11 Lensless Imaging and Inverse Problems

The previous chapter completed our tour through Maxwell's equations, showing how simple dielectric interfaces can serve as lenses to form images. While the power and simplicity of such a lens is appealing, it is also limiting. A confocal microscope scans the position of its sample, something that is not feasible if an image is being made of, say, a planet. Optics are available to image just a fraction of the range of energies and mechanisms that are useful for probing an object, and even then there is information in the interactions that can reveal information far beyond what is seen in a conventional image.

To address these problems we will now introduce intelligence into the apparatus. This will start by abstracting the operation performed by a lens into something that can be implemented by measuring and processing other kinds of signals, such as microwaves or sound. Then we will look beyond two-dimensional images to explore a number of tomographic techniques to recover three-dimensional structure, closing with an introduction to the general problem of inverting a set of measurements to deduce the signal sources.

11.1 MATCHED FILTERS AND SYNTHETIC LENSES

The first step is synthesizing lenses is to introduce the notion of a *matched filter*, which we will see is the optimal linear detector for a known signal. Matched filters will provide a way to generalize the response of a lens to other domains.

Consider a signal $x(t)$ with Fourier transform $X(\omega)$ passed through a linear filter that has an impulse response $f(t)$ [Gershenfeld, 1999], or equivalently, a frequency response $F(\omega)$ given by the Fourier transform of the impulse response. The frequency domain response of the output of the filter $Y(\omega)$ will be the product of the Fourier transforms

$$Y(\omega) = X(\omega)F(\omega) \quad , \tag{11.1}$$

and the time domain response will be the convolution

$$y(T) = x(T) * f(T) = \int_0^T x(T-t)f(t)\,dt \quad , \tag{11.2}$$

where the bounds of the integral are the time during which the signal has been applied to the filter.

The size of the output of the filter can be bounded by the integral version of the

Cauchy–Schwarz inequality: the magnitude of the integral of a product is less than or equal to the product of the integrals of the magnitude

$$y^2(T) = \left| \int_0^T x(T-t)f(t)\,dt \right|^2$$
$$\leq \int_0^T |x(T-t)|^2\,dt \int_0^T |f(t)|^2\,dt \quad . \tag{11.3}$$

By inspection, this bound will be saturated if

$$f(t) = Ax^*(T-t) \tag{11.4}$$

for arbitrary A. The filter will produce the maximum output for a given input signal if the filter's impulse response is equal to the complex conjugate of that signal reversed in time. This is called a matched filter, and is routinely used to detect and time known signals such as radar echoes.

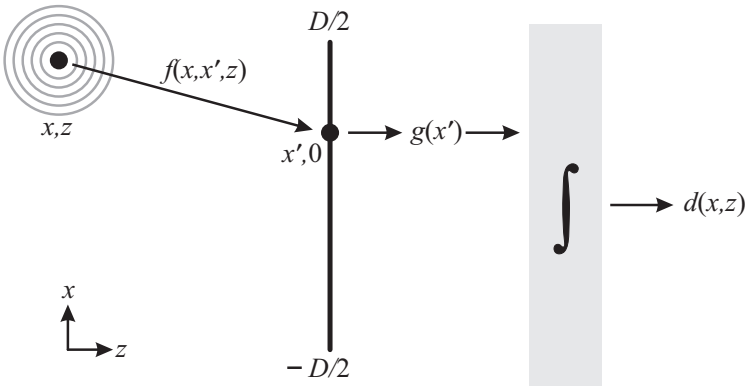

Figure 11.1. Mathematical model of a lens.

What do matched filters have to do with lenses? Consider the abstract view of a lens shown in Figure 11.1. The amplitude and phase of a periodic wave that starts at the point (x, z) and arrives at a point on the surface of the lens $(x', 0)$ will be changed by the wave's propagation. This change in amplitude and phase can be described by multiplying the wave by a complex factor $f(x, x', z)$ that depends on the distances. The signal seen by a detector on the far side of the lens will be equal to the integral over the surface of the lens of f multiplied by another term $g(x')$ that contains the amplitude and phase shifts introduced by passing through the lens and then traveling to the detector; let's call the overall factor giving the change in detected amplitude and phase $d(x, z)$.

f will be the product of a phase shift due to the wave traveling the distance from (x, z) to $(x', 0)$ and an amplitude change due to the wave spreading out. Since the amplitude change is a higher-order correction than the phase shift, in this calculation we will include just the phase shift, which is

$$f(x, x', z) = e^{ikr}$$
$$= e^{ik\sqrt{(x-x')^2+z^2}}$$

$$\begin{aligned} &\approx e^{ik[z+(x-x')^2/2z]} \\ &= e^{ikz} e^{ik(x-x')^2/2z} \quad . \end{aligned} \tag{11.5}$$

We have once again made the paraxial approximation ($x - x' \ll z$) for sources near the axis of the lens. The e^{ikz} term can also be dropped because we will be interested only in the functional form of the transverse dependence of f on x.

The response of the detector is found by integrating over the lens the propagation term f and the lens response function g:

$$d(x, z) = \int_{-D/2}^{D/2} g(x') f(x - x', z)\, dx' \quad . \tag{11.6}$$

This is just a convolution over x', therefore if we want the lens to produce the maximum output for a signal at a given location (i.e., to focus the waves from the source onto the detector) then from the last section we've learned that the lens should be designed to have a response function equal to the complex conjugate of the propagation factor, with the direction reversed

$$\begin{aligned} g(x') &= f^*(-x') \\ &= e^{-ik(x+x')^2/2z} \quad . \end{aligned} \tag{11.7}$$

For a source on the axis ($x = 0$) this reduces to

$$g(x') = e^{-ikx'^2/2z} \quad . \tag{11.8}$$

The matched-filter lens has a quadratic phase dependence on the distance from the axis x'.

Now let's look at how such a lens responds to a signal away from the focus. If the lens is matched to a source on-axis at $(0, z)$ but the source is actually at (x, z), the response is

$$\begin{aligned} d(x, z) &= \int_{-D/2}^{D/2} g(x') f(x - x', z)\, dx' \\ &= \int_{-D/2}^{D/2} e^{-ikx'^2/2z} e^{ik(x-x')^2/2z}\, dx' \\ &= e^{ikx^2/2z} \int_{-D/2}^{D/2} e^{-ikxx'/z}\, dx' \\ &= -De^{ikx^2/2z} \frac{\sin(kxD/2z)}{kxD/2z} \\ &= -De^{ikx^2/2z} \operatorname{sinc} \frac{xD}{z\lambda} \quad , \end{aligned} \tag{11.9}$$

where $\operatorname{sinc}(x) \equiv \sin(\pi x)/\pi x$. This is called the *point spread function*; it has a central peak and side lobes, and in the limit $D \to \infty$ it becomes a delta function (Figure 11.2). As we saw with Gaussian optics, the resolution of the lens increases as the aperture increases; Problem 12.2 looks at resolution in more detail.

Matched filter lenses can be implemented for a variety of purposes. By doing the delay and summation electronically rather than physically in a piece of glass it is possible to make lenses for a much broader range of types and wavelengths of radiation, and the properties of such lenses can quickly be modified. In the simplest implementation, the

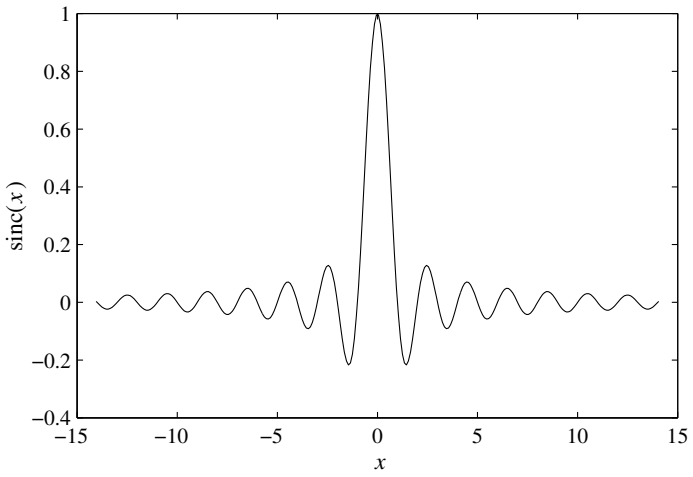

Figure 11.2. $\text{sinc}(x) \equiv \sin(\pi x)/\pi x$.

lens is made up of an array of transmitters and receivers. Doing the calculation that we just did in reverse shows that by introducing delays in the outgoing signals it is possible to produce an electronically steerable beam, and then with the same delays on the return signal it is possible to probe for an object at a particular point. This is how *phased-array radar* works; the number and spacing of the individual transceivers produce *grating lobes* that further blur the overall point spread function that results from the aperture size. Since the resolution of the lens will increase as the size of the aperture D increases it is desirable to make that as large as possible, but physical constraints such as the size of a satellite will limit the maximum aperture. However, a single moving transmitter and receiver can do the work of a much larger array by sending out a signal and collecting the return from many points along a trajectory, such as the satellite's orbit. This is called a *synthetic aperture* [Fitch, 1988], and is used in both satellite radars mapping the Earth and towed sonar transducers mapping the ocean (Problem 12.3).

11.2 COHERENT IMAGING

The preceeding calculation assumed that the radiation was *coherent*, so that interference is possible independent of the distance that a wave travels. We will shortly see the limits of that assumption, but if it is justified then a detector that records the wave's mean square amplitude will throw away the information stored in the phase. That can be recovered by *holography*, which saves the interference rather than just the intensity [Gabor, 1948; Gabor, 1966]

Consider two beams of coherent radiation, such as sound or light, incident on a piece of film or other medium which responds to the intensity in the (x, y) plane (Figure 11.3). One of the beams is unmodified, and the other one has been scattered from an object. After traveling a distance r, the phase of the reference beam in the plane of the film will

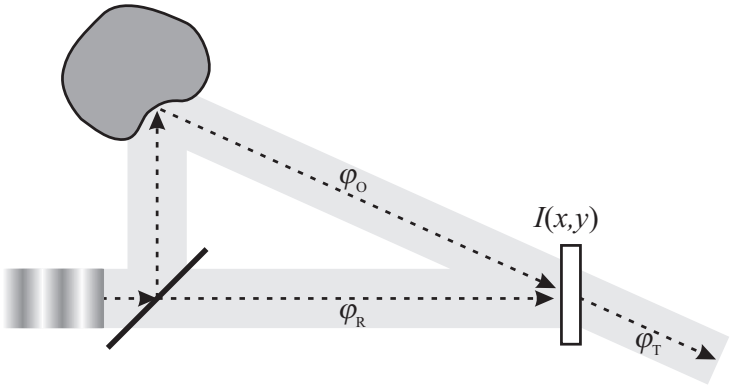

Figure 11.3. Creation of a hologram.

be

$$\varphi_R = e^{-i\vec{k}\cdot\vec{r}} = e^{-i(k_x x + k_y y)} \quad , \tag{11.10}$$

and the scattered beam will have its amplitude and phase modified by the object:

$$\varphi_O = A(x,y)e^{-i\theta(x,y)} \quad . \tag{11.11}$$

The recorded intensity on the film is the square magnitude of the sum of these:

$$\begin{aligned} I(x,y) &= |\varphi_O + \varphi_R|^2 \\ &= (\varphi_O + \varphi_R)(\varphi_O + \varphi_R)^* \\ &= \varphi_O\varphi_R^* + \varphi_O^*\varphi_R + \varphi_O\varphi_O^* + \varphi_R\varphi_R^* \\ &= A(x,y)e^{-i\theta(x,y)}e^{+i(k_x x + k_y y)} + A(x,y)e^{+i\theta(x,y)}e^{-i(k_x x + k_y y)} \\ &\quad + |A|^2 + 1 \quad . \end{aligned} \tag{11.12}$$

Now, if an identical reference beam is later sent through this film, the transmitted beam will have its intensity modulated by the recorded pattern:

$$\begin{aligned} \varphi_T &= I(x,y)e^{-i(k_x x + k_y y)} \\ &= A(x,y)e^{-i\theta(x,y)} + (\text{background terms}) \quad . \end{aligned} \tag{11.13}$$

The first term gives rise to a transmitted beam that is identical to the light arriving from the object, and the other terms give rise to various kinds of background (glare and unfocused virtual images) which can be eliminated by more sophisticated geometries. The transmitted beam that is identical to the arriving light truly is identical. There is no difference between it and what would be seen if the object really was there, up to the limits of the film. The most obvious difference between a hologram and a normal incoherent film image is parallax: the image looks different from different directions.

Instead of creating holograms by recording light bounced off a real object, it is also possible to create them numerically to make synthetic holograms either on film or in real time with a fast acousto-optic modulator (Chapter 13) [Lucente, 1997]. Since holograms can reproduce the light that would be generated by an arbitrary source, it is also possible to design holograms to serve as *Holographic Optical Elements* (*HOE*s), thin sheets that

can transform coherent light in ways that are impossible with conventional lenses [Stern, 1996]. HOEs are routinely used in applications such as laser beam scanning, wavelength multiplexing, and heads-up displays [Sweatt, 1979; Rallison, 1984; Herzig & Dandliker, 1987].

Real light sources are of course not perfectly coherent. If we approximate the spectrum as a Gaussian with a width $\Delta\omega$, then the temporal distribution will also have a Gaussian envelope with a width of $\Delta t = 1/\Delta\omega$ because the Fourier transform of a Gaussian is also Gaussian with the variance inverted (equation 5.23). Therefore,

$$\Delta\omega \, \Delta t = 1 \quad , \tag{11.14}$$

or

$$\Delta f \, \Delta t = \frac{1}{2\pi} \quad . \tag{11.15}$$

This is called the *frequency–time uncertainty relation*. Associated with the temporal spread is a distance $c\,\Delta t$, called the *coherence length*. This is the maximum path difference over which interference effects can be seen. The most stable lasers have linewidths below a hertz, giving a coherence length so large that it can be ignored for all but the most sensitive metrology measurements. Incoherent lights sources can have linewidths of tens to hundreds of nanometers, giving coherence lengths of a few microns.

If a hologram is illuminated by broadband light, each wavelength will scatter at a different angle, blurring the image. This can be partially compensated for by using a slit to restrict the beamwidth reflected from the object used to make the hologram, so that when it is read by incoherent light the images associated with different colors will not overlap [Benton, 1969].

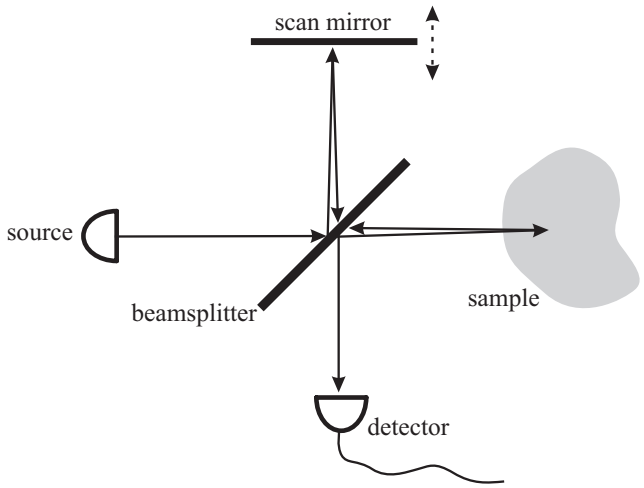

Figure 11.4. Optical Coherence Tomography (the beams have been displaced for clarity).

Limited coherence becomes a feature rather than a bug when it is used in *Optical Coherence Tomography* (*OCT*) [Huang *et al.*, 1991, Fujimoto, 2001, Fercher *et al.*, 2003]. The idea is shown in Figure 11.4. Light from a source with a short coherence length enters a *Michelson interferometer*. One beam reflects off of a mirror, and the

other one scatters off of the sample. These are recombined by the beamsplitter and detected by a photodetector. If the two path lengths exactly match the beams will add coherently, giving a strong signal. Because of the short coherence length, if the paths are slightly different then the detector will see an incoherent background. Scanning the mirror then gives a profile of the scattering as a function of depth in the sample; these measurements can be assembled into a three-dimensional image by moving the sample also. The resolution, speed, and sensitivity can all be improved by scanning the frequency rather than the geometry [Choma *et al*., 2003, Wojtkowski *et al*., 2004]. OCT is promising for biomedical applications because it can provide internal information with a benign light source.

11.3 COMPUTED TOMOGRAPHY

OCT is limited in the depth that can be probed because photons must be scattered back to the detector, but if they scatter multiple times in the sample the image will degrade. An opposite limit applies to higher-energy radiation such as X-rays: most photons either travel straight through the material or are absorbed. Measurement of this absorption as a function of orientation can reveal the internal structure via *Computed Tomography* (CT), also know as *Computerized Axial Tomography* or *Computer-Assisted Tomography* (CAT) [Kak & Slaney, 1988].

Let $\rho(x, y)$ be the distribution of the absorption coefficient of the material in a two-dimensional slice (Figure 11.5). Assume that the radiation is generated in a parallel beam inclined at an angle θ to the x axis, and let (s, t) be the coordinates in the direction of, and perpendicular to, the radiation. The total absorption in that direction, $A_\theta(t)$, is found by projecting along the s direction:

$$\begin{aligned} A_\theta(t) &= \int_{-\infty}^{\infty} \rho(s,t)\, ds \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(s,t')\, \delta(t'-t)\, ds\, dt' \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(x,y)\, \delta(x\cos\theta + y\sin\theta - t)\, dx\, dy \quad . \end{aligned} \tag{11.16}$$

This is called the *Radon transform* of the distribution [Radon, 1917]; it is a parallel projection onto the t axis.

Now consider the spatial Fourier transform of the absorption density:

$$R(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(x,y)\, e^{-i2\pi(ux+vy)}\, dx\, dy \quad . \tag{11.17}$$

Expressing the reciprocal space variables u and v in polar coordinates r and θ (but not changing the integral to polar coordinates),

$$\begin{aligned} R(r\cos\theta, r\sin\theta) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(x,y)\, e^{-i2\pi r(x\cos\theta + y\sin\theta)}\, dx\, dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(x,y) \times \end{aligned}$$

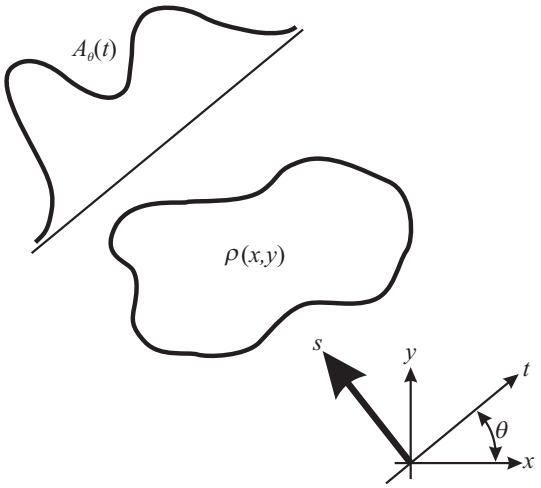

Figure 11.5. The geometry for tomographic reconstruction.

$$\left[\int_{-\infty}^{\infty} e^{-i2\pi rt}\delta(x\cos\theta + y\sin\theta - t)\,dt\right]\,dx\,dy$$
$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho(x,y)\,\delta(x\cos\theta + y\sin\theta - t)\,dx\,dy\right]$$
$$\times\, e^{-i2\pi rt}\,dt$$
$$R(r\cos\theta, r\sin\theta) = \int_{-\infty}^{\infty} A_\theta(t)\, e^{-i2\pi rt}\,dt \quad . \tag{11.18}$$

The Fourier transform of the Radon transform along the projection axis is equal to the Fourier transform of the absorption density distribution expressed in polar coordinates. This is called the *Fourier Slice Theorem*. If the parallel projection is measured for many values of θ, this relationship can then be inverted to find the internal distribution. This is how a CAT scan works. A similar calculation holds for a point source that generates a fan or a cone beam instead of the parallel propagation assumed here.

11.4 MAGNETIC RESONANCE IMAGING

The techniques we've discussed so far give information about where something is, but not what it is. *Magnetic Resonance Imaging* (*MRI*) probes the environment of nuclear spins to provide remarkable chemical as well as spatial information. It is based on *Nuclear Magnetic Resonance* (*NMR*), but was renamed to sound less frightening (even though the energies involved are only a factor of $\sim 10^{13}$ or so less than what is required for a nuclear reaction). To understand NMR we must introduce the basic features of spins; Chapters 12, 14 and 16 will continue this development.

Quantum mechanical magnetic particles and systems possess an intrinsic magnetic moment $\boldsymbol{\mu}$ and an angular momentum $\vec{J}$ that are related by the *gyromagnetic ratio* γ:

$$\vec{\mu} = \gamma\vec{J} \quad . \tag{11.19}$$

Such a particle acts just like a spinning top that is also a magnetic dipole, hence this property is called *spin* (even though nothing is actually spinning in the classical sense).

If a magnetic field is applied, it will exert a torque of $\vec{\tau} = \vec{\mu} \times \vec{B}$ on the magnetic moment. This must be equal to the rate of change of the angular momentum,

$$\begin{aligned} \frac{d\vec{J}}{dt} &= \vec{\tau} = \vec{\mu} \times \vec{B} \\ \frac{d\vec{\mu}}{dt} &= \vec{\mu} \times \gamma \vec{B} \quad . \end{aligned} \tag{11.20}$$

Therefore the spin precesses around the field with a frequency γB, called the *Larmor frequency*. Although this is a classical result, it matches the quantum calculation in the limit of non-interacting spins [Slichter, 1992].

The mysterious gyromagnetic ratio relating the magnetic moment and the angular momentum can be estimated semiclassically by considering a particle with a charge q traveling in an orbit of radius r with period T. The angular momentum is

$$J = mvr = m\frac{2\pi r}{T}r \quad , \tag{11.21}$$

the current is

$$I = \frac{q}{T} \quad , \tag{11.22}$$

the magnetic moment is

$$m = IA = \frac{q}{T}\pi r^2 \quad , \tag{11.23}$$

and so the gyromagnetic ratio is

$$m = \gamma J \;\Rightarrow\; \gamma = \frac{q}{2m} \quad . \tag{11.24}$$

This is independent of the radius, suggesting that the result might be more generally valid. In fact, this estimate is approximately equal to the correct formulas found from complete atomic and nuclear calculations. The inverse dependence on mass means that there will be many orders-of-magnitude difference in nuclear versus electronic resonance frequencies (Problem 12.4).

In an NMR experiment a strong magnetic field is applied, customarily taken to define the $\hat{z}$ directon. The resolution and sensitivity of the instrument will depend on the strength of this field; the largest spectrometers have static fields on the order of 20 T. In equilibrium the spins in the sample line up around this field with a thermal distribution. The fractional excess of spins aligned with the field, which is $\sim 10^{-5}$ for fields on the order of tesla, produces a magnetization in the sample. Transverse to the static field there is a coil that can apply an RF field to rotate the spins into the (x, y) plane (a *tipping pulse*). After this is turned off the magnetization will then precess around the static field, generating an RF signal back in the coil. This *Free Induction Decay* (*FID*) will continue until the spins thermalize again, on a time scale called T_1 that is typically on the order of seconds.

The signal strength will depend on the number of spins, the gyromagnetic ratio will depend on the spin species, and in a material the local magnetic field will vary due to internal fields. Therefore the RF spectrum provides detailed information about the

density and type of spins and their local chemical environments. To find out where they are, a gradient is added to the static field.

Let $f(\omega)$ be the NMR signal as a function of frequency, normalized to 1:

$$\int f(\omega)\, d\omega = 1 \quad , \tag{11.25}$$

and similarly let $\rho(\vec{r})$ be the normalized spatial distribution of the density of the spin being probed

$$\int \rho(\vec{r})\, d^3\vec{r} = 1 \quad . \tag{11.26}$$

If the gradient can be approximated as a constant slope along the z direction,

$$B(z) = B_0 + Gz \quad , \tag{11.27}$$

then the resonance frequency as a function of position is

$$\omega(z) = \gamma B = \gamma B_0 + \gamma G z \tag{11.28}$$

or

$$d\omega = \gamma G dz \quad . \tag{11.29}$$

Because of this gradient, the spin density integrated over a slice gives the corresponding component of the resonance spectrum

$$\begin{aligned} f(\omega)\, d\omega &= dz \int\int \rho(x, y, z(\omega))\, dx\, dy \\ f(\omega)\gamma G &= \int\int \rho(x, y, z(\omega))\, dx\, dy \quad . \end{aligned} \tag{11.30}$$

Therefore the Fourier transform of the FID provides a map of the spin density as a function of z, and additional gradients permit the full spin distribution to be reconstructed [Lauterbur, 1973; Stehling *et al*., 1991]. This projection of the spin density in a slice is similar to the Radon transform projection of an absoprtion coefficient along a slice. The spatial resolution depends on the strength of the gradients and the homogeneity of the field; in the best instruments this is on the order of microns.

The spatial and chemical analyses can be performed simultaneously; one of the most dramatic consequences of this has been the development of *functional MRI* (*fMRI*) [Ogawa *et al*., 1990; Kwong, 1995]. It is possible to resolve the change in the oxygenation state of hemoglobin to determine the amount of oxygen in the bloodstream, and it was found experimentally that the brain delivers more oxygen to active parts of the brain. This *Blood-Oxygen-Level-Dependent* (*BOLD*) contrast provides a means to spatially resolve the workings of the brain, and has revealed the location of not only gross motor control but also abstract thoughts as they happen.

11.5 INVERSE PROBLEMS

Each of the techniques we've covered has drawn on insight into a particular domain to use a set of measurements to estimate the sources of the signals. The most general approach

to imaging replaces these problem-specific details with a framework for working back from observations to inferences, the study of *inverse problems*.

Let $\vec{y}$ be a vector of measurements (pixel intensities, antenna waveforms, ...), let $\vec{x}$ be the desired internal state information (mass distribution, airplane location, ...), and let $\vec{y} = \vec{f}(\vec{x})$ be the *forward model* that relates them (geometrical optics, RF propagation, ...). It's tempting to simply write $\vec{x} = \vec{f}^{-1}(\vec{y})$, but there are at least two important problems with doing that: $\vec{f}$ may be a terribly complicated function, so that its inverse is not available, and $\vec{x}$ may have more elements than $\vec{y}$. For example, we might want to find a three-dimensional distribution from two-dimensional observations.

If the measurement has noise it's necessary to work with a probabilistic relationship $p(\vec{y}|\vec{x})$. This too must be inverted, because we want to find the most likely value of $\vec{x}$ given a measurement of $\vec{y}$. This can be done via Bayes' Theorem:

$$\max_{\vec{x}} p(\vec{x}|\vec{y}) = \max_{\vec{x}} \frac{p(\vec{y}|\vec{x})\ p(\vec{x})}{p(\vec{y})} \tag{11.31}$$

or

$$\begin{aligned} \max_{\vec{x}} \log p(\vec{x}|\vec{y}) &= \max_{\vec{x}} \log \frac{p(\vec{y}|\vec{x})\ p(\vec{x})}{p(\vec{y})} \\ &= \max_{\vec{x}} [\log p(\vec{y}|\vec{x}) + \log p(\vec{x}) - \log p(\vec{y})] \quad . \end{aligned} \tag{11.32}$$

$p(\vec{x}|\vec{y})$ is called the *posterior*, because it updates our knowledge of the internal state following a measurement of the observable. $p(\vec{y}|\vec{x})$ is the *likelihood*, giving the forward probability for the observations. The $p(\vec{y})$ term, the *evidence*, measures how reliable we think the data are, and has no influence on the maximization over a single data set. The middle term, called a *prior* or *regularizer*, is the important modification. It expresses our beliefs about the internal state in advance of seeing any data, such as whether the distribution should be smooth, or perhaps have sharp features.

Problems that can be written in terms of a Green's function (equation 7.45) are linear in the sources, letting the maximization be solved by matrix inversion (assuming that application of the prior is also linear). This inverse is well conditioned because of the presence of the regularizer. More generally, a nonlinear search is needed to find an acceptable solution; efficient approximate techniques have been developed to do this in high-dimensional parameter spaces [Besag *et al.*, 1995].

The maximization can be done to find the value of $\vec{x}$ at a discrete set of locations, or to find the parameters in a continuous representation of $\vec{x}$. For example, a least-squares likelihood term results from taking the logarithm of a Gaussian error model, and a *maximum entropy* prior finds the distribution that makes the weakest assumption about the data, so if we want to find $(x_1, \ldots, x_N)$ from $(y_1, \ldots, y_M)$ we must solve

$$0 = \frac{\partial}{\partial x_n} \left[-\sum_{m=1}^{M} \frac{[f_m(x_1, \ldots, x_N) - y_m]^2}{\sigma_m^2} - \lambda \sum_{n'=1}^{N} x_{n'} \log x_{n'} \right] \quad . \tag{11.33}$$

This gives a system of equations for the x_n. The measurement error is σ_m, and λ controls the essential tradeoff between trusting the data and trusting our advance beliefs. Note that the Lagrange multiplier could equivalently have been put in front of either term.

While the introduction of a prior might appear to violate the sanctity of the data, most any experimental procedure contains some kind of implicit prior even if it is not

written down, and this assumption is what makes it possible to generalize from limited measurements. In theory, the combination of a prior and a search algorithm makes it possible to deduce the most plausible signal sources from any set of measurements. In practice, salvation lies in the details of the implementation, but there are broadly-applicable lessons about search, constraint, and functional approximation that can help guide the development of entirely new approaches to imaging [Gershenfeld, 1999].

11.6 COMPUTATIONAL IMAGING

11.6.1 Confocal

Because light that enters a lens from sources away from its focal plane will blur the image, it's not possible to use a lens to see into a three-dimensional object even if light can propagate through it. But consider what happens if the illumination is focused onto a detector through a pinhole in the *confocal* geometry shown in Figure 11.6 [Minsky, 1957; Lichtman, 1994]. Now most of the light that scatters away from the focal plane is deflected away from the pinhole, so that by scanning the sample it is possible to reconstruct a three-dimensional image. The spatial resolution and chemical sensitivity can be further enhanced by using a two-photon process for illumination [Denk *et al*., 1990]. Such confocal microscopy has become a workhorse for biomedical imaging; Chapter 11 will continue with other kinds of imaging that don't require lenses.

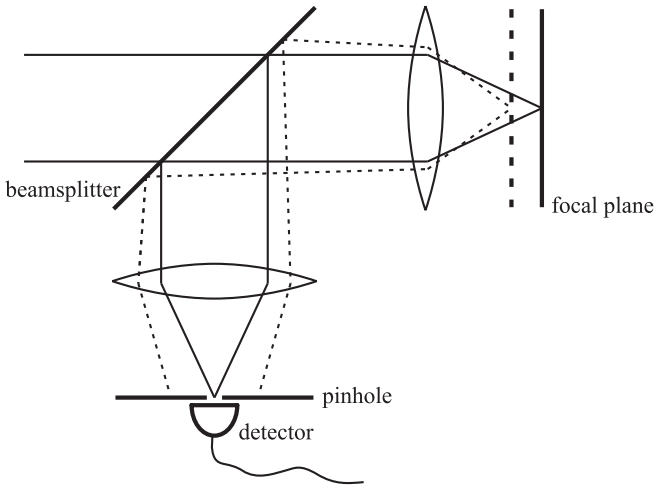

Figure 11.6. Confocal imaging.

11.6.2 Near-Field

11.6.3 Ultrafast

11.6.4 Single-Pixel

11.6.5 Light Fields

11.7 SELECTED REFERENCES

[Kino, 1987] Kino, Gordon S. (1987). *Acoustic Waves: Devices, Imaging, and Analog Signal Processing*. Englewood Cliffs: Prentice-Hall.

Advanced acoustics, including synthetic lenses, holography, and microscopy.

[Slichter, 1992] Slichter, Charles P. (1992). *Principles of Magnetic Resonance*. 3rd edn. New York: Springer-Verlag.

A great introduction to magnetic resonance (and quantum mechanics).

[Shung *et al*., 1992] Shung, K. Kirk, Smith, Michael B., & Tsui, Benjamin. (1992). *Principles of Medical Imaging*. San Diego: Academic Press.

CAT, PET, and all that.

[Press *et al*., 2007] Press, William H., Teukolsky, Saul A., Vetterling, William T., & Flannery, Brian P. (2007). *Numerical Recipes in C: The Art of Scientific Computing*. 3nd edn. Cambridge: Cambridge University Press.

Recipes has a good practical introduction to inverse problems and transforms.

11.8 Problems

(12.1) Is a thin spherical lens (like the ones we studied in the last chapter) a matched filter with a response like equation (11.8)?

(12.2) The resolving power of a lens can be defined in terms of the distance between the two points at which the point spread function has decreased from its maximum by 3 dB. Two objects can be resolved if they are separated by this distance. This is also sometimes defined by *Rayleigh's criterion*: the principal maximum of the point spread function from one object is at the first minimum of the point spread function of the other.

(*a*) For equation (11.9) what is the resolving power in terms of the wavelength, the lens aperture size, and the distance from the lens?

(*b*) For an ultrasonic signal (100 kHz) in air ($\sim$ 350 m/s), what size aperture is needed to resolve 1 cm at a distance of 1 m?

(*c*) For a radar satellite (10 GHz) in Low Earth Orbit ($\sim$ 200 km) what size aperture is needed to resolve 1 cm? What is the angle subtended by this aperture relative to the Earth's surface?

(12.3) Work out the delay function $g(x')$ to implement a matched filter in the plane for a paraxial synthetic aperture radar. Assume that the transceiver is moving along a straight line, at each point sending out a spherical wave and accumulating the return signal convolved by $g(x')$, and assume that the transceiver velocity is slow compared to the wave speed.

(12.4) Estimate the typical resonance frequency for a nuclear spin (NMR) and an electronic spin (ESR) in a 1 T field.

(12.5) This problem is harder than the others. Consider a point charge above an infinite conducting ground plane:

(*a*) Using the method of images, find the surface charge distribution on the plane.

(*b*) Assume that the plane is divided up into a grid of square electrodes, and analytically integrate the charge density to find the measured charge at each electrode.

(*c*) Numerically evaluate the electrode charge distribution generated by the point charge.

(*d*) Use these measurements to estimate the source charge distribution as a function of height above the center of the surface charge distribution. Take a least-squares likelihood function, and for a regularizer use the sum of squares of the source charges. Plot the charge distribution as the relative weight of the regularizer term is varied, showing the minimum amount of charge compatible with the measurements for a given total error.