ ms and $\Delta = 71$ ms, at a b-value equivalent of 14,000 s mm^{-2}; compared with 353,000 s mm^{-2} in a true **q**-space experiment (Biton *et al.*, 2006), and just 1000 s mm^{-2} in a typical DTI acquisition. However, it has been shown that even under these circumstances, meaningful information about the displacement distribution can be recovered (Lori *et al.*, 2003).

Although DMRI is unique as a technique for studying structural connectivity and white matter integrity, functional magnetic resonance imaging (FMRI), which gives an indication of the level of activity across the brain, provides complementary information. By looking for consistent patterns of correlated activity in different parts of the brain, a degree of *functional* connectivity between regions can be inferred. There have been a number of attempts to combine FMRI and DMRI data acquired from the same subject together (e.g. Cherubini *et al.*, 2007; Guye *et al.*, 2003; Staempfli *et al.*, 2008), and this is likely to remain an active research area for some time.

4.6 Summary

We have discussed the physical process of diffusion, and the means by which diffusion displacement distributions of varying complexity can be indirectly measured with MRI. A number of scalar indices indicating the shape of the diffusion tensor have been described—notably the widely used fractional anisotropy. The uses to which these methods have been put in the clinic, including studies of ageing and stroke, have also been briefly surveyed. The existence and measurability of anisotropic diffusion in the white matter of the brain are crucial prerequisites for DMRI-based tractography; and it is to that application that we turn next.

[d] The magnetisation transfer ratio is a metric derived from magnetisation transfer MRI, a method which has not been described above. It is sensitive to changes in large molecules such as myelin.