# The Mathematics of Medical Imaging by <br> Charles L. Epstein 

(c) Charles L. Epstein, 2001-2002. All rights reserved.

## Chapter 1

## Measurements and modeling

A quantitative model of a physical system is expressed in the language of mathematics. A qualitative model often precedes a quantitative model. For many years clinicians used medical X-rays without employing a precise quantitative model. X-rays were thought of as high frequency 'light' with three very useful properties:
(1). If X-rays are incident on a human body, some fraction of the incident radiation is absorbed, though a sizable fraction is transmitted. The fraction absorbed is proportional to the total 'density' of the material encountered.
(2). A 'beam' of X-ray light travels in a straight line.
(3). X-rays darken photographic film. The opacity of the film is, in some sense, proportional to the incident energy.


Figure 1.1: The world of old fashioned X-rays.

Taken together, these properties mean that using X-rays one could "see through" a human body to obtain a shadow or projection of the internal anatomy on a sheet of film. ${ }^{1}$.

The model was adequate given the available technology. In their time, X-rays led to a revolution in the practice of medicine because they opened the door to non-destructive examination of internal anatomy. They are still useful for locating bone fractures, dental caries, and foreign objects but their ability to visualize soft tissues and more detailed anatomic structure is very limited. There are several reasons for this. An X-ray image is a two-dimensional representation of a three-dimensional object. In figure 1.1(a), the opacity of the film at a point on the film plane is inversely proportional to an average of the density of the object, measured along the line joining the point to the X-ray source. This renders it impossible to deduce the spatial ordering in the missing third dimension. Photographic film is not very sensitive to X-rays. To get a usable image, a light emitting phosphor is sandwiched with the film. This increases the sensitivity of the overall 'detector,' but even so, large changes in the intensity of the incident X-rays still produce small differences in the opacity of film. This means that the contrast between different soft tissues is poor. Because of these limitations a qualitative theory was adequate for the interpretation of X-ray images.

A desire to improve upon this situation led Alan Cormack, [10] and Godrey Hounsfield, [30] to independently develop X-ray tomography or slice imaging. The first step in their work was to use a quantitative theory for the absorption of X-rays. Such a theory already existed and is little more than a quantitative restatement of (1) and (2). It is not needed for old fashioned X-rays because they are read "by eye," no further processing is done after the film is developed. Both Cormack and Hounsfield realized that mathematics could be used to infer 3-dimensional anatomic structure from a large collection of different 2-dimensional projections. The possibility for making this idea work relied on two technological advances:
(1). The availability of scintillation crystals to use as detectors.
(2). Powerful, digital computers to process the tens of thousands of measurements needed to form a usable image.

A detector using a scintillation crystal is about a hundred times more sensitive than photographic film. Increasing the dynamic range in the basic measurements makes possible much finer distinctions. As millions of arithmetic operations are needed for each image, fast computers are a necessity for reconstructing an image from the available measurements. It is an interesting historical note that the mathematics underlying X-ray tomography was done in 1917 by Johan Radon, [59]. It had been largely forgotten and both Hounsfield and Cormack worked out solutions to the problem of reconstructing an image from its projections. Indeed, this problem had arisen and been solved in contexts as diverse as radio astronomy and statistics.

This book is a detailed exploration of the mathematics that underpins the reconstruction of images in X-ray tomography. While our emphasis is on understanding these mathematical foundations, we constantly return to the practicalities of X-ray tomography. Of particular interest is the relationship between the mathematical treatment of a problem and

[^0]the realities of numerical computation and physical measurement. There are many different imaging modalities in common use today, X-ray computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), ultrasound, optical imaging, electrical impedence imaging, etc. Because each relies on a different physical principle, each provides different information. In every case the mathematics needed to process and interpret the data has a large overlap with that used in X-ray CT. We concentrate on X-ray CT because of the simplicity and clarity of the physical principles underlying the measurement process. Detailed descriptions of the other modalities can be found in [39] or [4].

### 1.1 Mathematical modeling

Mathematics is the language in which any quantitative theory or model is eventually expressed. In this introductory chapter we consider a variety of examples of physical systems, measurement processes, and the mathematical models used to describe them. These models illustrate different aspects of more complicated models used in medical imaging.

Mathematics is used to model physical systems from the formation of the universe to the structure of the atomic nucleus, from the function of the kidney to the opinions of voters. The first step in giving a mathematical description of a "system" is to isolate that system from the universe in which it sits. While it is no doubt true that a butterfly flapping its wings in Siberia in mid-summer will effect the amount of rainfall in the Amazon rain forest a decade hence, it is surely a tiny effect, impossible to accurately quantify. To obtain a practical model such effects are ignored, though they may come back, as measurement error and noise to haunt the model. After a system is isolated, we need to find a collection of numerical parameters which describe its state. In this generality these parameters are called state variables. In the idealized world of an isolated system the exact measurement of the state variables should uniquely determine the state of the system. It may happen that the state parameters which give a convenient description of the system are not directly measurable. The mathematical model then describes relations among the state variables. Using these relations the state of the system can often be determined from feasible measurements. A simple example will clarify these very abstract sounding concepts.
Example 1.1.1. Suppose the system is a ball on a rod. For simplicity we assume the ball has radius zero. The state of the system is described by $(x, y)$, the coordinates of the ball. These are the state variables. If the rod is of length $r$ and one end of it is fixed at the point $(0,0)$, then the state variables satisfy the relation

$$
\begin{equation*}
x^{2}+y^{2}=r^{2} . \tag{1.1}
\end{equation*}
$$

Imagine now that one dimensional creatures, living on the $x$-axis $\{y=0\}$ can observe a shadow of the ball, cast by very distant light sources so that the rays of light are perpendicular to the $x$-axis. The line creatures want to predict whether or not the ball is about to collide with their world. Locating the shadow determines the $x$-coordinate of the ball and using equation (1.1) gives

$$
y= \pm \sqrt{r^{2}-x^{2}} .
$$

To determine the sign of the $y$-coordinate requires additional information not available in the model. On the other hand this information is adequate if one only wants to predict if the
ball is about to collide with the $x$-axis. If the $x$-axis is illuminated by red light from above and blue light from below, then a ball approaching from below would cast of red shadow while a ball approaching from above would cast a blue shadow. With this additional data, the location of the ball is completely determined.

Ordered pairs of real numbers, $\{(x, y)\}$ are the state variables for the system in example 1.1.1. Because of the constraint (1.1) not every pair defines a state of this system. Generally we define the state space to be values of state variables which correspond to actual states of the system. The state space in example 1.1.1 is the circle of radius $r$ centered at $(0,0)$.

## Exercises

Exercise 1.1.1. Suppose that in example 1.1.1 light sources are located at $(0, \pm R)$. What is the relationship between the $x$-coordinate and the shadow?

Exercise 1.1.2. Suppose that in example 1.1.1 the ball is tethered to $(0,0)$ by a string of length $r$. What relations do the state variables $(x, y)$ satisfy? Is there a measurement the line creatures can make to determine the location of the ball? What is the state space for this system?

Exercise 1.1.3. Suppose that the ball is untethered, but is constrained to lie in the region $\{(x, y): 0 \leq y<R\}$. Assume the points $\left\{\left(x_{1}, y_{1}\right),\left(x_{2}, y_{2}\right),\left(x_{3}, y_{3}\right)\right\}$ do not lie on a line and have $y_{j}>R$. Show that the shadows cast on the line $y=0$ by light sources located at these three points determine the location of the ball. Find a formula for $(x, y)$ in terms of the shadow locations. Why are three sources needed?

### 1.1.1 Finitely many degrees of freedom

## See: A.1, B.5, B.6, B.7.

The collection of ordered $n$-tuples of real numbers

$$
\left\{\left(x_{1}, \ldots, x_{n}\right): x_{j} \in \mathbb{R}, j=1, \ldots, n\right\}
$$

is called Euclidean $n$-space and is denoted by $\mathbb{R}^{n}$. We often use boldface letters $\mathbf{x}, \mathbf{y}$ to denote points in $\mathbb{R}^{n}$ which we sometimes call vectors. Recall that if $\mathbf{x}=\left(x_{1}, \ldots, x_{n}\right)$ and $\mathbf{y}=\left(y_{1}, \ldots, y_{n}\right)$ then their $\operatorname{sum} \mathbf{x}+\mathbf{y}$ is defined by

$$
\begin{equation*}
\mathbf{x}+\mathbf{y}=\left(x_{1}+y_{1}, \ldots, x_{n}+y_{n}\right) \tag{1.2}
\end{equation*}
$$

and if $a \in \mathbb{R}$ then $a \mathbf{x}$ is defined by

$$
\begin{equation*}
a \mathbf{x}=\left(a x_{1}, \ldots, a x_{n}\right) \tag{1.3}
\end{equation*}
$$

These two operations make $\mathbb{R}^{n}$ into a real vector space. This space is the simplest state space for a system with $n$ degrees of freedom.

If the state of a system is described by a finite collection of real numbers then the system has finite many degrees of freedom. Most of the systems encountered in elementary physics and electrical engineering have this property. Suppose that the state of a system is specified by a point $\mathbf{x} \in \mathbb{R}^{n}$ then the mathematical model is expressed as relations that these variables satisfy. These often take the form of functional relations,

$$
\begin{array}{ccc}
f_{1}\left(x_{1}, \ldots, x_{n}\right) & = & 0 \\
\vdots & \vdots  \tag{1.4}\\
f_{m}\left(x_{1}, \ldots, x_{n}\right) & = & 0 .
\end{array}
$$

The state space for the system is then the subset of $\mathbb{R}^{n}$ consisting of solutions to this system of equations. Example 1.1.1 considers a system with one degree of freedom. The state space for this system is the subset of $\mathbb{R}^{2}$ consisting of points satisfying (1.1). If the state variables satisfy constraints then this generally reduces the number of degrees of freedom.

A function $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ is linear if satisfies the conditions

$$
\begin{align*}
f(\mathbf{x}+\mathbf{y}) & =f(\mathbf{x})+f(\mathbf{y}) \text { for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{n} \text { and } \\
f(a \mathbf{x}) & =a f(\mathbf{x}) \text { for all } a \in \mathbb{R} \text { and } \mathbf{x} \in \mathbb{R}^{n} . \tag{1.5}
\end{align*}
$$

Recall that the dot or inner product is the map from $\mathbb{R}^{n} \times \mathbb{R}^{n} \rightarrow \mathbb{R}$ defined by

$$
\begin{equation*}
\langle\mathbf{x}, \mathbf{y}\rangle=\sum_{j=1}^{n} x_{j} y_{j} . \tag{1.6}
\end{equation*}
$$

Sometimes it is denoted by $\mathbf{x} \cdot \mathbf{y}$. The Euclidean length of $\mathbf{x} \in \mathbb{R}^{n}$ is defined to be

$$
\begin{equation*}
\|\mathbf{x}\|=\sqrt{\langle\mathbf{x}, \mathbf{x}\rangle}=\left[\sum_{j=1}^{n} x_{j}^{2}\right]^{\frac{1}{2}} \tag{1.7}
\end{equation*}
$$

From the definition it is easy to establish that

$$
\begin{align*}
\langle\mathbf{x}, \mathbf{y}\rangle & =\langle\mathbf{y}, \mathbf{x}\rangle \text { for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{n}, \\
\langle a \mathbf{x}, \mathbf{y}\rangle & =a\langle\mathbf{x}, \mathbf{y}\rangle \text { for all } a \in \mathbb{R} \text { and } \mathbf{x} \in \mathbb{R}^{n}, \\
\left\langle\mathbf{x}_{1}+\mathbf{x}_{2}, \mathbf{y}\right\rangle & =\left\langle\mathbf{x}_{1}, \mathbf{y}\right\rangle+\left\langle\mathbf{x}_{2}, \mathbf{y}\right\rangle \text { for all } \mathbf{x}_{1}, \mathbf{x}_{2}, \mathbf{y} \in \mathbb{R}^{n} .  \tag{1.8}\\
\|c \mathbf{x}\| & =|c|\|\mathbf{x}\| \text { for all } c \in \mathbb{R} \text { and } \mathbf{x} \in \mathbb{R}^{n} .
\end{align*}
$$

For $\mathbf{y}$ a point in $\mathbb{R}^{n}$ define the function

$$
f_{\mathbf{y}}(\mathbf{x})=\langle\mathbf{x}, \mathbf{y}\rangle
$$

The second and third relations in (1.8) show that $f_{\mathbf{y}}$ is linear. Indeed every linear function has a such a representation.

Proposition 1.1.1. If $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ is a linear function then there is a unique vector $\mathbf{y}_{f}$ such that $f(\mathbf{x})=\left\langle\mathbf{x}, \mathbf{y}_{f}\right\rangle$.

This fact is proved in exercise 1.1.5.
The inner product satisfies a basic inequality called the Cauchy-Schwarz inequality.
Proposition 1.1.2 (Cauchy-Schwarz inequality). If $\mathbf{x}, \mathrm{y} \in \mathbb{R}^{n}$ then

$$
\begin{equation*}
|\langle\mathbf{x}, \mathbf{y}\rangle| \leq\|\mathbf{x}\|\|\mathbf{y}\| . \tag{1.9}
\end{equation*}
$$

A proof of this result is outlined in exercise 1.1.6. The Cauchy-Schwarz inequality shows that if neither $\mathbf{x}$ nor $\mathbf{y}$ is zero then

$$
-1 \leq \frac{\langle\mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1
$$

this in turn allows us the define the angle between two vectors.
Definition 1.1.1. If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n}$ are both non-vanishing then the angle $\theta \in[0, \pi]$, between $\mathbf{x}$ and $\mathbf{y}$ is defined by

$$
\begin{equation*}
\cos \theta=\frac{\langle\mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} \tag{1.10}
\end{equation*}
$$

In particular two vector are orthogonal if $\langle\mathbf{x}, \mathbf{y}\rangle=0$.
The Cauchy-Schwarz inequality implies that the Euclidean length satisfies the triangle inequality.

Proposition 1.1.3. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n}$ the following inequality holds

$$
\begin{equation*}
\|\mathbf{x}+\mathbf{y}\| \leq\|\mathbf{x}\|+\|\mathbf{y}\| . \tag{1.11}
\end{equation*}
$$

Suppose that the state of a system is specified by a point in $\mathbb{R}^{n}$ subject to the constraints in (1.4). If all the functions $\left\{f_{1}, \ldots, f_{m}\right\}$ are linear then we say that this is a linear model. This is the simplest type of model and also the most common in applications. For a linear model it is a simple matter to determine the number of degrees of freedom. Suppose there is a single equation, in light of Proposition 1.1.1 it can be expressed in the form

$$
\begin{equation*}
\left\langle\mathbf{a}_{1}, \mathbf{x}\right\rangle=0, \tag{1.12}
\end{equation*}
$$

with $\mathbf{a}_{1}$ non-zero. This is the equation of a hyperplane in $\mathbb{R}^{n}$. The solutions are the vectors in $\mathbb{R}^{n}$ orthogonal to $\mathbf{a}_{1}$. Recall the following definition:

Definition 1.1.2. The vectors $\left\{\mathbf{v}_{1}, \ldots, \mathbf{v}_{k}\right\}$ are linearly independent if the only linear combination, $c_{1} \mathbf{v}_{1}+\cdots+c_{k} \mathbf{v}_{k}$ which vanishes has all its coefficients, $\left\{c_{i}\right\}$ equal to zero. Otherwise the vectors are linearly dependent.

There is a collection of $(n-1)$ linearly independent $n$-vectors $\left\{\mathbf{v}_{1}, \ldots, \mathbf{v}_{n-1}\right\}$ so that $\left\langle\mathbf{a}_{1}, \mathbf{x}\right\rangle=0$ if and only if

$$
\mathbf{x}=\sum_{i=1}^{n-1} c_{i} \mathbf{v}_{i}
$$

The system described by this equation has $n-1$ degrees of freedom.

The general case is not much harder. Suppose that the state space is the solution set of the system of linear equations

$$
\begin{array}{ccc}
\left\langle\mathbf{a}_{1}, \mathbf{x}\right\rangle & = & 0 \\
\vdots & \vdots &  \tag{1.13}\\
\left\langle\mathbf{a}_{m}, \mathbf{x}\right\rangle & = & 0 .
\end{array}
$$

Suppose that $k \leq m$ is the largest number of linear independent vectors in the collection $\left\{\mathbf{a}_{1}, \ldots, \mathbf{a}_{m}\right\}$. By renumbering we can assume that $\left\{\mathbf{a}_{1}, \ldots, \mathbf{a}_{k}\right\}$ are linearly independent and for any $l>k$, the vector $\mathbf{a}_{l}$ is a linear combination of these vectors. Hence if $\mathbf{x}$ satisfies

$$
\left\langle\mathbf{a}_{i}, \mathbf{x}\right\rangle=0 \text { for } 1 \leq i \leq k
$$

then it also satisfies $\left\langle\mathbf{a}_{l}, \mathbf{x}\right\rangle=0$ for any $l$ greater than $k$. The argument in the previous paragraph can be applied recursively to conclude that there is a collection of $n-k$ linearly independent vectors $\left\{\mathbf{u}_{1}, \ldots, \mathbf{u}_{n-k}\right\}$ so that $\mathbf{x}$ solves (1.13) if and only if

$$
\mathbf{x}=\sum_{i=1}^{n-k} c_{i} \mathbf{u}_{i}
$$

Thus the system has $n-k$-degrees of freedom.
A non-linear model can often be approximated by a linear model. If $f$ is a differentiable function then the gradient of $f$ at $\mathbf{x}$ is defined to be

$$
\nabla f(\mathbf{x})=\left(\frac{\partial f}{\partial x_{1}}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_{n}}(\mathbf{x})\right)
$$

From the definition of the derivative it follows that

$$
\begin{equation*}
f\left(\mathbf{x}_{0}+\mathbf{x}_{1}\right)=f\left(\mathbf{x}_{0}\right)+\left\langle\mathbf{x}_{1}, \nabla f\left(\mathbf{x}_{0}\right)\right\rangle+e\left(\mathbf{x}_{1}\right), \tag{1.14}
\end{equation*}
$$

where the error $e\left(\mathbf{x}_{1}\right)$ satisfies

$$
\lim _{\mathbf{x}_{1} \rightarrow 0} \frac{\left|e\left(\mathbf{x}_{1}\right)\right|}{\left\|\mathbf{x}_{1}\right\|}=0
$$

In this case we write

$$
\begin{equation*}
f\left(\mathbf{x}_{0}+\mathbf{x}_{1}\right) \approx f\left(\mathbf{x}_{0}\right)+\left\langle\mathbf{x}_{1}, \nabla f\left(\mathbf{x}_{0}\right)\right\rangle . \tag{1.15}
\end{equation*}
$$

Suppose that the functions in (1.4) are differentiable and $f_{j}\left(\mathbf{x}_{0}\right)=0$ for $j=1, \ldots, m$ then

$$
f_{j}\left(\mathbf{x}_{0}+\mathbf{x}_{1}\right) \approx\left\langle\mathbf{x}_{1}, \nabla f_{j}\left(\mathbf{x}_{0}\right)\right\rangle
$$

For small values of $\mathbf{x}_{1}$ the system of equations (1.4) can be approximated, near to $\mathbf{x}_{0}$, by a system of linear equations,

$$
\begin{array}{ccc}
\left\langle\mathbf{x}_{1}, \nabla f_{1}\left(\mathbf{x}_{0}\right)\right\rangle & = & 0 \\
\vdots & \vdots &  \tag{1.16}\\
\left\langle\mathbf{x}_{1}, \nabla f_{m}\left(\mathbf{x}_{0}\right)\right\rangle & = & 0 .
\end{array}
$$

This provides a linear model which approximates the non-linear model. The accuracy of this approximation depends in subtle way on the collection of vectors $\left\{\nabla f_{1}(\mathbf{x}), \ldots, \nabla f_{m}(\mathbf{x})\right\}$,
for $\mathbf{x}$ near to $\mathbf{x}_{0}$. The simplest situation is when these vectors are linearly independent at $\mathbf{x}_{0}$. In this case the solutions to

$$
f_{j}\left(\mathbf{x}_{0}+\mathbf{x}_{1}\right)=1, \quad j=1, \ldots, m
$$

are well approximated, for small $\mathbf{x}_{1}$, by the solutions of (1.16). This is a consequence of the Implicit Function Theorem, see [?].

Often the state variables for a system are divided into two sets, the input variables, $\left(w_{1}, \ldots, w_{k}\right)$ and output variables, $\left(z_{1}, \ldots, z_{m}\right)$, with constraints re-written in the form

$$
\begin{array}{ccc}
F_{1}\left(w_{1}, \ldots, w_{k}\right) & = & z_{1}  \tag{1.17}\\
\vdots & \vdots & \\
F_{m}\left(w_{1}, \ldots, w_{k}\right) & = & z_{m}
\end{array}
$$

The output variables are thought of as the being measured; the remaining variables must then be determined by solving this system of equations. For a linear model this amounts to solving a system of linear equations. We now consider some examples of physical systems and their mathematical models.
Example 1.1.2. We would like to find the height of a mountain without climbing it. To that end, the distance $x$ between the point $P$ and the base of the mountain, as well as the angle $\theta$ are measured, see figure 1.2. If $x$ and $\theta$ are measured exactly then the height $h$, of the mountain is given by

$$
\begin{equation*}
h(x, \theta)=x \tan \theta \tag{1.18}
\end{equation*}
$$

Measurements are never exact, using the model and elementary calculus we can relate the error in the measurement $\theta$ to the error in the computed value of $h$. Suppose that $x$ is measured exactly but there is an uncertainty $\Delta \theta$ in the value of $\theta$. Equation (1.15) gives the linear approximation

$$
h(x, \theta+\Delta \theta)-h(x, \theta) \approx \frac{\partial h}{\partial \theta}(x, \theta) \Delta \theta
$$

As $\partial_{\theta} h=x \sec ^{2} \theta$, the height, $h_{m}$ predicted from the measurement of the angle is given by

$$
h_{m}=x \tan (\theta+\Delta \theta) \approx x\left(\tan \theta+\sec ^{2} \theta \Delta \theta\right)
$$

The approximate value of the absolute error is

$$
h_{m}-h \approx x \frac{\Delta \theta}{\cos ^{2} \theta} .
$$

The absolute error is a number with the same units as $h$; in general it is not a very interesting quantity. If, for example the true measurement were $10,000 \mathrm{~m}$ then an error of size 1 m would not be too significant. If the true measurement were 2 m , then this error would be significant. To avoid this obvious pitfall one normally considers the relative error. In this problem the relative error is

$$
\frac{h_{m}-h}{h}=\frac{\Delta \theta}{\cos ^{2} \theta \tan \theta}=\frac{\Delta \theta}{\sin \theta \cos \theta} .
$$



Figure 1.2: Using to trigonometry to find the height of a mountain.
Generally the relative error is the absolute error divided by the correct value. It is a dimensionless quantity that gives a quantitative assessment of the accuracy of the measurement. If the angle $\theta$ is measured from a point too near to or too far from the mountain, i.e. $\theta$ is very close to 0 or $\pi / 2$ then small measurement errors result in a substantial loss of accuracy. A useful feature of a precise mathematical model is the possibility of estimating how errors in measurement affect the accuracy of the parameters we wish to determine. In exercise 1.1.10 we consider how to estimate the error entailed in using a linear approximation.
Example 1.1.3. In a real situation we cannot measure the distance to the base of the mountain. Suppose that we measure the angles, $\theta_{1}$ and $\theta_{2}$ from two different points, $P_{1}$ and $P_{2}$ as well as the distance $x_{2}-x_{1}$ between the two points, as shown in the figure below.


Figure 1.3: A more realistic measurement.
From the previous example we know that

$$
\begin{align*}
& h=x_{1} \tan \theta_{1},  \tag{1.19}\\
& h=x_{2} \tan \theta_{2} .
\end{align*}
$$

Using these equations and elementary trigonometry we deduce that

$$
\begin{equation*}
x_{1}=\frac{x_{2}-x_{1}}{\left[\frac{\tan \theta_{1}}{\tan \theta_{2}}-1\right]} \tag{1.20}
\end{equation*}
$$

which implies that

$$
\begin{align*}
h & =x_{1} \tan \theta_{1} \\
& =\left(x_{2}-x_{1}\right) \frac{\sin \theta_{1} \sin \theta_{2}}{\sin \left(\theta_{1}-\theta_{2}\right)} . \tag{1.21}
\end{align*}
$$

Thus $h$ can be determined from $\theta_{1}, \theta_{2}$ and $x_{2}-x_{1}$. With $d=x_{2}-x_{1}$, equation (1.21) expresses $h$ as a function of $\left(d, \theta_{1}, \theta_{2}\right)$. At the beginning of this example ( $x_{1}, \theta_{1}, x_{2}, \theta_{2}, h$ ) were the state variables describing our system. By the end these were replaced by $\left(d, \theta_{1}, \theta_{2}, h\right)$. The first three are directly measurable and the last is an explicit function of the others. The models in this and the previous example, as expressed by the equations (1.21) and (1.18) respectively, are non-linear models.

In this example there are many different ways that the model may fail to capture important features of the physical situation. We now consider a few potentials problems.
(1). If the shape of a mountain looks like that in figure 1.4 and we measure the distance and angle at the point $P$, we are certainly not finding the real height of the mountain. Some a priori information is always incorporated in a mathematical model.


Figure 1.4: Not exactly what we predicted!
(2). The curvature of the earth is ignored. A more sophisticated geometric model is needed to correct for such errors. This becomes a significant problem as soon as the distances, $x, x_{1}, x_{2}$ are large compared to the distance to the horizon (about 25 km for a 2 meter tall person). The approximations used in the model must be adapted to the actual physical conditions of the measurements.
(3). The geometry of the underlying measurements could be very different from the simple Euclidean geometry used in the model. To measure the angles $\theta_{1}, \theta_{2}$ one would normally use a transit to sight the peak of the mountain. If the mountain is far away then the light, traveling from the mountain to the transit, passes through air of varying density. The light is refracted by the air and therefore the ray path is not a straight line, as assumed in the model. To include this effect would vastly complicate the model. This is an important consideration in the very similar problem of creating a map of the sky from earth based observations of stars and planets.

Analogous problems arise in medical imaging. If the wavelength of the energy used to probe the human anatomy is very small compared to the size of the structures that are present then it is reasonable to assume that the waves are not refracted. For example X-rays can be assumed to travel along straight lines. For energies with wavelengths comparable to the size of structures present in the human anatomy, this assumption is simply wrong. The waves are then bent and diffracted by the medium and the difficulty of modeling the ray paths is considerable. This is an important issue in ultrasound imaging which remains largely unresolved.

Example 1.1.4. Refraction provides another example of a simple physical system. Suppose that we have two fluids in a tank as shown in the figure and would like to determine the height of the interface between them. Suppose first of all that the refractive indices of the fluids are known. Let $n_{1}$ be the refractive index of the upper fluid and $n_{2}$ the refractive index of the lower one, Snell's law states that

$$
\frac{\sin \left(\theta_{1}\right)}{\sin \left(\theta_{2}\right)}=\frac{n_{2}}{n_{1}}
$$

Let $h$ denote the total height of the fluid, then


Figure 1.5: Using refraction to determine the height of an interface.

$$
h_{1}+h_{2}=h
$$

The measurement we make is the total displacement $l$, of the light ray as it passes through the fluids. It satisfies the relationship

$$
h_{1} \tan \left(\theta_{1}\right)+h_{2} \tan \left(\theta_{2}\right)=l
$$

The heights $h_{1}$ and $h_{2}$ are easily determined from these three formulæ. The assumption that we know $n_{1}$ implies, by Snell's law, that we can determine $\theta_{1}$ from a measurement of the angle of the light ray above the fluid. If $n_{2}$ is also known, then using these observations we can determine $\theta_{2}$ as well:

$$
\sin \left(\theta_{2}\right)=\frac{n_{1}}{n_{2}} \sin \left(\theta_{1}\right)
$$

The pair $\left(h_{1}, h_{2}\right)$ satisfies the $2 \times 2$-linear system

$$
\left(\begin{array}{cc}
1 & 1  \tag{1.22}\\
\tan \left(\theta_{1}\right) & \tan \left(\theta_{2}\right)
\end{array}\right)\binom{h_{1}}{h_{2}}=\binom{h}{l} .
$$

In example 2.1.2 we consider a slightly more realistic situation where the refractive index of the lower fluid in not known. By using more measurements $n_{2}$ can also be determined. It is the solution of a non-linear equation.

## Exercises

Exercise 1.1.4. Prove the formulæ in (1.8).

Exercise 1.1.5. Let $\mathbf{e}_{j} \in \mathbb{R}^{n}, j=1, \ldots, n$ denote the vector with a 1 in the $j^{\text {th }}$ place and otherwise zero,

$$
\mathbf{e}_{1}=(1,0,0, \ldots, 0), \mathbf{e}_{2}=(0,1,0, \ldots, 0), \ldots, \mathbf{e}_{n}=(0, \ldots, 0,1) .
$$

(1). Show that if $\mathbf{x}=\left(x_{1}, \ldots, x_{n}\right)$ then

$$
\mathbf{x}=\sum_{j=1}^{n} x_{j} \mathbf{e}_{j}
$$

(2). Use the previous part to prove the existence statement in Proposition 1.1.1, that is, show that there is a vector $\mathbf{y}_{f}$ so that $f(\mathbf{x})=\left\langle\mathbf{x}, \mathbf{y}_{f}\right\rangle$. Give a formula for $\mathbf{y}_{f}$.
(3). Show that the uniqueness part of the proposition is equivalent to the statement: "If $\mathbf{y} \in \mathbb{R}^{n}$ satisfies

$$
\langle\mathbf{x}, \mathbf{y}\rangle=0 \text { for all } \mathbf{x} \in \mathbb{R}^{n}
$$

then $\mathbf{y}=0 . "$ Prove this statement.
Exercise 1.1.6. In this exercise we use calculus to prove the Cauchy-Schwarz inequality. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n}$ be non-zero vectors, define the function

$$
F(t)=\langle\mathbf{x}+t \mathbf{y}, \mathbf{x}+t \mathbf{y}\rangle
$$

Use calculus to find the value of $t$ where $F$ assumes its minimum value. By using the fact that $F(t) \geq 0$ for all $t$ deduce the Cauchy-Schwarz inequality.

Exercise 1.1.7. Show that (1.11) is a consequence of the Cauchy-Schwarz inequality. Hint: Consider $\|\mathbf{x}+\mathbf{y}\|^{2}$.
Exercise 1.1.8. Let a be a non-zero $n$-vector. Show that there is a collection of $n-1$ linearly independent $n$-vectors, $\left\{\mathbf{v}_{1}, \ldots, \mathbf{v}_{n-1}\right\}$ so that $\mathbf{x}$ solves $\langle\mathbf{a}, \mathbf{x}\rangle=0$ if and only if

$$
\mathbf{x}=\sum_{i=1}^{n-1} c_{i} \mathbf{v}_{i}
$$

for some real constants $\left\{c_{1}, \ldots, c_{n-1}\right\}$.
Exercise 1.1.9. Let $\left\{\mathbf{a}_{1}, \ldots, \mathbf{a}_{k}\right\}$ be linearly independent $n$-vectors. Show that there is a collection $n-k$ linearly independent $n$-vectors, $\left\{\mathbf{v}_{1}, \ldots, \mathbf{v}_{n-k}\right\}$ so that $\mathbf{x}$ solves

$$
\left\langle\mathbf{a}_{j}, \mathbf{x}\right\rangle=0 \text { for } j=1, \ldots, k
$$

if and only if

$$
\mathbf{x}=\sum_{i=1}^{n-k} c_{i} \mathbf{v}_{i}
$$

for some real constants $\left\{c_{1}, \ldots, c_{n-k}\right\}$. Hint: Use the previous exercise and an induction argument.

Exercise 1.1.10. If a function $f$ has two derivatives then Taylor's theorem gives a formula for the error $e(y)=f(x+y)-\left[f(x)+f^{\prime}(x) y\right]$. There exists a $z$ between 0 and $y$, such that

$$
e(z)=\frac{f^{\prime \prime}(z) y^{2}}{2}
$$

see (B.11). Use this formula to bound the error made in replacing $h(x, \theta+\Delta \theta)$ with $h(x, \theta)+\partial_{\theta} h(x, \theta) \Delta \theta$. Hint: Find the value of $z$ between 0 and $\Delta \theta$ which maximizes the error term.

Exercise 1.1.11. In example 1.1.3 compute the gradient of $h$ to determine how the absolute and relative errors depend on $\theta_{1}, \theta_{2}$ and $d$.

### 1.1.2 Infinitely many degrees of freedom

```
See: A.3, A.6.
```

In the previous section we examined some simple physical systems with finitely many degrees of freedom. In these examples, the problem of determining the state of the system from feasible measurements reduces to solving systems of finitely many equations in finitely many unknowns. In imaging applications the state of a system is usually described by a function or functions of continuous variables. These systems have infinitely many degrees of freedom. In this section we consider several examples.
Example 1.1.5. * Suppose that we would like to determine the shape of a planar region, $D$ that cannot be seen. The object is lying inside a disk and we can fire particles at the object. Assume that the particles bounce off according to a simple scattering process. Each particle strikes the object once and is then scattered along a straight line off to infinity. The outline of the object can be determined by knowing the correspondence between incoming lines, $l_{\text {in }}$ and outgoing lines, $l_{\text {out }}$. Each intersection point $l_{\text {in }} \cap l_{\text {out }}$ lies on the boundary of the object. Measuring $\left\{l_{\text {out }}^{j}\right\}$ for finitely many incoming directions $\left\{l_{\text {in }}^{j}\right\}$ determines finitely many points $\left\{l_{\text {in }}^{j} \cap l_{\text {out }}^{j}\right\}$ on the boundary of $D$. In order to use this finite collection of points to make any assertions about the rest of the boundary of $D$, more information is required. If we know that $D$ consists of a single piece or component then these points would lie on a single closed curve, though it might be difficult to decide in what order they should appear on the curve. On the other hand, these measurements provide a lot of information about convex regions.

Definition 1.1.3. A region $D$ in the plane is convex if it has the following property: for each pair of points $p$ and $q$ lying in $D$ the line segment $\overline{p q}$ is also contained in $D$.

Convex regions have many special properties. If $p$ and $q$ are on the boundary of $D$ then the line segment $\overline{p q}$ lies inside of $D$. From this observation one can show that if $\left\{p_{1}, \ldots, p_{N}\right\}$ are points on the boundary of a convex region then the smallest polygon with these points as vertices lies entirely within $D$, see figure $1.7(\mathrm{a})$. Convexity can also be defined by a property of the boundary of $D$ : For each point $p$ on the boundary of $D$ there is a line $l_{p}$ which passes through $p$ but is otherwise disjoint from $D$. This line is called a support line


Figure 1.6: Convex and non-convex regions.
through $p$. If the boundary is smooth at $p$ then the tangent line to the boundary is the unique support line. A line divides the plane into two half planes. Let $l_{p}$ be a support line to $D$ at $p$. Since $D$ does not meet $p$ it must lies entirely in one of the half planes determined by $l_{p}$, see figure 1.7.(b). If each support line meets the boundary of $D$ at exactly one point then the region is strictly convex.


Figure 1.7: Further properties of convex regions.
Suppose that the object is convex and more is known about the scattering process, for example that the angle of incidence is equal to the angle of reflection. From a finite number of incoming and outgoing pairs, $\left\{\left(l_{\text {in }}^{i}, l_{\text {out }}^{i}\right): i=1, \ldots, N\right\}$ we can now determine an approximation to $D$ with an estimate for the error. The intersection points, $p_{i}=l_{\mathrm{in}}^{i} \cap l_{\text {out }}^{i}$ lie on the boundary of the convex region, $D$. If we use these points as the vertices of a
polygon, $P_{N}^{\text {in }}$ then as remarked above, $P_{N}^{\text {in }}$ is completely contained within $D$. On the other hand, as the angle of incidence equals the angle of reflection we can also determine the tangent lines $\left\{l_{p_{i}}\right\}$ to the boundary of $D$ at the points $\left\{p_{i}\right\}$. These lines are support lines for $D$. Hence by intersecting the half planes which contain $D$, defined by these tangent lines, we obtain another convex polygon, $P_{N}^{\text {out }}$ which contains $D$. Thus with these $N$-measurements we obtain the both an inner and outer approximation to $D$ :

$$
P_{N}^{\mathrm{in}} \subset D \subset P_{N}^{\mathrm{out}}
$$

A convex region is determined by its boundary and each point on the boundary is, in effect a state variable. Therefore the collection of convex regions is a system with infinitely many degrees of freedom. A nice description for the state space of convex regions is developed in section 1.2 .2 . As we have seen, a convex region can be approximated by polygons. Once the number of sides is fixed, then we are again considering a system with finitely many degrees of freedom. In all practical problems, a system with infinitely many degrees of freedom must eventually be approximated by a system with finitely many degrees of freedom.

(a) The angle of incidence equals the angle of reflection.

(b) The outer approximation as an intersection of half spaces.

Figure 1.8: Using particle scattering to determine the boundary of a convex region.

Remark 1.1.1. For a non-convex body the above method does not work as the correspondence between incoming and outgoing lines can be quite complicated: some incoming lines may undergo multiple reflections before escaping, in fact some lines might become permanently trapped.
Example 1.1.6. Suppose that the surface of a sea is mapped by coordinates $(x, y)$ belonging to a region $D \subset \mathbb{R}^{2}$. The depth of the bottom of the sea is described by a function $h(x, y)$. One way to determine $h$ would be to drop a weighted string until it hits the bottom. There are problems with this method: 1. It is difficult to tell when the weight hits the bottom. 2 . Unknown underwater currents may carry the string so that it does not go straight down. A somewhat less direct approach would be to use sonar to measure the distance to the
bottom. The physical principle underlying the measurement is that the speed of sound is determined by the density and temperature of the water which are in turn determined by the depth. Let $c(z)$ denote the known speed of sound, as a function of the depth. A speaker underneath the boat emits a loud, short pulse of sound and the time it takes for the sound to return is measured. Here we assume that the sound travels in a straight line to the bottom and the microphone only detects the direct reflection, traveling back along the straight line. Using $c(z)$ the transit time can be related to the depth.


Figure 1.9: Using sound to measure depth.
A simple model, valid for shallow seas, is that the speed of sound is a constant, $c$. The measurement, $T$ is the time it takes for the sound pulse to go down and back,

$$
2 h=c T .
$$

This assumes that the boat is stationary from the time the pulse is emitted until the return is received. If $T(x, y)$ is the return time at $(x, y) \in D$ then the depth $h(x, y)$ is given by

$$
h(x, y)=\frac{c T(x, y)}{2} .
$$

In reality such continuous measurements are not possible. Instead the boat is placed at a finite set of locations $P=\left\{\left(x_{j}, y_{j}\right): j=1, \ldots, N\right\}$ and $T\left(x_{j}, y_{j}\right)$ is measured. These measurements determine the finite set of values

$$
h\left(x_{j}, y_{j}\right)=\frac{c T\left(x_{j}, y_{j}\right)}{2}, \quad j=1, \ldots, N .
$$

Again, what use is a finite set of values? Without qualitative, a priori information about the nature of the function $h$, this finite data set is indeed useless! On the other hand it is reasonable to assume that $h$ is a continuous function of $(x, y)$. With this assumption, values of $h$ for points not in $P$ can be interpolated from the measured values. The minimum necessary separation between the points in $P$ is determined by a quantitative assessment of how continuous $h$ is expected to be. Suppose it is known that there is a constant $M$ so that

$$
\left|h(x, y)-h\left(x^{\prime}, y^{\prime}\right)\right| \leq M \sqrt{\left(x-x^{\prime}\right)^{2}+\left(y-y^{\prime}\right)^{2}} .
$$

If every point $(x, y)$ is within $d$ of a point $\left(x_{j}, y_{j}\right)$ in $P$ then we have the estimate

$$
\left|h(x, y)-h\left(x_{j}, y_{j}\right)\right| \leq M d
$$

This then gives an estimate for the accuracy of the interpolated values. A small value of $M$ indicates that the depth is varying slowly, while a large value indicates rapid variations. In the former case a larger value of $d$ provides acceptable results, while in the latter case a smaller value of $d$ is needed to get an accurate picture of the bottom.
Example 1.1.7. Now assume that the sea, in the previous example, is one-dimensional, but that the sound speed is not constant. To use the measurements described above to determine the depth $h(x)$ requires more mathematical apparatus. Let $z(t)$ denote the depth of the sound pulse at a time $t$ after it is emitted. Using calculus we can express the assertion that the 'speed of sound at depth $z$ is $c(z)$ ' as a differential equation

$$
\begin{equation*}
\frac{d z}{d t}(t)=c(z(t)) \tag{1.23}
\end{equation*}
$$

Formally this is equivalent to

$$
\frac{d z}{c(z)}=d t
$$

The integral of $1 / c(z)$ defines a function,

$$
\begin{equation*}
G(h) \stackrel{d}{=} \int_{0}^{h} \frac{d z}{c(z)} \tag{1.24}
\end{equation*}
$$

From the derivation of (1.23) it follows that the transit time, $T(h)$ for a pulse to reach depth $h$ and return is $2 G(h)$.

The function $G$ is increasing and therefore its inverse is well defined. Using $G^{-1}$ we can determine the depth, $h$ from the available measurement, $T$

$$
h=G^{-1}\left(\frac{T}{2}\right)
$$

To use this model, the function $G^{-1}$ needs to be explicitly determined. If $c(z)$ is simple enough then an analytic formula for $G$ might be available. Otherwise the integral defining $G$ is computed for a finite collection of depths $\left\{h_{1}, \ldots, h_{m}\right\}$, with $t_{i}=G\left(h_{i}\right)$. From this table of values, the inverse function is also known for a finite collection of times

$$
h_{i}=G^{-1}\left(t_{i}\right)
$$

If $c(z)$ is a differentiable function, then a linear approximation of the form

$$
c(z) \approx a+b z
$$

is valid for small values of $z$. Integrating gives

$$
G(h) \approx \log \left(1+\frac{b h}{a}\right)
$$

solving for $G^{-1}(T)$ we find

$$
h(T) \approx \frac{a}{b}\left(e^{b \frac{T}{2}}-1\right)
$$

Using Taylor's formula, (B.5) for $e^{x}$ gives

$$
h(T) \approx a \frac{T}{2}+\frac{b a T^{2}}{8}+O\left(T^{3}\right) .
$$

Here as usual $O\left(T^{3}\right)$ is an error term which goes to zero, as $T$ goes to zero, at the same rate as $T^{3}$. This agrees, to leading order with the previous computation.
Example 1.1.8. The one-dimensional model in the previous example can be used to solve the two-dimensional problem. Suppose that the area we are interested in mapping corresponds to the rectangle $[-1,1] \times[-1,1]$ in the $(x, y)$-map coordinates. For each $y$ define the function of one variable

$$
h_{y}(x) \stackrel{d}{=} h(x, y) .
$$

Knowing the collection of functions $\left\{h_{y}: y \in[-1,1]\right\}$ for $x \in[-1,1]$ is evidently exactly the same thing as a knowing $h$, for $(x, y) \in[-1,1] \times[-1,1]$. Because the measuring apparatus observes only the sound returning on the straight line from the boat to the bottom of the sea, the analysis in the previous example applies to allow the determination of $h_{y}(x)$ from a measurement of $T_{y}(x)$,

$$
h(x, y)=h_{y}(x)=G^{-1}\left(\frac{T_{y}(x)}{2}\right) .
$$

In this way a two-dimensional problem is sliced into simpler one dimensional problems. In real applications, only finitely many measurements are made. A typical strategy if to pick an equally spaced set of $y$-values,

$$
y_{k}=\frac{k}{N}, \quad k=-N, \ldots, N
$$

and determine $h_{y_{k}}\left(x_{j}\right)$ at finitely many equally spaced $x$-values

$$
x_{j}=\frac{j}{N} \quad j=-N, \ldots, N .
$$

These examples capture many of the features that we will encounter in X-ray tomography: by using a mathematical model for the measurements, an inaccessible, physical quantity can be determined using feasible measurements. The model is itself an approximation, but is subject to improvements.

## Exercises

Exercise 1.1.12. Find state variables to describe the set of polygons with $n$-vertices in the plane. For the case of triangles, find the relations satisfied by your variables. Extra credit: Find a condition, in terms of your parameters implying that the polygon is convex.
Exercise 1.1.13. Suppose that $D_{1}$ and $D_{2}$ are convex regions in the plane. Show that their intersection $D_{1} \cap D_{2}$ is also a convex region.

Exercise 1.1.14. * Suppose that $D$ is a possibly non-convex region in the plane. Define a new region $D^{\prime}$ as the intersection of all the half spaces which contain $D$. Show that $D=D^{\prime}$ if and only if $D$ is convex.

Exercise 1.1.15. Find an example of a planar region such that at least one particle trajectory is trapped forever.

Exercise 1.1.16. Why is $G$ in example 1.1.7 a monotonely increasing function?
Exercise 1.1.17. Suppose that $c(z)$ is piecewise constant, so that

$$
c(z)= \begin{cases}c_{1} & \text { if } 0 \leq z \leq z_{1}, \\ c_{2} & \text { if } z_{1}<z .\end{cases}
$$

Find $G$ and $G^{-1}$.
Exercise 1.1.18. In example 1.1 .7 why is it reasonable to model $c(z)$ as a linear function under the assumption that it is a differentiable function? Suggest a method for determining $b$.

Exercise 1.1.19. In examples 1.1.6-1.1.8 it is assumed that all returns not arriving on the straight line path from the bottom of the ocean are ignored. Analyze the problems that result if return signals are accepted from all directions. What impact would this have on using the slicing method to reduce the dimensionality of the problem?

Exercise 1.1.20. Repeat the analysis in example 1.1.7 assuming that the boat is traveling at constant velocity $v$. Continue assuming that only returns meeting the bottom of the boat at right angles are detected.

### 1.2 A simple model problem for image reconstruction

The problem of image reconstruction in X-ray tomography is sometimes described as reconstructing an object from its "projections." Of course these are projections under the illumination of X-ray "light." In this section we consider the analogous, but simpler problem, of determining the outline of an convex object from its shadows. As is also the case in medical applications, we consider a two-dimensional problem. Let $D$ be a convex region in the plane. Imagine that a light source is placed very far from $D$. Since the light source is very far away, the rays of light are all traveling in essentially the same direction. We can think of them as a collection of parallel lines. We want to measure the shadow that $D$ casts for each position of the light source. To describe the measurements imagine that a screen is placed on the "other side" of $D$ perpendicular to the direction of the light rays, see the figure below. In a real apparatus sensors would be placed on the screen, allowing us to determine where the shadow begins and ends.


Figure 1.10: The shadow of a convex region
The region, $D$ blocks a certain collection of light rays and allows the rest to pass. Locating the shadow amounts to determining the "first" and "last" lines in this family of parallel lines to intersect $D$. To completely describe the object we need to rotate the source and detector through $\pi$ radians, measuring, at each angle, where the shadow begins and ends.

The first and last lines to intersect a region just meet it along its boundary. These lines are therefore tangent to the boundary of $D$. The problem of reconstructing a region from its shadows is mathematically the same as the problem of reconstructing a region from a knowledge of the tangent lines to its boundary. As a first step in this direction we need a good way to organize our measurements. To that end we give a description for the space of lines in the plane.

### 1.2.1 The space of lines in the plane ${ }^{\star}$

A line in the plane is a set of points which satisfies an equation of the form

$$
a x+b y=c
$$

where $a^{2}+b^{2} \neq 0$. We could use $(a, b, c)$ to parametrize the set of lines, but note that we get the same set of points if we replace this equation by

$$
\frac{a}{\sqrt{a^{2}+b^{2}}} x+\frac{b}{\sqrt{a^{2}+b^{2}}} y=\frac{c}{\sqrt{a^{2}+b^{2}}} .
$$

The coefficients, $\left(\frac{a}{\sqrt{a^{2}+b^{2}}}, \frac{b}{\sqrt{a^{2}+b^{2}}}\right.$ ) define a point $\omega$ on the unit circle, $S^{1} \subset \mathbb{R}^{2}$, and the constant $\frac{c}{\sqrt{a^{2}+b^{2}}}$ can be any number. The lines in the plane are parametrized by a pair consisting of a unit vector $\omega=\left(\omega_{1}, \omega_{2}\right)$ and a real number $t$. The line $l_{t, \omega}$ is the set of points satisfying the equation

$$
\langle(x, y), \omega\rangle=t .
$$

Very often it is convenient to parametrize the points on the unit circle by a real number, to that end we set

$$
\begin{equation*}
\omega(\theta)=(\cos (\theta), \sin (\theta)) \tag{1.25}
\end{equation*}
$$

Since cos and sin are $2 \pi$-periodic it clear that $\omega(\theta)$ and $\omega(\theta+2 \pi)$ are the same point on the unit circle. Using this notation the line $l_{t, \theta}=l_{t, \omega(\theta)}$ is the set of solutions to the equation

$$
\langle(x, y),(\cos (\theta), \sin (\theta))\rangle=t .
$$

Both notations are used in the sequel.
While the parameterization provided by $(t, \omega)$ is much more efficient than that provided by $(a, b, c)$ note that the set of points satisfying this equation is unchanged if $(t, \omega)$ is replaced by $(-t,-\omega)$. Thus, as sets,

$$
\begin{equation*}
l_{t, \omega}=l_{-t,-\omega} . \tag{1.26}
\end{equation*}
$$

It is not difficult to show that these are the only pairs $(t, \omega),\left(t^{\prime}, \omega^{\prime}\right)$ for which $l_{t, \omega}=l_{t^{\prime}, \omega^{\prime}}$. The vector

$$
\hat{\omega}=\left(-\omega_{2}, \omega_{1}\right),
$$

is perpendicular to $\omega$. For any real number $s$,

$$
\langle\omega,(t \omega+s \hat{\omega})\rangle=t
$$

and therefore we can describe $l_{t, \omega}$ parametrically as the set of points

$$
l_{t, \omega}=\{t \omega+s \hat{\omega} \mid s \in(-\infty, \infty)\} .
$$

Both $\hat{\omega}$ and $-\hat{\omega}$ are unit vectors which are perpendicular to $\omega$; $\hat{\omega}$ is singled out by the condition that the $2 \times 2$ matrix,

$$
\left(\begin{array}{cc}
\omega_{1} & -\omega_{2} \\
\omega_{2} & \omega_{1}
\end{array}\right),
$$

has determinant +1 . The vector $\hat{\omega}$ defines the "positive" direction or orientation along the line $l_{t, \omega}$. This shows that the pair $(t, \omega)$ determines an oriented line. We summarize these computations in a proposition.
Proposition 1.2.1. The pairs $(t, \omega) \in \mathbb{R} \times S^{1}$ are in one-to-one correspondence with the set of oriented lines in the plane.


Figure 1.11: Parameterization of oriented lines in the plane.
The vector $\omega$ is the direction orthogonal to the line and the number $t$ is called the affine parameter of the line, $|t|$ is the distance from the line to the origin of the coordinate system. The pair $(t, \omega)$ defines two half planes

$$
\begin{equation*}
H_{t, \omega}^{+}=\left\{\mathbf{x} \in \mathbb{R}^{2} \mid\langle\mathbf{x}, \omega\rangle>t\right\} \text { and } H_{t, \omega}^{-}=\left\{\mathbf{x} \in \mathbb{R}^{2} \mid\langle\mathbf{x}, \omega\rangle<t\right\}, \tag{1.27}
\end{equation*}
$$

the line $l_{t, \omega}$ is the common boundary of these half planes. Facing along the line $l_{t, \omega}$ in the direction specified by $\hat{\omega}$, the half plane $H_{t, \omega}^{-}$lies to the left.

## Exercises

Exercise 1.2.1. Suppose that $(t, \omega)$ and $\left(t^{\prime}, \omega^{\prime}\right)$ are different points in $\mathbb{R} \times S^{1}$ such that $l_{t, \omega}=l_{t^{\prime}, \omega^{\prime}}$. Show that $(t, \omega)=\left(-t^{\prime},-\omega^{\prime}\right)$.
Exercise 1.2.2. Show that

$$
|t|=\min \left\{\sqrt{x^{2}+y^{2}}:(x, y) \in l_{t, \omega}\right\}
$$

Exercise 1.2.3. Show that if $\omega$ is fixed then the lines in the family $\left\{l_{t, \omega}: t \in \mathbb{R}\right\}$ are parallel.

Exercise 1.2.4. Show that every line in the family $\left\{l_{t, \hat{\omega}}: t \in \mathbb{R}\right\}$ is orthogonal to every line in the family $\left\{l_{t, \omega}: t \in \mathbb{R}\right\}$.

Exercise 1.2.5. Each choice of direction $\omega$ defines a coordinate system on $\mathbb{R}^{2}$,

$$
(x, y)=t \omega+s \hat{\omega} .
$$

Find the inverse, expressing $(t, s)$ as functions of $(x, y)$. Show that the area element in the plane satisfies

$$
d x d y=d t d s
$$

### 1.2.2 Reconstructing an object from its shadows

Now we can quantitatively describe the shadow. Because there are two lines in each family of parallel lines which are tangent to the boundary of $D$ we need a way to select one of them. To do this we choose an orientation for the boundary of $D$; this operation is familiar from Green's theorem in the plane. The positive direction on the boundary is selected so that, when facing in this direction the region lies to the left; the counterclockwise direction is, by convention the positive direction, see figure 1.12.

Fix a source position $\omega(\theta)$; in the family of parallel lines $\left\{l_{t, \omega(\theta)}: t \in \mathbb{R}\right\}$ there are two values of $t, t_{0}<t_{1}$ such that the lines $l_{t_{0}, \omega(\theta)}$ and $l_{t_{1}, \omega(\theta)}$ are tangent to the boundary of $D$, see figure 1.12. Examining the diagram it is clear that the orientation of the boundary at the point of tangency and that of the oriented line agree for $t_{1}$, and are opposite for $t_{0}$. Define $h_{D}$, the shadow function of $D$ by setting

$$
\begin{equation*}
h_{D}(\theta)=t_{1} \text { and } h_{D}(\theta+\pi)=-t_{0} \tag{1.28}
\end{equation*}
$$

The shadow function is completely determined by source positions covering half the circle, or by values of $\theta$ belonging to an interval of length $\pi$. Because $\omega(\theta)=\omega(\theta+2 \pi)$ the shadow function is also regarded as a $2 \pi$-periodic function defined on the whole real line. The mathematical formulation of reconstruction problem is: Can the boundary of the region $D$ be determined from $h_{D}$ ?

The line $l_{h_{D}(\theta), \omega(\theta)}$ is given parametrically by

$$
\left\{h_{D}(\theta)(\cos (\theta), \sin (\theta))+s(-\sin (\theta), \cos (\theta)) \mid s \in(-\infty, \infty)\right\} .
$$



Figure 1.12: The measurement of the shadow

To determine the boundary of $D$ it would suffice to determine the point of tangency of $l_{h_{D}(\theta), \omega(\theta)}$ with the boundary of $D$, in other words we would like to find the function $s(\theta)$ so that for each $\theta$,

$$
\begin{equation*}
(x(\theta), y(\theta))=h_{D}(\theta)(\cos (\theta), \sin (\theta))+s(\theta)(-\sin (\theta), \cos (\theta)) \tag{1.29}
\end{equation*}
$$

is a point on the boundary of $D$.
The function $s(\theta)$ is found by recalling that, at the point of tangency, the direction of the tangent line to $D$ is $\hat{\omega}(\theta)$. For a curve in the plane given parametrically by $(x(\theta), y(\theta))$ the direction of the tangent line at a point $\theta_{0}$ is the same as that of the vector $\left(x^{\prime}\left(\theta_{0}\right), y^{\prime}\left(\theta_{0}\right)\right)$. Differentiating the expression given in (1.29) and using the fact that $\partial_{\theta} \omega=\hat{\omega}$ we find that

$$
\begin{equation*}
\left(x^{\prime}(\theta), y^{\prime}(\theta)\right)=\left(h_{D}^{\prime}(\theta)-s(\theta)\right) \omega(\theta)+\left(h_{D}(\theta)+s^{\prime}(\theta)\right) \hat{\omega}(\theta) . \tag{1.30}
\end{equation*}
$$

Since the tangent line at $(x(\theta), y(\theta))$ is parallel to $\hat{\omega}(\theta)$ it follows from (1.30) that

$$
\begin{equation*}
h_{D}^{\prime}(\theta)-s(\theta)=0 . \tag{1.31}
\end{equation*}
$$

This gives a parametric representation for the boundary of a convex region in terms of its shadow function: If the shadow function is $h_{D}(\theta)$ then the boundary of $D$ is given parametrically by

$$
\begin{equation*}
(x(\theta), y(\theta))=h_{D}(\theta) \omega(\theta)+h_{D}^{\prime}(\theta) \hat{\omega}(\theta) . \tag{1.32}
\end{equation*}
$$

Note that we have assumed that $D$ is strictly convex and the $h_{D}(\theta)$ is a differentiable function. This is not always true, for example if the region $D$ is a polygon then neither assumption holds.

Let $D$ denote a convex region and $h_{D}$ its shadow function. We can think of $D \mapsto h_{D}$ as a mapping from convex regions in the plane to $2 \pi$-periodic functions. It is reasonable to enquire if every $2 \pi$-periodic function is the shadow function of a convex region. The answer to this question is no. For strictly convex regions with smooth boundaries we are able to
characterize the range of this mapping. If $h$ is twice differentiable then the tangent vector to the curve defined by

$$
\begin{equation*}
(x(\theta), y(\theta))=h(\theta) \omega(\theta)+h^{\prime}(\theta) \hat{\omega}(\theta) \tag{1.33}
\end{equation*}
$$

is given by

$$
\left(x^{\prime}(\theta), y^{\prime}(\theta)\right)=\left(h^{\prime \prime}(\theta)+h(\theta)\right) \hat{\omega}(\theta)
$$

In our construction of the shadow function we observed that the tangent vector to the curve at $(x(\theta), y(\theta))$ and the vector $\hat{\omega}(\theta)$ point in the same direction. From our formula for the tangent vector we see that this implies that

$$
\begin{equation*}
h^{\prime \prime}(\theta)+h(\theta)>0 \text { for all } \theta \in[0,2 \pi] . \tag{1.34}
\end{equation*}
$$

This gives a necessary condition for a twice differentiable function $h$ to be the shadow function for a strictly convex region with a smooth boundary. Mathematically we are determining the range of the map that takes a convex body $D \subset \mathbb{R}^{2}$ to its shadow function $h_{D}$, under the assumption that $h_{D}$ is twice differentiable. This is a convenient mathematical assumption, though in an applied context it is likely to be overly restrictive.

## Exercises

Exercise 1.2.6. Justify the definition of $h_{D}(\theta+\pi)$ in (1.28) by showing that the orientation of the boundary at the point of tangency with $l_{t_{0}, \omega(\theta)}$ agrees with that of $l_{-t_{0}, \omega(\theta+\pi)}$.
Exercise 1.2.7. Suppose that $D_{n}$ is a regular $n$-gon. Determine the shadow function $h_{D_{h}}(\theta)$.

Exercise 1.2.8. Suppose that $h$ is a $2 \pi$-periodic, twice differentiable function that satisfies (1.34). Show that the curve given by (1.33) is the boundary of a strictly convex region.

Exercise 1.2.9. How is the assumption that $D$ is strictly convex used in the derivation of (1.31)?

Exercise 1.2.10. If $h$ is a differentiable function then equation (1.33) defines a curve. By plotting examples, determine what happens if the condition (1.34) is not satisfied.

Exercise 1.2.11. Suppose that $h$ is a function satisfying (1.34). Show that the area of $D_{h}$ is given by the

$$
\operatorname{Area}\left(D_{h}\right)=\frac{1}{2} \int_{0}^{2 \pi}\left[(h(\theta))^{2}-\left(h^{\prime}(\theta)\right)^{2}\right] d \theta
$$

Explain why this implies that a function satisfying (1.34) also satisfies the estimate

$$
\int_{0}^{2 \pi}\left(h^{\prime}(\theta)\right)^{2} d \theta<\int_{0}^{2 \pi}(h(\theta))^{2} d \theta
$$

Exercise 1.2.12. Let $h$ be a smooth $2 \pi$-periodic function which satisfies (1.34). Prove that the curvature of the boundary of the region with this shadow function, at the point $h(\theta) \omega(\theta)+h^{\prime}(\theta) \hat{\omega}(\theta)$ is given by

$$
\begin{equation*}
\kappa(\theta)=\frac{1}{h(\theta)+h^{\prime \prime}(\theta)} . \tag{1.35}
\end{equation*}
$$

Exercise 1.2.13. Suppose that $h$ is a function satisfying (1.34). Show that another parametric representation for the boundary of the region with this shadow function is

$$
\theta \mapsto\left(-\int_{0}^{\theta}\left(h(s)+h^{\prime \prime}(s)\right) \sin (s) d s, \int_{0}^{\theta}\left(h(s)+h^{\prime \prime}(s)\right) \cos (s) d s\right) .
$$

Exercise 1.2.14. In this exercise we determine which positive functions $\kappa$ defined on $S^{1}$ are the curvatures of closed strictly convex curves. Prove the following result: A positive function $\kappa$ on $S^{1}$ is the curvature of a closed, strictly convex curve (parametrized by its tangent direction) if and only if

$$
\int_{0}^{\infty} \frac{\sin (s) d s}{\kappa(s)}=0=\int_{0}^{\infty} \frac{\cos (s) d s}{\kappa(s)}
$$

Exercise 1.2.15. Let $D$ be a convex region with shadow function $h_{D}$. For a vector $\mathbf{v} \in \mathbb{R}^{2}$ define the translated region

$$
D^{\mathbf{v}}=\{(\mathbf{x}+\mathbf{v}: \mathbf{x} \in D\} .
$$

Find the relation between $h_{D}$ and $h_{D^{\mathbf{v}}}$. Explain why this answer is inevitable in light of the formula (1.35) for the curvature.

Exercise 1.2.16. Let $D$ be a convex region with shadow function $h_{D}$. For a rotation $A=\left(\begin{array}{cc}\cos \phi & -\sin \phi \\ \sin \phi & \cos \phi\end{array}\right)$ define the rotated region

$$
D^{A}=\{A \mathbf{x}:: \mathbf{x} \in D\} .
$$

Find the relation between $h_{D}$ and $h_{D^{A}}$.
Exercise 1.2.17. ${ }^{*}$ If $h_{1}$ and $h_{2}$ are $2 \pi$-periodic functions satisfying (1.34) then they are the shadow functions of convex regions $D_{1}$ and $D_{2}$. The sum, $h_{1}+h_{2}$ also satisfies (1.34) and so is the shadow function of a convex region, $D_{3}$. Describe, geometrically how $D_{3}$ is determined by $D_{1}$ and $D_{2}$.

Exercise 1.2.18. ${ }^{*}$ Suppose that $D$ is non-convex planar region. The shadow function $h_{D}$ is defined as before. What information about $D$ in encoded in $h_{D}$ ?

### 1.2.3 Approximate reconstructions

## See: A.7.2.

In a realistic situation the shadow function is measured at a finite set of angles $\left\{\theta_{1}, \ldots, \theta_{m}\right\}$. How can the data, $\left\{h_{D}\left(\theta_{1}\right), \ldots, h_{D}\left(\theta_{m}\right)\right\}$, be used to construct an approximation to the region $D$. We consider two different strategies; each relies on the special geometric properties of convex regions. Recall that a convex region always lies in one of the half planes determined by the support line at any point of its boundary. Since the boundary of $D$ and $l_{h(\theta), \omega(\theta)}$ have the same orientation at the point of contact, it follows that $D$ lies in each of the half planes

$$
H_{h\left(\theta_{j}\right), \omega\left(\theta_{j}\right)}^{-}, \quad j=1, \ldots, m,
$$

see (1.27). As $D$ lies in each of these half planes it also lies in their intersection. This defines a convex polygon

$$
P_{m}=\bigcap_{j=1}^{m} H_{h\left(\theta_{j}\right), \omega\left(\theta_{j}\right)}^{-}
$$

which contains $D$. This polygon provides one sort of approximation for $D$ from the measurement of a finite set of shadows. It is a stable approximation to $D$ because small errors in the measurements of either the angles $\theta_{j}$ or the corresponding affine parameters $h\left(\theta_{j}\right)$ lead to small changes in the approximating polygon.

The difficulty with using the exact reconstruction formula (1.32) is that $h$ is only known at finitely many values, $\left\{\theta_{j}\right\}$. From this information it is not possible to exactly compute the derivatives, $h^{\prime}\left(\theta_{j}\right)$. We could use a finite difference approximation for the derivative to determine a finite set of points which approximate points on the boundary of $D$ :

$$
\left(x_{j}, y_{j}\right)=h\left(\theta_{j}\right) \omega\left(\theta_{j}\right)+\frac{h\left(\theta_{j}\right)-h\left(\theta_{j+1}\right)}{\theta_{j}-\theta_{j+1}} \hat{\omega}\left(\theta_{j}\right) .
$$

If the measurements were perfect, the boundary of $D$ smooth and the numbers $\left\{\left|\theta_{j}-\theta_{j+1}\right|\right\}$ small then the finite difference approximations to $h^{\prime}\left(\theta_{j}\right)$ would be accurate and these points would lie close to points on the boundary of $D$. Joining these points, in the given order gives a polygon, $P^{\prime}$ which approximates $D$. If the points could be computed exactly then $P^{\prime}$ would be contained in $D$. With approximate values this cannot be asserted with certainty, though under the assumptions above, $P^{\prime}$ should be largely contained within $D$.

This gives a different way to reconstruct an approximation to $D$ from a finite set of measurements. This method is not as robust as the first technique because it requires the measured data to be differentiated. In order for the finite difference $\frac{h\left(\theta_{j}\right)-h\left(\theta_{j+1}\right)}{\theta_{j}-\theta_{j+1}}$ to be a good approximation to $h^{\prime}\left(\theta_{j}\right)$ it is generally necessary for $\left|\theta_{j}-\theta_{j+1}\right|$ to be small. Moreover the errors in the measurements of $h\left(\theta_{j}\right)$ and $h\left(\theta_{j+1}\right)$ must also be small compared to $\left|\theta_{j}-\theta_{j+1}\right|$. This difficulty arises in solution of the reconstruction problem in X-ray CT , the exact reconstruction formula calls for the measured data to be differentiated. In general, measured data is corrupted by random noise, and random noise is usually "nondifferentiable."

This means that measurements of a function must be regularized before they can be used to approximate derivatives. One way to handle this is to improve the quality of individual measurements. One assumes that the errors in individual measurements have "mean zero:" if the same measurement is repeated many times then the average of the individual measurements should approach the true value. This is the approach taken in magnetic resonance imaging. Another possibility is to make a large number of measurements at closely spaced angles $\left\{\left(h_{j}, j \Delta \theta\right): j=1, \ldots, N\right\}$ which are then "averaged" to give less noisy approximations on a coarser grid. There are many ways to do the averaging. One way is to find a differentiable function, $H$ belonging to a family of functions of dimension $M<N$ that minimizes the square error

$$
e(H)=\sum_{j=1}^{N}\left(h_{j}-H(j \Delta \theta)\right)^{2} .
$$

For example $H$ could be taken to be a polynomial of degree $M-1$, or a continuously differentiable, piecewise cubic function. The reconstruction formula can be applied to $H$ to obtain a different approximation to $D$. The use of averaging reduces the effects of noise but fine structure in the boundary is also blurred by any such procedure.

## Exercises

Exercise 1.2.19. Suppose that the angles $\left\{\theta_{j}\right\}$ can be measured exactly but there is an uncertainty of size $\epsilon$ in the measurement of the affine parameters, $h\left(\theta_{j}\right)$. Find a polygon $P_{m, \epsilon}$ which gives the best possible approximation to $D$ and certainly contains $D$.

Exercise 1.2.20. Suppose that we know that $\left|h^{\prime \prime}(\theta)\right|<M$, and the measurement errors are bounded by $\epsilon>0$. For what angle spacing is the error in using a finite difference approximation for $h^{\prime}$ due to the uncertainty in the measurements equal to that caused by the non-linearity of $h$ itself.

### 1.2.4 Can an object be reconstructed from its width?

To measure the location of the shadow requires an expensive detector which can accurately locate a transition from light to dark. It would be much cheaper to build a device, similar to the exposure meter in a camera, to measure the length of the shadow region without determining its precise location. It is therefore an interesting question whether or not the boundary of a region can be reconstructed from measurements of the widths of its shadows. Let $w_{D}(\theta)$ denote the width of the shadow in direction $\theta$, a moments consideration shows that

$$
\begin{equation*}
w_{D}(\theta)=h_{D}(\theta)+h_{D}(\theta+\pi) . \tag{1.36}
\end{equation*}
$$

Using this formula and exercise 1.2 .8 it is easy to show that $w_{D}$ does not determine $D$. From exercise 1.2 .8 we know that if $h_{D}$ has two derivatives such that $h_{D}^{\prime \prime}+h_{D}>0$ then $h_{D}(\theta)$ is the shadow function of a strictly convex region. Let $e$ be an odd smooth function, i.e. $e(\theta)+e(\theta+\pi) \equiv 0$ such that

$$
h_{D}^{\prime \prime}+h_{D}+e^{\prime \prime}+e>0 .
$$

If $e \not \equiv 0$ then $h_{D}+e$ is the shadow function for $D^{\prime}$, a different strictly convex region. Observe that $D^{\prime}$ has the same width of shadow for each direction as $D$, that is

$$
w_{D}(\theta)=\left(h_{D}(\theta)+e(\theta)\right)+\left(h_{D}(\theta+\pi)+e(\theta+\pi)\right)=w_{D^{\prime}}(\theta) .
$$

To complete this discussion note that any function with a Fourier representation of the form

$$
e(\theta)=\sum_{j=0}^{\infty}\left[a_{j} \sin (2 j+1) \theta+b_{j} \cos (2 j+1) \theta\right]
$$

is an odd function. This is an infinite-dimensional space of functions. This implies that if $w_{D}(\theta)$ is the "width of the shadow" function for a convex region $D$ then there is an infinite dimensional set of regions with the same "width of the shadow" function. Consequently the simpler measurement is inadequate to reconstruct the boundary of a convex region. Figure 1.13 shows the unit disk and another region which has constant "shadow width" equal to 2 .


Figure 1.13: Two regions of constant width 2

## Exercises

Exercise 1.2.21. Justify the formula (1.36) for the shadow width.
Exercise 1.2.22. Show that the width function satisfies $w_{D}^{\prime \prime}+w_{D}>0$.
Exercise 1.2.23. Is it true that every twice differentiable, $\pi$-periodic function, $w$ satisfying $w^{\prime \prime}+w>0$ is the width function of a convex domain?

Exercise 1.2.24. Our motivation for considering whether or not a convex body is determined by the width of its shadows was to replace our expensive detector, which can determine where a shadow begins and ends, with a less expensive detector. The cheaper detector can only measure the width of the covered region. Can you find a way to use a detector which only measures the length of an illuminated region to locate the edge of the shadow? Hint: Cover only half of the detector with photosensitive material.

### 1.3 Conclusion

By examining a large collection of examples we have seen how physical systems can be described using mathematical models. The problem of determining the state of the system from measurements is replaced by that of solving equations or systems of equations. It is important to keep in mind that mathematical models are just that, models, often toy models. A good model must satisfy two opposing requirements: the model should accurately depict the system under study while at the same time being simple enough to be usable. In addition it must also have accurate, finite dimensional approximations.

In mathematics, problems of determining the state of a physical system from feasible measurements are gathered under the rubric of inverse problems. The division of problems into inverse problems and direct problems is often a matter of history. Usually a physical theory which models how the state of the system determines feasible measurements preceded a description of the inverse process: how to use measurements to determine the state of the system. Although very simple, example 1.1.6 is typical. Formula (2.17) describes the solution to the direct problem: the determination of the transit time from a knowledge of the sound speed and the depth. The inverse problem asks for a determination of the depth from a knowledge of the sound speed and the transit time. While many of the problems which arise in medical imaging are considered to be inverse problems, we do not give any systematic development of this subject. The curious reader is referred to the very nice article by Joe Keller, [41] which contains analyses of many classical inverse problems.

The models used in medical imaging usually involve infinitely many degrees of freedom. The state of the system is described by a function of continuous variables. Ultimately of course only a finite number of measurements can be made and only a finite amount of time is available to process them. Our analysis of the reconstruction process in X-ray CT passes through several stages. We begin with a description of the complete, perfect data situation. The measurement is described by a function on the space of lines. By finding a very explicit inversion formula we show that the state of the system can be determined from these measurements. The main tool in this analysis is the Fourier transform. We next consider the consequences of having only discrete samples of these measurements. This leads us to the Fourier series. After introducing the basic concepts, we consider the relationship between the Fourier transform and the Fourier series. With all the basic tools in hand we next study the actual algorithms used in practical imaging machines. Up to this point we have not considered the effects of noise. Probability theory is the language of noise analysis. The book concludes with a brief introduction to this subject and an analysis of the effects of noise on the quality of an approximate image, reconstructed from finitely many measurements.

In the next chapter we quickly review linear algebra, recasting some familiar concepts in the language of measurement. We also introduce the convolution product which provides a very flexible mathematical model for measurement and filtering processes.

## Chapter 2

## Some fundamental mathematics

As we have seen, using measurements to determine other quantities requires the solution of systems of equations. Sometimes we need to solve differential equations and sometimes algebraic equations. In almost all practical applications one is eventually reduced to solving systems of linear equations. This is true even for physical systems which are described by non-linear equations. As we saw in section 1.1.1, non-linear equations may be approximated by linear equations. A non-linear equation is usually solved iteratively where the iteration step involves the solution of the approximating linear system. There are many reasons why linear equations and linear models are ubiquitous. From the pragmatic point of view, there is a "complete" mathematical theory for systems of linear equations. One has necessary and sufficient conditions for linear equations to have solutions, a description of the space of solutions when they exist and practical algorithms for finding them. This is not true even for the simplest families of non-linear equations. On a more conceptual level, for systems with some sort of intrinsic smoothness, a linear model often suffices to describe small deviations from an equilibrium state.

In this chapter we review some basic concepts from linear algebra, in particular the theory of systems of linear equations. This is not intended to serve as a text on finite dimensional linear algebra but rather to situate these familiar concepts in the context of measurement problems. A more complete introduction to theory of vector spaces is presented in Appendix A.2. The systems which arise in medical imaging are not usually finite dimensional. We close this chapter by introducing the convolution product. This is a bilinear operation which provides a model for many of the measurement and filtering processes one encounters in medical imaging.

### 2.1 Linear equations and linear maps

Suppose that the state of a system is described by the variables $\left(x_{1}, \ldots, x_{n} ; y_{1}, \ldots, y_{m}\right)$ or more concisely by $(\mathbf{x} ; \mathbf{y})$ with $\mathbf{x} \in \mathbb{R}^{n}$ and $\mathbf{y} \in \mathbb{R}^{m}$. Here $\mathbf{x}$ is regarded as an input variable and $\mathbf{y}$ is regarded as an output variable. The model for the system is expressed by the
linear equations,

$$
\begin{array}{ccc}
a_{11} x_{1}+a_{12} x_{2}+ & \cdots+a_{1 n} x_{n} & =y_{1} \\
a_{21} x_{1}+a_{22} x_{2}+ & \cdots+a_{2 n} x_{n} & =y_{2} \\
\vdots & \vdots & \vdots  \tag{2.1}\\
a_{m 1} x_{1}+a_{m 2} x_{2}+ & \cdots+a_{m n} x_{n} & =y_{m}
\end{array}
$$

Before proceeding with our analysis we first need to simplify the notation. It is very cumbersome to have to work with complicated expressions like (2.1), instead we use standard matrix and vector notation. Let a denote the $m \times n$ array of numbers $\left(a_{i j}\right)_{i=1 \ldots m, j=1 \ldots n}$, $\mathbf{x}$ an $n$-vector and $\mathbf{y}$ an $m$-vector. The system of equations (2.1) is concisely expressed as

$$
\begin{equation*}
\mathrm{ax}=\mathbf{y} \tag{2.2}
\end{equation*}
$$

where ax denotes the matrix product. We briefly recall the properties of matrix multiplication. Let $\mathbf{x}_{1}$ and $\mathbf{x}_{2}$ be $n$-vectors, then

$$
\mathbf{a}\left(\mathbf{x}_{1}+\mathbf{x}_{2}\right)=\mathbf{a} \mathbf{x}^{1}+\mathbf{a x}_{2}
$$

and for any number $c$

$$
\mathbf{a}\left(c \mathbf{x}_{1}\right)=c\left(\mathbf{a x}_{1}\right) .
$$

In other words the map from $\mathbb{R}^{n}$ to $\mathbb{R}^{m}$ defined by $\mathbf{x} \mapsto \mathbf{a x}$ is linear. The state space for the system is the graph of this linear transformation,

$$
\left\{(\mathbf{x} ; \mathbf{a x}): \mathbf{x} \in \mathbb{R}^{n}\right\} .
$$

Let $\mathbf{a}_{i}=\left(a_{i 1}, \ldots, a_{i n}\right)$ denote the $i^{\text {th }}$ row of $\mathbf{a}$. The equations in (2.1) can also be rewritten

$$
\begin{gather*}
\left\langle\mathbf{a}_{1}, \mathbf{x}\right\rangle=y_{1}  \tag{2.3}\\
\vdots \\
\left\langle\mathbf{a}_{m}, \mathbf{x}\right\rangle=y_{m}
\end{gather*}
$$

The inner product $\left\langle\mathbf{a}_{i}, \mathbf{x}\right\rangle$ is interpreted as the outcome of the $i^{\text {th }}$ measurement when the system is in the state described by $\mathbf{x}$. The matrix $\mathbf{a}$ is called the measurement matrix. The state of the system is determined by $\mathbf{x}$, which can take any value in $\mathbb{R}^{n}$. The output variables $\mathbf{y}$ are determined by (2.2); the question of principal interest is the extent to which these measurements determine the state. If the rows of a are linearly dependent then there is some index $l$ and constants $\left\{c_{i}\right\}$ so that

$$
\mathbf{a}_{l}=\sum_{i \neq l} c_{i} \mathbf{a}_{i}
$$

By linearity this means that

$$
\left\langle\mathbf{a}_{l}, \mathbf{x}\right\rangle=\sum_{i \neq l} c_{j}\left\langle\mathbf{a}_{i}, \mathbf{x}\right\rangle .
$$

In other words, the outcome of the $l^{\text {th }}$ measurement is already determined by the others. It is not an independent measurement and, in pure mathematics, would be regarded as
redundant information. In applications, measurement error and noise make it useful to repeat experiments, so such a measurement might very well be retained.

There is a final way to rewrite (2.1). If

$$
\mathbf{a}^{j}=\left(\begin{array}{c}
a_{1 j} \\
\vdots \\
a_{m j}
\end{array}\right)
$$

denotes the $j^{\text {th }}$ column of $\mathbf{a}$ then (2.1) is equivalent to

$$
x_{1} \mathbf{a}^{1}+\cdots+x_{n} \mathbf{a}^{n}=\mathbf{y} .
$$

In this form it is clear that the set of possible outcomes of these measurements is the subspace of $\mathbb{R}^{m}$ spanned by the columns of $\mathbf{a}$. This is the range of the linear transformation a.

There are three questions which require answers:

## Q1: Existence:

For a given $m$-vector $\left(y_{1}, \ldots, y_{m}\right)$ does there exist an $n$-vector $\left(x_{1}, \ldots, x_{n}\right)$ which satisfies the equations in (2.1)?

Q2: Uniqueness:
When a solution exists is it unique? More generally, describe the space of solutions.
Q3: Stability:
How sensitive is the solution to small variations in the measurement matrix a or the right hand side $\mathbf{y}$ ?

It is a somewhat unexpected, but very important fact that these issues are, in practice, rather independent of one another. For applications it is also necessary to have an algorithm to find approximate solutions of (2.1) and criteria to select a solution when there is more than one.

### 2.1.1 Solving linear equations

Suppose that $\mathbf{x}_{0}$ is a solution of the equation $\mathbf{a x}=0$ and $\mathbf{x}_{1}$ is a solution of the equation $\mathbf{a x}_{1}=\mathbf{y}$. By linearity it follows that

$$
\mathbf{a}\left(c \mathbf{x}_{0}+\mathbf{x}_{1}\right)=c \mathbf{a x}_{0}+\mathbf{a x}_{1}=\mathbf{a x}_{1}=\mathbf{y} \text { for any } c \in \mathbb{R} .
$$

If $\mathbf{y}=0$ as well then we conclude that the set of solutions to the equation

$$
\mathrm{ax}=0
$$

is a linear subspace, that is if $\mathbf{x}_{0}$ and $\mathbf{x}_{1}$ solve this equation then so does $\mathbf{x}_{0}+\mathbf{x}_{1}$ as well as $c \mathbf{x}_{0}$, for any number $c$. This subspace is called the null space or kernel of $\mathbf{a}$. It is denoted by $\operatorname{ker}(\mathbf{a})$ and always contains, at least the zero vector $\mathbf{0}=(0, \ldots, 0)$. These observations question Q2.

Theorem 2.1.1. Let a be an $m \times n$ matrix. Suppose that $\mathbf{x}_{1}$ satisfies $\mathbf{a x}_{1}=\mathbf{y}$ then every other solution to this equation is of the form $\mathbf{x}_{1}+\mathbf{x}_{0}$ where $\mathbf{x}_{0} \in \operatorname{ker}(\mathbf{a})$. Moreover, every vector of this form solves the equation $\mathbf{a x}=\mathbf{y}$.

The solution of the equation $\mathbf{a x}=\mathbf{y}$ is unique only if the null space of a contains only the 0 -vector or, in other words, the columns of a are linearly independent.

We now turn to question Q1. Suppose that $\mathbf{a}$ is an $m \times n$-matrix, $\mathbf{x}$ is an $n$-vector and $\mathbf{y}$ is an $m$-vector then ax is an $m$-vector and

$$
\langle\mathbf{a x}, \mathbf{y}\rangle=\sum_{i=1}^{m} \sum_{j=1}^{n} a_{i j} y_{i} x_{j} .
$$

The transpose of the matrix $\mathbf{a}$ is the $n \times m$ matrix $\mathbf{a}^{t}$ whose $i j$-entry is $a_{j i}$. From the previous formula it follows that that

$$
\langle\mathbf{a x}, \mathbf{y}\rangle=\left\langle\mathbf{x}, \mathbf{a}^{t} \mathbf{y}\right\rangle .
$$

Suppose that $\mathbf{b}$ is a non-zero vector in the null space of the transpose, $\mathbf{a}^{t}$ and the the equation $\mathbf{a x}=\mathbf{y}$ has a solution. Using the calculations above we see that

$$
\langle\mathbf{y}, \mathbf{b}\rangle=\langle\mathbf{a x}, \mathbf{b}\rangle=\left\langle\mathbf{x}, \mathbf{a}^{t} \mathbf{b}\right\rangle=0
$$

The last equality follows from the fact that $\mathbf{a}^{t} \mathbf{b}=0$. This gives a necessary condition for existence of a solution to the equation $\mathbf{a x}=\mathbf{y}$. The vector $\mathbf{y}$ must satisfy the conditions

$$
\langle\mathbf{b}, \mathbf{y}\rangle=0,
$$

for every solution of the homogeneous equation $\mathbf{a}^{t} \mathbf{b}=0$. This also turns out to be sufficient.
Theorem 2.1.2. Let $\mathbf{a}$ be an $m \times n$-matrix and $\mathbf{y}$ and $m$-vector. The equation $\mathbf{a x}=\mathbf{y}$ has a solution if and only if

$$
\langle\mathbf{b}, \mathbf{y}\rangle=0
$$

for every vector $\mathbf{b} \in \operatorname{ker}\left(\mathbf{a}^{t}\right)$.
The equation $\mathbf{a}^{t} \mathbf{b}=0$ has non-trivial solutions if and only if the rows of a are linearly dependent. This means that the outcomes of the measurements $\left\{\left\langle\mathbf{a}_{i}, \mathbf{x}\right\rangle: i=1, \ldots, m\right\}$ are not independent of one another. The condition for the solvability of $\mathbf{a x}=\mathbf{y}$ is simply that entries of $\mathbf{y}$ should satisfy the same relations as the measurements themselves.

The corollary summarizes these results.
Corollary 2.1.1. Let $\mathbf{a}$ be an $m \times n$-matrix the equation $\mathbf{a x}=\mathbf{y}$ has a unique solution for any vector $\mathbf{y}$ if and only if $\operatorname{ker}(\mathbf{a})=\{0\}$ and $\operatorname{ker}\left(\mathbf{a}^{t}\right)=\{0\}$. In particular, this implies that $m=n$.

Example 2.1.1. Suppose we have a collection of photons sources, labeled by $1 \leq i \leq n$ and an array of detectors, labeled by $1 \leq j \leq m$. The matrix $\mathbf{P}$ has entries $0 \leq p_{i j} \leq 1$. The
$i j$-entry is the probability that a particle emitted from source $i$ is detected by detector $j$. Since a given photon can be detected by at most one detector it follows that

$$
\sum_{j=1}^{m} p_{i j} \leq 1 \text { for } i=1, \ldots n
$$

If $d_{j}, j=1, \ldots, m$ is the number of photons detected at detector $j$ and $s_{i}, i=1, \ldots, n$ is the number of photons emitted by source $i$ then our model predicts that

$$
\mathbf{P s}=\mathbf{d}
$$

If $m=n$ and $\mathbf{P}$ is an invertible matrix then we can use the measurements $\mathbf{d}$ to obtain a unique vector $\mathbf{s}$. Since the model is probabilistic this should be regarded as an expected value for the distribution of sources. If $m>n$ then we have more measurements than unknowns, so any measurement errors or flaws in the model could make it impossible to find a vector $\mathbf{s}$ so that $\mathbf{P s}=\mathbf{d}$. This is a frequent situation in image reconstruction problems. One chooses a way to measure the error, usually a function of the form $e(\mathbf{P s}-\mathbf{d})$ and seeks a vector $\mathbf{s}$ which minimizes the error. Finally we may have more sources than detectors. The measurements are then inadequate, in principle to determine their distribution. This is also a common circumstance in image reconstruction problems and is resolved by making some a priori assumptions about the allowable distribution of sources to obtain a determined (or even overdetermined) problem.

As illustrated by this example and explained in the theorem there are essentially 3 types of linear models for systems with finitely many degrees of freedom.

## Determined:

The simplest case arises when the number of independent measurements and parameters describing the state of the system are the same. This implies that $n=m$. In this case the measurements uniquely determine the state of the system. Mathematically we say that the matrix, a is invertible. For a square matrix this is equivalent to the statement that the homogeneous equation, $\mathbf{a x}=0$ has only the trivial solution, $\mathbf{x}=0$. The inverse matrix is denoted by $\mathbf{a}^{-1}$, it is both a left and a right inverse to a,

$$
\mathbf{a}^{-1} \mathbf{a}=\mathrm{Id}_{n}=\mathbf{a a}^{-1}
$$

Here $\mathrm{Id}_{n}$ denotes the $n \times n$ identity matrix, that in

$$
\left(\operatorname{Id}_{n}\right)_{i j}=\left\{\begin{array}{l}
1 \text { if } i=j, \\
0 \text { if } i \neq j .
\end{array}\right.
$$

From the mathematical point of view, the unique solution is obtained by setting

$$
\mathbf{x}=\mathbf{a}^{-1} \mathbf{y}
$$

Except in special cases, the inverse matrix $\mathbf{a}^{-1}$ is not computed directly.

## Overdetermined:

In this case we have more measurements than parameters, i.e. $m>n$. If the model and measurements are perfect then there should be a unique $\mathbf{x}$ with $\mathbf{a x}=\mathbf{y}$. In general, neither is true and there will not exist any $\mathbf{x}$ exactly satisfying this equation. Having more measurements than parameters can be used to advantage in several different ways. In example 2.1.2 we explain how to use the conditions for solvability given in Theorem 2.1.2 to determine physical parameters. Often times measurements are noisy. A model for the noise in the measurements can be used to select a criterion for a "best approximate solution." The most common way to measure the error is to use the norm defined in (1.7) setting

$$
e(\mathbf{x})=\|\mathbf{a x}-\mathbf{y}\|^{2} .
$$

There are two reasons why this measure of the error is often employed: 1 . It is a natural choice if the noise is normally distributed, 2 . The problem of minimizing $e(\mathbf{x})$ can be reduced to the problem of solving a system of linear equations.

## Underdetermined:

Most of the problems in image reconstruction are underdetermined, that is we do not have enough data to uniquely determine a solution. In mathematical tomography a "perfect reconstruction" requires an infinite number of exact measurements. These are, of course never available. In a linear algebra problem, this is the case where $m<n$. When the measurements $\mathbf{y}$ do not uniquely determine the state $\mathbf{x}$, additional criteria are needed to determine which solution to actually use. For example one might use the solution to $\mathbf{a x}=\mathbf{y}$ which is of smallest norm. Another approach is to assume that $\mathbf{x}$ belongs to a subspace whose dimension is equal to the number of independent measurements. Both of these approaches are used in medical imaging.

Example 2.1.2. In the refraction problem considered in example 1.1.4 we remarked that the refractive index of the lower fluid $n_{2}$ could be determined by an additional measurement. Suppose that we shine a beam of light in at a different angle, so that the upper angle is $\phi_{1}$ and the lower angle is $\phi_{2}$. This light beam is displaced by $l_{2}$ as it passes through the fluid. We now have 3 equations for the two unknowns:

$$
\left(\begin{array}{cc}
1 & 1  \tag{2.4}\\
\tan \left(\theta_{1}\right) & \tan \left(\theta_{2}\right) \\
\tan \left(\phi_{1}\right) & \tan \left(\phi_{2}\right)
\end{array}\right)\binom{h_{1}}{h_{2}}=\left(\begin{array}{l}
h \\
l_{1} \\
l_{2}
\end{array}\right) .
$$

In order for this equation to have a solution the measurements $\left(h, l_{1}, l_{2}\right)$ must satisfy the condition

$$
\left(\begin{array}{c}
1 \\
\tan \left(\theta_{1}\right) \\
\tan \left(\phi_{1}\right)
\end{array}\right) \times\left(\begin{array}{c}
1 \\
\tan \left(\theta_{2}\right) \\
\tan \left(\phi_{2}\right)
\end{array}\right) \cdot\left(\begin{array}{l}
h \\
l_{1} \\
l_{2}
\end{array}\right)=0 .
$$

Here $\times$ is the vector cross product. Since

$$
\frac{\sin \left(\theta_{1}\right)}{\sin \left(\theta_{2}\right)}=\frac{\sin \left(\phi_{1}\right)}{\sin \left(\phi_{2}\right)}=\frac{n_{2}}{n_{1}}
$$

and the angles $\theta_{1}$ and $\phi_{1}$ as well as $\left(h, l_{1}, l_{2}\right)$ are assumed known, this solvability conditions gives a non-linear equation which allows the determination of $\frac{n_{2}}{n_{1}}$ from the measured data.

The problem of stability, Q3 is considered in the next section. This material is not essential for the remainder of this chapter and can safely be skipped.

## Exercises

Exercise 2.1.1. Let a be an $m \times n$ matrix. Show that if $\operatorname{ker} \mathbf{a}=\operatorname{ker} \mathbf{a}^{t}=\mathbf{0}$ then $n=m$. Is the converse true?

Exercise 2.1.2. Let $\mathbf{a}: \mathbb{R}^{n} \rightarrow \mathbb{R}^{n}$ be an invertible linear transformation show that $\mathbf{a}^{-1}$ is also linear.

Exercise 2.1.3. Suppose that the state of a system is described by the vector $\mathbf{x}$. The measurements are modeled as inner products $\left\{\left\langle\mathbf{a}_{i}, \mathbf{x}\right\rangle: i=1, \ldots, m\right\}$. However the measurements are noisy and each is repeated $m_{i}$ times leading to measured values $\left\{y_{i}^{1}, \ldots, y_{j}^{m_{i}}\right\}$. Define an error function by

$$
e(\mathbf{x})=\sum_{i=1}^{m} \sum_{k=1}^{m_{j}}\left(\left\langle\mathbf{a}_{i}, \mathbf{x}\right\rangle-y_{i}^{k}\right)^{2} .
$$

Show that $e(\mathbf{x})$ is minimized by the vector which satisfies the averaged equations

$$
\left\langle\mathbf{a}_{i}, \mathbf{x}\right\rangle=\frac{1}{m_{i}} \sum_{k=1}^{m_{i}} y_{i}^{k} .
$$

### 2.1.2 Stability of solutions*

Suppose that $\mathbf{a}$ is an $n \times n$ invertible matrix that models a measurement process. If $\mathbf{x}_{1}$ and $\mathbf{x}_{2}$ are two states of our system then, because the model is linear the difference in the measurements can easily be computed

$$
\mathbf{y}_{1}-\mathbf{y}_{2}=\mathbf{a x}_{1}-\mathbf{a x}_{2}=\mathbf{a}\left(\mathbf{x}_{1}-\mathbf{x}_{2}\right) .
$$

When studying stability it is very useful to have a measure for the "size" of a linear transformation.

Definition 2.1.1. Let a be an $m \times n$-matrix. The Euclidean operator norm of a is defined to be

$$
\|\mathbf{a}\|=\max \|\mathbf{a x}\|: \mathbf{x} \in \mathbb{R}^{n} \text { with }\|\mathbf{x}\|=1
$$

This method for measuring the size of linear transformations has many useful properties. Let $\mathbf{a}_{1}$ and $\mathbf{a}_{2}$ by $m \times n$-matrices and $c$ a real number, then

$$
\begin{align*}
\left\|c \mathbf{a}_{1}\right\| & =|c|\left\|\mathbf{a}_{1}\right\|, \\
\left\|\mathbf{a}_{1}+\mathbf{a}_{2}\right\| & \leq\left\|\mathbf{a}_{1}\right\|+\left\|\mathbf{a}_{2}\right\|,  \tag{2.5}\\
\left\|\mathbf{a}_{1} \mathbf{x}\right\| & \leq\left\|\mathbf{a}_{1}\right\|\|\mathbf{x}\| \text { for all } \mathbf{x} \in \mathbb{R}^{n}
\end{align*}
$$

Using the operator norm we obtain the estimate

$$
\left\|\mathbf{y}_{1}-\mathbf{y}_{2}\right\| \leq\|\mathbf{a}\|\| \| \mathbf{x}_{1}-\mathbf{x}_{2} \| .
$$

Therefore nearby states result in nearby measurements. However the reverse is often not true. There may exist states $\mathbf{x}_{1}$ and $\mathbf{x}_{2}$ which are not nearby, in the sense that $\left\|\mathbf{x}_{1}-\mathbf{x}_{2}\right\|$ is large but $\left\|\mathbf{a}\left(\mathbf{x}_{1}-\mathbf{x}_{2}\right)\right\|$ is small. Physically, the measurements performed are not sufficiently independent to distinguish certain pairs of states, which are not, in fact very close together. In numerical analysis this is known as an ill-conditioned equation. Briefly, a small error in the measurement process can be magnified by applying $\mathbf{a}^{-1}$ to the measurement vector. For an ill-conditioned problem even a good algorithm for solving linear equations can produce meaningless results.
Example 2.1.3. For example, consider the system with $m=n=2$ and

$$
\mathbf{a}=\left(\begin{array}{cc}
1 & 0 \\
1 & 10^{-5}
\end{array}\right) .
$$

Then $\mathbf{x}$ is given by $\mathbf{a}^{-1} \mathbf{y}$ where

$$
\mathbf{a}=\left(\begin{array}{cc}
1 & 0 \\
-10^{5} & 10^{5}
\end{array}\right) .
$$

If the actual data is $\mathbf{y}=(1,1)$ but we make an error in measurement and measure, $\mathbf{y}_{m}=$ $(1,1+\epsilon)$ then the relative error is

$$
\frac{\left|\mathbf{y}_{m}-\mathbf{y}\right|}{|\mathbf{y}|}=\epsilon 10^{5} .
$$

Even though the measurements uniquely determine the state of the system, a small error in measurement is vastly amplified.

The uncertainty in the model can be itself be modeled by using $\mathbf{a}+\delta \mathbf{a}$. for the measurement matrix. Here $\delta \mathbf{a}$ represents an aggregation of the errors in the model. The measurements are also subject to error and should be considered to have the form $\mathbf{y}+\delta \mathbf{y}$. A more realistic problem is therefore to solve the system of equations

$$
\begin{equation*}
(\mathbf{a}+\delta \mathbf{a}) \mathbf{x}=\mathbf{y}+\delta \mathbf{y} \tag{2.6}
\end{equation*}
$$

But what does this mean?
We consider only the simplest case where $\mathbf{a}$ is an $n \times n$, invertible matrix. Suppose that we can bound the uncertainty in both the model and the measurements in the sense that we have constants $\epsilon>0$ and $\eta>0$ such that

$$
\|\delta \mathbf{y}\|<\epsilon \text { and }\|\delta \mathbf{a}\|<\eta
$$

In the absence of more detailed information about the systematic errors, "the solution" to (2.6) should be defined as the set of vectors

$$
\{\mathbf{x}: \mid(\mathbf{a}+\delta \mathbf{a}) \mathbf{x}=\mathbf{y}+\delta \mathbf{y} \text { for some choice of } \delta \mathbf{a}, \delta \mathbf{y} \text { with }\|\delta \mathbf{y}\|<\epsilon,\|\delta \mathbf{a}\|<\eta\}
$$

This is a little cumbersome. In practice one finds a vector which satisfies

$$
a x=y
$$

and a bound for the error one makes in asserting that the actual state of the system is $\mathbf{x}$.
To proceed with this analysis we assume that all the possible model matrices, $\mathbf{a}+\delta \mathbf{a}$ are invertible. If $\|\delta \mathbf{a}\|$ is sufficiently small then this condition is satisfied. As a is invertible the number

$$
\begin{equation*}
\mu=\min _{\mathbf{x} \neq 0} \frac{\|\mathrm{ax}\|}{\|\mathbf{x}\|} \tag{2.7}
\end{equation*}
$$

is a positive. If $\|\delta \mathbf{a}\|<\mu$ then $\mathbf{a}+\delta \mathbf{a}$ is also invertible. The proof of this statement is given in section A.2.7. In the remainder of this discussion we assume that $\eta$, the bound on the uncertainty in the model is smaller than $\mu$.

An estimate on the error in $\mathbf{x}$ is found in two steps. First, fix the model and consider only errors in measurement. Suppose that $\mathbf{a x}=\mathbf{y}$ and $\mathbf{a}(\mathbf{x}+\delta \mathbf{x})=\mathbf{y}+\delta \mathbf{y}$. Taking the difference of these two equations gives

$$
\mathbf{a} \delta \mathbf{x}=\delta \mathbf{y}
$$

and therefore $\delta \mathbf{x}=\mathbf{a}^{-1} \delta \mathbf{y}$. Using (2.5) we see that

$$
\|\delta \mathbf{x}\| \leq\left\|\mathbf{a}^{-1}\right\|\|\delta \mathbf{y}\| .
$$

This is a bound on the absolute error; it is more meaningful to bound the relative error $\|\delta \mathbf{x}\| /\|\mathbf{x}\|$. To that end observe that

$$
\|\mathbf{y}\| \leq\|\mathbf{a}\|\|\mathbf{x}\|
$$

and therefore

$$
\begin{equation*}
\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq\|\mathbf{a}\|\left\|\mathbf{a}^{-1}\right\| \frac{\|\delta \mathbf{y}\|}{\|\mathbf{y}\|} \tag{2.8}
\end{equation*}
$$

This is a very useful estimate: it estimates the relative uncertainty in the state in terms of the relative uncertainty in the measurements. The coefficient

$$
\begin{equation*}
c_{\mathbf{a}}=\|\mathbf{a}\|\| \| \mathbf{a}^{-1} \| \tag{2.9}
\end{equation*}
$$

is called the condition number of the matrix a. It is very useful measure of the stability of a model of this type.

To complete our analysis we need to incorporate errors in the model. Suppose that $\mathbf{x}+\delta \mathbf{x}$ solves

$$
(\mathbf{a}+\delta \mathbf{a})(\mathbf{x}+\delta \mathbf{x})=\mathbf{y}+\delta \mathbf{y} .
$$

Subtracting this from $\mathbf{a x}=\mathbf{y}$ gives

$$
(\mathbf{a}+\delta \mathbf{a}) \delta \mathbf{x}=\delta \mathbf{y}-\delta \mathbf{a x} .
$$

Proceeding as before we see that

$$
\begin{equation*}
\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq\left\|(\mathbf{a}+\delta \mathbf{a})^{-1}\right\|\| \| \mathbf{a}\left\|\frac{\|\delta \mathbf{y}\|}{\|\mathbf{y}\|}+\right\|(\mathbf{a}+\delta \mathbf{a})^{-1}\| \|\|\delta \mathbf{a}\| . \tag{2.10}
\end{equation*}
$$

If $\delta \mathbf{a}$ is very small (relative to $\mu$ ) then

$$
(\mathbf{a}+\delta \mathbf{a})^{-1}=\mathbf{a}^{-1}-\mathbf{a}^{-1} \delta \mathbf{a} \mathbf{a}^{-1}+O\left([\delta \mathbf{a}]^{2}\right) .
$$

The triangle inequality implies that

$$
\left\|(\mathbf{a}+\delta \mathbf{a})^{-1}\right\| \lesssim\left\|\mathbf{a}^{-1}\right\|+\left\|\mathbf{a}^{-1} \delta \mathbf{a} \mathbf{a}^{-1}\right\| .
$$

Ignoring quadratic error terms this gives the estimate

$$
\begin{equation*}
\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq c_{\mathbf{a}}\left[\frac{\|\delta \mathbf{y}\|}{\|\mathbf{y}\|}+\frac{\|\delta \mathbf{a}\|}{\|\mathbf{a}\|}\right] . \tag{2.11}
\end{equation*}
$$

Once again, it is the condition number of a which relates the relative error in the predicted state to the relative errors in the model and measurements.

This analysis considers a very special case, but it indicates how gross features of the model constrain the accuracy of its predictions. We have discussed neither the effects of using a particular algorithm to solve the system of equations or round-off error, a similar analysis applies to study these problems. A very good reference for this material is [78]. In image reconstruction the practical problems of solving systems of linear equations are considerable. It is not uncommon to have 10,000 -equations in 10,000 -unknowns. These huge systems arise as finite dimensional approximations to linear equations for functions of continuous variables. We close this section with a short discussion of linear algebra in in infinite dimensional spaces. This is a theme which occupies a large part of this book.

## Exercises

Exercise 2.1.4. Show that $\mu$, defined in (2.7) equals $\left\|\mathbf{a}^{-1}\right\|$.
Exercise 2.1.5. Show that the condition number is given by the following ratio

$$
\begin{equation*}
c_{\mathbf{a}}=\frac{\max _{\mathbf{x} \neq 0} \frac{\|\mathbf{a x}\|}{\|\mathbf{x}\|}}{\min _{\mathbf{x} \neq 0} \frac{\|\mathbf{a x}\|}{\|\mathbf{x}\|}} . \tag{2.12}
\end{equation*}
$$

This shows that the condition number of any matrix is at least 1 .

### 2.2 Infinite dimensional linear algebra

The state of a 'system' in medical imaging is described by a function of continuous variables. In this introductory section we consider real valued functions defined on the real line. Let $f(x)$ describe the state of the system. A linear measurement of the state is usually described as an integral

$$
\mathcal{M}(f)(x)=\int_{-\infty}^{\infty} m(x, y) f(y) d y
$$

Here $m(x, y)$ is a function on $\mathbb{R} \times \mathbb{R}$ which provides a model for the measurement process. It can be thought of as an infinite 'matrix' with indices $x$ and $y$. A linear transformation
of an infinite dimensional space is called a linear operator. A linear transformation which can be expressed as an integral is called an integral operator.

Suppose that the function $g(x)$ is the output of the measurement process, to reconstruct $f$ means solving the linear equation

$$
\mathcal{M} f=g .
$$

This is a concise way to write a system of infinitely many equations in infinitely many unknowns. Theorems 2.1.1 and 2.1.2 contain the complete theory for the existence and uniqueness of solutions to linear equations in finitely many variables. These theorems are entirely algebraic in character. No such theory exists for equations in infinitely many variables. It is usually a very complicated problem to describe both the domain and range of such a transformation. We close this section with a few illustrative examples.
Example 2.2.1. Perhaps the simplest linear operator is the indefinite integral

$$
\mathcal{I}(f)(x)=\int_{0}^{x} f(y) d y
$$

If we use the continuous functions on $\mathbb{R}$ as the domain of $\mathcal{I}$ then every function in the range is continuously differentiable. Moreover the null-space of $\mathcal{I}$ is the zero function. Observe that the domain and range of $\mathcal{I}$ are fundamentally different spaces. Because $\mathcal{I}(f)(0)=0$ not every continuously differentiable function is in the range of $\mathcal{I}$. The derivative is a left inverse to $\mathcal{I}$ as the Fundamental Theorem of Calculus states that if $f$ is continuous then

$$
\frac{d}{d x} \circ \mathcal{I}(f)(x)=f(x)
$$

On the other hand it is not quite a right inverse because

$$
\mathcal{I}\left(\frac{d f}{d x}\right)(x)=f(x)-f(0) .
$$

The domain of $\mathcal{I}$ can be enlarged to include all locally integrable functions. These are functions such that

$$
\int_{0}^{x}|f(y)| d y<\infty
$$

for every $x \in \mathbb{R}$. Enlarging the domain also enlarges the range. For example the function $|x|$ lies in the enlarged range of $\mathcal{I}$,

$$
|x|=\int_{0}^{x} \operatorname{sign}(y) d y
$$

where $\operatorname{sign}(y)=1$ if $y \geq 0$ and -1 if $y<0$. Even though $|x|$ is not differentiable at $x=0$ it is still the indefinite integral of a locally integrable function, however the formula

$$
\frac{d|x|}{d x}=\operatorname{sign}(x)
$$

does not make sense at $x=0$.

Example 2.2.2. Changing the lower limit of integration to $-\infty$ leads to a very different sort of linear transformation. Initially $\mathcal{I}_{\infty}$ is defined for continuous functions $f$, vanishing for sufficiently negative $x$ by

$$
\mathcal{I}_{\infty}(f)(x)=\int_{-\infty}^{x} f(y) d y
$$

Once again the null-space of $\mathcal{I}_{\infty}$ consists of the zero function alone. The domain can be enlarged to include locally integrable functions such that

$$
\begin{equation*}
\lim _{R \rightarrow \infty} \int_{-R}^{0}|f(y)| d y<\infty \tag{2.13}
\end{equation*}
$$

If $f$ is continuous then we can apply the F.T.C. to obtain

$$
\frac{d}{d x} \circ \mathcal{I}(f)=f
$$

If a function $g$ belongs to the range of $\mathcal{I}$ then

$$
\begin{equation*}
\lim _{x \rightarrow-\infty} g(x)=0 \tag{2.14}
\end{equation*}
$$

There are once differentiable functions satisfying this condition which do not belong to the range of $\mathcal{I}_{\infty}$. For example,

$$
f(x)=\frac{x \cos x-\sin x}{x^{2}}=\frac{d}{d x} \frac{\sin x}{x}
$$

satisfies (2.14) but $\frac{\cos x}{x}$ does not satisfy (2.13). With the domain defined by (2.13) the precise range of $\mathcal{I}_{\infty}$ is rather difficult to describe.

This example illustrates how a integral operator may have a simple definition on a certain domain, which by a limiting process can be extended to a larger domain. The domain of such an operator is often characterized by a size condition like (2.13).
Example 2.2.3. A real physical measurement is always some sort of an average. If the state of the system is described by a function $f$ of a single variable $x$ then the average of $f$ over an interval of length $2 \delta$ is

$$
\mathcal{M}_{\delta}(f)(x)=\frac{1}{2 \delta} \int_{x-\delta}^{x+\delta} f(y) d y
$$

A natural domain for $\mathcal{M}_{\delta}$ is all locally integrable functions. To what extent is $f$ determined by $\mathcal{M}_{\delta}(f)$ ? Suppose that $f$ and $g$ are two states, then, because the integral is linear

$$
\mathcal{M}_{\delta}(f)-\mathcal{M}_{\delta}(g)=\mathcal{M}_{\delta}(f-g)
$$

The extent to which $\mathcal{M}_{\delta}(f)$ determines $f$ is characterized by the null-space of $\mathcal{M}_{\delta}$,

$$
\mathcal{N}_{\delta}=\left\{f: \mathcal{M}_{\delta}(f)=0\right\} .
$$

Proceeding formally, we can differentiate $\mathcal{M}_{\delta}(f)$ to obtain

$$
\begin{equation*}
\frac{d \mathcal{M}_{\delta}(f)}{d x}=f(x+\delta)-f(x-\delta) \tag{2.15}
\end{equation*}
$$

If $f \in \mathcal{N}_{\delta}$ then $\mathcal{M}_{\delta}(f)$ is surely constant and therefore

$$
f \in \mathcal{N}_{\delta} \Rightarrow f(x+\delta)-f(x-\delta)=0,
$$

in other words $f$ is periodic with periodic $2 \delta$. A periodic function has an expansion in terms of sines and cosines, that is

$$
f(x)=a_{0}+\sum_{j=1}^{\infty}\left[a_{j} \cos \left(\frac{\pi j x}{\delta}\right)+b_{j} \sin \left(\frac{\pi j x}{\delta}\right)\right] .
$$

If $a_{0}=0$ then $\mathcal{M}_{\delta}(f)=0$. This shows that the null-space of $\mathcal{M}_{\delta}$ is infinite dimensional.
In applications one often has additional information about the state of the system, for example one might know that

$$
\begin{equation*}
\lim _{|x| \rightarrow \infty} f(x)=0 \tag{2.16}
\end{equation*}
$$

A periodic function that tends to zero at infinity must be identically zero, so among such functions the measurements $\mathcal{M}_{\delta}(f)$ would appear to determine $f$ completely. To prove this statement we need to know somewhat more about $f$ than (2.16). With a more quantitative condition like

$$
\begin{equation*}
\|f\|_{p}=\left[\int_{-\infty}^{\infty}|f(y)|^{p} d y\right]^{\frac{1}{p}}<\infty \tag{2.17}
\end{equation*}
$$

for a $p$ between 1 and 2 , it is possible to show that $\mathcal{M}_{\delta}(f)=0$ implies that $f=0$. For such functions the measurement $\mathcal{M}_{\delta}(f)$ uniquely determines $f$. However, $f$ cannot be stably reconstructed from $\mathcal{M}_{\delta}(f)$. A small error in measurement can lead to a very large error in the reconstructed state.

The integral in (2.17) defines a measure for the size of $f$ called the $L^{p}$-norm. It is a generalization of the notion of a norm on a finite dimensional vector space and satisfies the familiar conditions for a norm:

$$
\|a f\|_{p}=|a|\|f\|_{p} \text { and }\|f+g\|_{p} \leq\|f\|_{p}+\|g\|_{p} .
$$

The first step in analyzing linear transformations of infinite dimensional spaces is the introduction of norms on the domain and range. This was not necessary in finite dimensions but is absolutely essential in the infinite dimensional case. In medical image reconstruction there is a small list of linear transformations that are very important, the Fourier transform, Radon transform and Abel transform. A large part of this text is devoted to the analysis of these operators.

## Exercises

Exercise 2.2.1. Prove that the null-space of $\mathcal{I}$ acting on $\mathcal{C}^{0}(\mathbb{R})$ is the zero function.

### 2.3 Conclusion

The text by Peter D. Lax, [43] provides an execellent introduction to the theoretical aspects of linear algebra while Trefethen and Bau, [78] covers the numerical side of the subject.


[^0]:    ${ }^{1}$ Chest x-ray provided courtesy of Dr. David S. Feigen, ENS Sherri Rudinsky and Dr. James G. Smirniotopoulos of the Uniformed Services University of the Health Sciences, Dept. of Radiology, Bethesda, MD

