MATH 5001 - MATHEMATICS OF MEDICAL IMAGING

JASON MURPHY, MISSOURI S&T

Abstract. These notes accompanied Math 5001 ('Mathematics of Medical Imaging') in Fall 2020 at Missouri University of Science & Technology. The notes are based primarily off of the textbook of C. L. Epstein [1], with a number of figures imported from that text.

Contents

References

[1] Charles L. Epstein, *Introduction to the Mathematics of Medical Imaging.*

Overview: X-ray CT and the associated mathematical problems

X-ray computerized tomography (CT) for non-invasive medical imaging applications was pioneered by Hounsfield in the early 1970s, with the possibility of such an application proposed several years earlier by Cormack. Since then, X-ray CT has been developed extensively and has become one of the primary medical imaging modalities in use today.

In ordinary X-ray imaging, one essentially obtains an image of a shadow of a three-dimensional object, which we can view as the superposition of the shadow of each two-dimensional slice of the object. In X-ray CT, we instead obtain precise imaging of each individual slice by taking multiple measurements of the 'shadows' of individual lines/beams of X-rays. There are different options for precisely how to configure the beams (e.g. parallel beams versus fan beams).

The goal in X-ray CT is to reconstruct the *attenuation coefficient* $\mu : \mathbb{R}^3 \to [0, \infty)$ of some part of the body. This function quantifies the tendency of tissue to absorb versus scatter X-rays. We work with one slice at a time, each of which is represented by a function on $\mathbb{R}^2$, using a coordinate system determined by the X-ray machine.

What we actually measure in X-ray CT essentially determines the value of

$$\int_L \mu \, ds,$$

where L is the line traversed by the X-ray beam. To derive this model of measurement, we use some basic physics of X-rays and a few simplifying assumptions. In particular, we ignore some effects like refraction/diffusion and assume infinitely thin and monochromatic X-rays, and the main underlying physical principle is that of Beer's law. Some of the physical assumptions are more realistic than others. The monochromatic assumption is not very realistic, but it plays a key role in what follows because it leads to a *linear* model for measurements. More accurate models would also take into account that the integrals above should really be taken over strips, or even $3d$ objects.

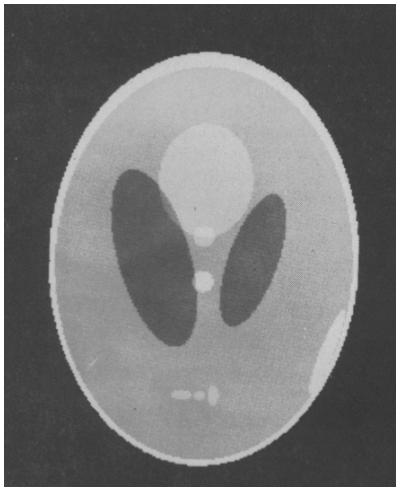
The map $\mu \mapsto \int_\bullet \mu \, ds$ is described mathematically by the *Radon transform*. This is a linear transformation, but must be interpreted in the setting of infinite-dimensional vector spaces. For the question of reconstructing μ from its Radon transform, we are immediately led to questions such as invertibility and (if invertibility holds) inversion formulas. In fact, we will show that the Radon transform is invertible for suitable types of functions, with accompanying inversion formulas (in particular, the filtered back-projection formula).

In any real application, there will be approximations made at many levels. It will therefore be important to understand stability/continuity properties of the Radon transform and its inverse. We will need to understand how to approximate inversion formulas (e.g. with finite sums replacing integrals), how to perform inversion with incomplete data, how to account for failures of uniqueness, and so on. This mathematical understanding will inform the development of several approaches to reconstruction.

For the problems of inverting the Radon transform and designing reconstruction algorithms, we rely largely on Fourier-based approaches. In particular, we will need to develop mathematics related to convolution and filtering, Fourier series, and the Fourier transform. As we will see, the choice of scanner geometry (e.g. parallel

beam versus fan beam geometry) plays an important role in the approach taken to reconstruction.

To test reconstruction algorithms, one can study so-called *mathematical phantoms* (as introduced by Shepp). The idea is to simulate a body section by a mathematical function that can be described explicitly. In this way, we remove all 'measurement error' and can understand precisely what errors are introduced by our reconstruction algorithms. An example of a phantom is given in the following figure (from L. A. Shepp and J. B. Kruskal, *Computerized Tomography: The New Medical X-Ray Tomography*, American Mathematical Monthly **85** (1978), no. 6, 420–439).



After developing the theory needed to describe several reconstruction algorithms in detail, we will turn to several other important topics that may be studied mathematically. The first such topic concerns various imaging artifacts that arise in X-ray CT. We will also study the effect of noise in measurements, which entails the introduction of some topics from probability theory.

In the final few sections of these notes, we will cover some additional topics, including an introduction to magnetic resonance imaging (MRI).

Primer on finite- and infinite-dimensional linear algebra

We will work with many different vector spaces in this class, including the standard finite-dimensional spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, as well the infinite-dimensional Lebesgue function spaces $L^p(\mathbb{R}^n)$ and sequence spaces $\ell^p(\mathbb{C})$ (with $1 \leq p \leq \infty$). All of these vector spaces admit standard *norms* (or lengths), which in turn define standard *metrics* (or distances), which in turn define *topologies* that give rise to notions of open sets, closed sets, convergence, continuous functions, and so on. Some of the vector spaces we consider additionally admit *inner products* which may be used to define the norm (and hence all of the things just mentioned), but which also give us a notion of angle/orthogonality.

We will assume familiarity with the usual spaces $\mathbb{R}^n$ and $\mathbb{C}^n$. We summarize what we need in the following example (in the setting of $\mathbb{C}^n$).

Example 1. *An element of $\mathbb{C}^n$ may be denoted by z, which we view as a column vector*

$$z = (z_1, z_2, \dots, z_n)^T,$$

where ^{T} denotes the transpose and each entry z_j is a complex number (we write $z_j \in \mathbb{C}$). This means $z_j = x_j + iy_j$ for some real numbers x_j, y_j (we write $x_j, y_j \in \mathbb{R}$) and i is the imaginary unit obeying $i^2 = -1$. In general, we will not use bold text for vectors or do any decorations with arrows. The inner product *of $z \in \mathbb{C}^n$ and $w \in \mathbb{C}^n$ is defined by*

$$\langle z, w \rangle := z^* w = \sum_{j=1}^{n} \bar{z}_j w_j = \bar{z}_1 w_1 + \cdots + \bar{z}_n w_n.$$

Here $^-$ denotes complex conjugation, defined by $\overline{x+iy} = x - iy$, and * denotes the conjugate transpose (also called the adjoint). In the setting of $\mathbb{C}^n$ (or $\mathbb{R}^n$) we may use the more familiar 'dot product' notation and write $\langle z, w \rangle = z \cdot w$.

For a complex number $z = x + iy$, we have $\bar{z}z = x^2 + y^2$, which corresponds to the euclidean length of the vector $(x, y)^T$ as an element of $\mathbb{R}^2$ and is denoted in this context by $|z|^2$.

The norm *of a vector $z \in \mathbb{C}^n$ is defined by*

$$\|z\| := \sqrt{\langle z, z \rangle} = \sqrt{\sum_{j=1}^{n} |z_j|^2} = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

In the setting of $\mathbb{C}^n$ and $\mathbb{R}^n$, we commonly write $|z|$ instead of $\|z\|$, leaving the $\|\cdot\|$ notation primarily for norms defined on function spaces. However, we will continue to use the notation $\|\cdot\|$ for the remainder of this example.

We have the Cauchy–Schwarz and triangle inequalities, given respectively by

$$|\langle z, w \rangle| \leq \|z\| \, \|w\| \quad \text{and} \quad \|z + w\| \leq \|z\| + \|w\|.$$

The distance *between two vectors $z, w \in \mathbb{C}^n$ is defined by*

$$d(z, w) = \|z - w\|.$$

Any norm gives rise to a distance (or metric*) in this way.*

We then consider the topology *on $\mathbb{C}^n$ generated by the basis of open balls in $\mathbb{C}^n$. In particular, given $z_0 \in \mathbb{C}^n$ and $r > 0$, we define the open ball $B(z_0, r)$ by*

$$B(z_0, r) = \{z \in \mathbb{C}^n : \|z - z_0\| < r\},$$

and we call a set $S \subset \mathbb{C}^n$ open *if for every* $z_0 \in S$*, there exists* $r > 0$ *such that* $B(z_0, r) \subset S$*. A set is* closed *if its complement is open. A set* $S \subset \mathbb{C}^n$ *is called bounded if there exists* $R > 0$ *such that* $S \subset B(0, R)$*.*

A sequence (z_k) *in* $\mathbb{C}^n$ *is* Cauchy *if for any* $\varepsilon > 0$*, there exists* $N > 0$ *such that*

$$j, k \geq N \implies \|z_j - z_k\| < \varepsilon,$$

and the sequence converges to z *if for any* $\varepsilon > 0$*, there exists* $N > 0$ *such that*

$$j \geq N \implies \|z_j - z\| < \varepsilon.$$

In the latter case, we write $z_j \to z$ *as* $j \to \infty$ *or* $\lim_{j\to\infty} z_j = z$*. The metric space* $\mathbb{C}^n$ *is* complete*, which means that any Cauchy sequence converges to some element of* $\mathbb{C}^n$*. When we need to refer to the components of a collection of vectors, we may use notation such as* $z_j = (z_{j,1}, \ldots, z_{j,n})^T$*.*

Finally, a mapping $T : \mathbb{C}^n \to \mathbb{C}^m$ *is called* continuous *if for all* $z \in \mathbb{C}^n$*, we have that*

$$z_n \to z \implies T(z_n) \to T(z).$$

The spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ are examples of *Banach spaces*: normed vector spaces that are complete with respect to the metric induced by the norm. In fact, they are *Hilbert spaces*: inner product spaces that are (i) complete with respect to the induced metric and (ii) separable. (A metric space (V, d) is separable if there exists a countable set $\{\phi_j\} \subset V$ such that for any $v \in V$ and any $\varepsilon > 0$, there exists j such that $d(v, \varphi_j) < \varepsilon$.)

The vector spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ are n-dimensional. Indeed, one has the standard basis $\{e_j\}_{j=1}^n$, where

$$e_{j,k} = \begin{cases} 1 & k = j \\ 0 & k \neq j. \end{cases}$$

We may 'send $n \to \infty$' to obtain the vector space of sequences of real or complex numbers, which may be denoted by $\mathbb{R}^\omega$ or $\mathbb{C}^\omega$, respectively. That is, an element of $\mathbb{C}^\omega$ is a sequence $(z_j)_{j=1}^\infty$ of complex numbers. We can define a family of subspaces of $\mathbb{C}^\omega$ (or $\mathbb{R}^\omega$) by introducing the standard ℓ^p-norms.

Example 2. *Let* $1 \leq p \leq \infty$*. We define* $\ell^p(\mathbb{C})$ *to be the subspace of* $\mathbb{C}^\omega$ *consisting of sequences* $z = (z_j)_{j=1}^\infty$ *such that*

$$\|z\|_{\ell^p} < \infty,$$

where the ℓ^p*-norm* $\|\cdot\|_{\ell^p}$ *is defined by*

$$\|z\|_{\ell^p} = \begin{cases} \left(\sum_{j=1}^\infty |z_j|^p\right)^{\frac{1}{p}} & 1 \leq p < \infty, \\ \sup_{1\leq j\leq\infty} |z_j| & p = \infty. \end{cases}$$

The space ℓ^p *is a Banach space for every* $1 \leq p \leq \infty$*. Moreover,* ℓ^2 *is a Hilbert space with the inner product*

$$\langle z, w \rangle = \sum_{j=1}^\infty \bar{z}_j w_j.$$

We will also make frequent use of the Lebesgue spaces on $\mathbb{R}^n$. To define these carefully requires the introduction of Lebesgue measure. Let us briefly recall the definitions and some of the key properties.

Lebesgue measure can be constructed by first defining a notion of an 'outer measure' on subsets of $\mathbb{R}^n$. In particular, this is defined so that it returns the correct volume for any set defined by the cartesian product of intervals. A set in $\mathbb{R}^n$ is then called measurable if it satisfies a certain regularity condition, namely, that it is well-approximated from without (in the sense of this outer measure) by an open set. The measure of a measurable set is simply its outer measure. Essentially any set you will encounter in 'everyday life' will be measurable, but nonmeasurable sets certainly do exist. Lebesgue measure agrees with our intuitive notion of 'n-dimensional volume'. We do not really need to go through all the definitions, but it is worth describing one special case in detail. A set $E \subset \mathbb{R}^n$ has measure zero, written $|E| = 0$, if for any $\varepsilon > 0$, there exists a countable collection of balls B_k in $\mathbb{R}^n$ such that

$$E \subset \bigcup_k B_k \quad \text{and} \quad \sum_k |B_k| < \varepsilon.$$

Lebesgue measure is used to define a theory of integration. In particular, given a measurable set E of finite measure, one first defines the *characteristic function* of E to be

$$\chi_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$$

and defines the integral to be $\int \chi_E \, dx = |E|$. One then extends this definition to act linearly on finite linear combinations of characteristic functions of disjoint sets (called 'simple functions'), and subsequently uses approximation by simple functions to define the integral of more general functions. While this sounds more complicated than Riemann integration, it turns out that when a function is integrable in both the Lebesgue and Riemann senses, the two integrals return the same value. What is gained by Lebesgue integration is the ability to ignore the behavior of functions on sets of measure zero. The classical example is the Dirichlet function, defined by the characteristic function of the rational numbers. This function is not Riemann integrable (it is 'too discontinuous'), but it is easily integrated in the Lebesgue sense. Indeed, because the rational numbers have measure zero (this is true of any countable set), the integral of this function is zero.

The inability of the Lebesgue integral to 'see' sets of measure zero also means that when we later define Lebesgue spaces, we must implicitly identify any function f with the equivalence class of all functions that equal f 'almost everywhere' (that is, the set of functions g such that the set $\{x \in \mathbb{R}^n : f(x) \neq g(x)\}$ has measure zero).

We can now define the Lebesgue spaces on $\mathbb{R}^n$.

Example 3. *For $1 \leq p < \infty$, we define $L^p(\mathbb{R}^n)$ to be the set of all functions $f : \mathbb{R}^n \to \mathbb{C}$ such that*

$$\|f\|_{L^p} := \left(\int_{\mathbb{R}^n} |f(x)|^p \, dx \right)^{\frac{1}{p}} < \infty.$$

We define $L^\infty(\mathbb{R}^n)$ to be the set of all functions $f : \mathbb{R}^n \to \mathbb{C}$ such that

$$\|f\|_{L^\infty} := \inf\{\alpha > 0 : |\{x \in \mathbb{R}^n : |f(x)| > \alpha\}| = 0\} < \infty.$$

In particular, $\|f\|_{L^\infty}$ is the infimal M such that $|f(x)| \leq M$ almost everywhere.

The spaces $L^p(\mathbb{R}^n)$ are Banach spaces for all $1 \leq p \leq \infty$. In addition, $L^2(\mathbb{R}^n)$ is a Hilbert space with the inner product

$$\langle f, g \rangle = \int_{\mathbb{R}^n} \overline{f(x)} g(x)\, dx.$$

As mentioned above, we must view elements of $L^p(\mathbb{R}^n)$ as equivalence classes of functions that are equal almost everywhere.

To verify the claims above, one needs to check that the quantities $\|\cdot\|_{L^p}$ do in fact define norms (e.g. they obey the triangle inequality). In addition, one must prove completeness (i.e. that Cauchy sequences converge). These are well-known facts in analysis and we will take them for granted here.

The Lebesgue spaces above were defined for complex-valued functions, although one can restrict to real-valued functions as well. If necessary, the codomain of the functions may be indicated by using notation such as $L^p(\mathbb{R}^n; \mathbb{C})$ or $L^p(\mathbb{R}^n; \mathbb{R})$.

One of the main tools from Lebesgue integration that we will need is the following theorem, called the dominated convergence theorem.

Theorem 1 (Dominated Convergence Theorem)**.** *Suppose $f_k : \mathbb{R}^n \to \mathbb{C}$ is a sequence of measurable functions that converge to some function $f : \mathbb{R}^n \to \mathbb{C}$ almost everywhere. If there exists a function $g \in L^1(\mathbb{R}^n)$ such that $|f_k(x)| \leq |g(x)|$ almost everywhere, then*

$$\lim_{k \to \infty} \int_{\mathbb{R}^n} f_k(x)\, dx = \int_{\mathbb{R}^n} f(x)\, dx.$$

It is also useful to work with function spaces that are defined in terms of the regularity of the functions in the space.

Example 4. *We define $C(\mathbb{R}^n)$ to be the vector space of continuous functions $f : \mathbb{R}^n \to \mathbb{C}$. This is a Banach space under the supremum norm (which agrees with the L^∞-norm for continuous functions).*

For $k = 1, 2, 3, \ldots$, we define $C^k(\mathbb{R}^n)$ to be the vector space of functions $f : \mathbb{R}^n \to \mathbb{C}$ that are k-times continuously differentiable. The space $C^\infty(\mathbb{R}^n)$ consists of functions that are k-times differentiable for all $k \in \mathbb{N}$.

At times, it will be useful to restrict to functions with 'bounded support', i.e. functions such that $\{x \in \mathbb{R}^n : f(x) \neq 0\}$ is a bounded set. We can indicate this with the subscript $_b$, e.g. $C_b(\mathbb{R}^n)$ denotes the vector space of continuous functions of bounded support. We may also write $C_0(\mathbb{R}^n)$ for the vector space of continuous functions f such that $\lim_{|x| \to \infty} f(x) = 0$.

We next discuss linear transformations on vector spaces. Given two vector spaces V and W over $\mathbb{C}$, we call a mapping $T : V \to W$ a linear transformation if

$$T(\alpha v + \beta w) = \alpha T(v) + \beta T(w)$$

for all $\alpha, \beta \in \mathbb{C}$ and $v, w \in V$. For finite-dimensional vector spaces, all linear transformations are represented by matrices and so the situation is somewhat familiar. In the infinite-dimensional setting, the situation can change quite a bit. In what follows, let us work through a few examples that are representative of what we will encounter later on.

Example 5. *Let $V = L^1(\mathbb{R}^n)$, $W = L^\infty(\mathbb{R}^n)$ and let $K \in L^\infty(\mathbb{R}^n \times \mathbb{R}^n)$. Define the mapping $T : V \to W$ via*

$$Tf(x) = \int_{\mathbb{R}^n} K(x,y) f(y)\, dy.$$

The fact that T maps V into W follows from the inequalities

$$|Tf(x)| \leq \int |K(x,y)|\, |f(y)|\, dy \leq \|K\|_{L^\infty(\mathbb{R}^n \times \mathbb{R}^n)} \|f\|_{L^1}$$

uniformly over $x \in \mathbb{R}^n$. We call T an integral transform *and we call K the* (integral) kernel *of T. The inequalities above also show that T is a* bounded *operator from L^1 to L^∞; in particular,*

$$\|T\|_{L^1 \to L^\infty} := \sup\big\{ \|Tf\|_{L^\infty} : \|f\|_{L^1} = 1 \big\} \leq \|K\|_{L^\infty(\mathbb{R}^n \times \mathbb{R}^n)}.$$

For linear transformations, boundedness is equivalent to continuity.

In the special case that $K(x,y) = \psi(x-y)$ for some function $\psi : \mathbb{R}^n \to \mathbb{C}$, we call T a convolution operator *and ψ the* (convolution) kernel *of T.*

Example 6. *Suppose V is a vector space over $\mathbb{C}$ with an inner product $\langle \cdot, \cdot \rangle$ and that W is a closed subspace of V. Then for each $v \in V$, there exists a unique element $P(v) \in W$ such that*

$$\|v - P(v)\| = \inf_{w \in W} \|v - w\|.$$

The mapping $P : V \to W$ defined by $v \mapsto P(v)$ is a linear transformation, called the orthogonal projection of V onto W. *The transformation P is bounded/continuous, with norm equal to 1. For any $v \in V$, we have that $P(v)$ is orthogonal to $v - P(v)$, that is,*

$$\langle P(v), v - P(v) \rangle = 0 \quad \textit{for all} \quad v \in V.$$

Next suppose that we have an orthonormal basis $\{\phi_j\}$ for W. This means that the span of the ϕ_j is dense in W and that

$$\langle \phi_j, \phi_k \rangle = \begin{cases} 1 & j = k \\ 0 & j \neq k. \end{cases}$$

Then the orthogonal projection of v onto W is computed via

$$P(v) = \sum_j \langle \phi_j, v \rangle \phi_j.$$

Given a linear transformation $T : V \to W$, we define the kernel (or null-space) of T by

$$N(T) = \{x \in V : T(x) = 0\}$$

and the image (or range) of T by

$$R(T) = \{T(x) : x \in V\}.$$

We call T injective if $N(T) = \{0\}$ and surjective of $R(T) = W$.

In finite-dimensional linear algebra, one is often faced with linear systems of the form $Ax = b$ that may or may not have solutions. Viewing $x \mapsto Ax$ as a linear transformation (which we still denote by A), one can instead solve the associated *normal system* $A^T Ax = A^T b$, which is always consistent. In fact, solving the normal system is equivalent to solving

$$Ax = \text{proj}_{R(A)}(b), \tag{1}$$

where $\text{proj}_{R(A)}$ denotes the orthogonal projection onto $R(A)$. This system is evidently consistent, and solutions to this equation are minimizers of the problem

$$\text{minimize } \|Ax - b\|^2$$

over all choices of x. For this reason, we call solutions to (1) *least-squares solutions* to the original system. If A has full rank, the least-squares solution is unique.

The physics of X-ray CT

The goal of X-ray tomography is to reconstruct a three-dimensional object from its two-dimensional slices. The object is described by its attenuation coefficient, which is a function $\mu : \mathbb{R}^3 \to [0, \infty)$ that quantifies the tendency of the object to absorb or scatter X-rays. We assume that air is transparent to X-rays, so that $\mu \equiv 0$ outside of the patient.

Remark 1. *We will work directly with the attenuation coefficient. In radiology one instead works with a dimensionless quantity called a Hounsfield unit, which compares the attenuation coefficient to that of water (and takes on both negative and positive values). From the following table (imported from [1]), we can see that the range encountered in a CT measurement may be around 2000 units, while one needs to be able to reconstruct the attenuation coefficient accurately to within around 10 units to distinguish between the different types of tissues.*

Material	**Attenuation coefficient in Hounsfield units**
water	0
air	−1000
bone	1086
blood	53
fat	−61
brain white/gray	−4
breast tissue	9
muscle	41
soft tissue	51

Let us now describe a simplified classical model for the interaction of X-rays with matter. With the interpretation of X-rays as high-energy electromagnetic radiation, we assume:

(i) No refraction/diffraction: X-ray beams travel along straight lines and are not bent by the objects they pass through.
(ii) The X-rays are monochromatic (i.e. all of a single frequency).
(iii) Beer's law: Each material encountered has a characteristic attenuation coefficient μ for X-rays of a given energy/frequency. The intensity I of the X-ray beam satisfies
$$\tfrac{dI}{ds} = -\mu I,$$
where s is the arc-length along the straight-line trajectory of the X-ray beam.

Assumption (i) is reasonable for high energy X-rays, while assumption (ii) is not. However, it is necessary if we wish to work with a linear model for measurement (more on this later).

Beer's law is essentially a probabilistic one. Under the monochromatic assumption, the intensity is proportional to the number of photons per second, and Beer's law posits that $\mu(s)\Delta s$ is the probability that a photon incident on the material at coordinate s is absorbed over a length Δs.

Using Beer's law, we may deduce that the intensity of an X-ray beam is attenuated on the line segment joining points $x_0 + av$ and $x_0 + bv$ by
$$\exp\Big\{-\int_a^b \mu(x_0 + sv)\, ds\Big\}.$$

When we use Beer's law, we implicitly assume that attenuation is isotropic (that is, independent of the direction of travel).

Example 1 (Beam Spreading)**.** *Suppose we have a point source of X-rays of intensity I_0 (measured in electron-volts/second) placed at the origin in the plane, with the same outgoing flux in all directions. Writing $I(r)$ for the intensity at distance r from the origin, we may derive that $I(r) = \frac{1}{2\pi r} I_0$ through an argument using conservation of energy.*

This effect can also be modeled by Beer's law. We introduce the function $\mu_s = 1/r$ to account for the attenuation due to the spreading of the beam. Then Beer's law implies $rI(r) = r_0 I(r_0)$ for any $r, r_0 > 0$, which agrees with the identity above.

We say that we have a non-diverging source *of X-rays whenever the attenuation due to beam spreading is very small compared to the attenuation due to the object.*

To make measurements, we turn on an X-ray source for some fixed period of time. We therefore know the total energy, denoted I_i, incident on the object along a line ℓ. We measure the total energy, denoted I_o, emerging from the object along ℓ using an X-ray detector.

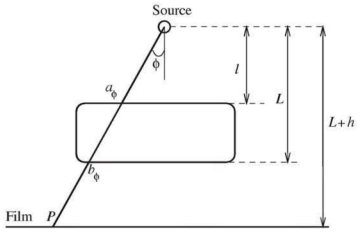
If we integrate Beer's law along this line, then we obtain

$$\log\left[\tfrac{I_0}{I_i}\right] = -\int_\ell \mu \, ds.$$

We model the measurements we make via these line integrals.

If we send X-rays along many parallel lines, we obtain a projection (or shadow) of the object. If we place the X-ray source at many different angles, we can better distinguish between various arrangements of objects. This is the basic idea underlying X-ray CT, with the goal of reconstructing each two-dimensional slice of the object under consideration.

Example 2. *Suppose we have a point source of X-rays, a photographic plate, and an attenuating body in between (as in the figure, imported from [1]). We may use polar coordinates (r, ϕ) based at the source.*



The attenuation coefficient is the sum of μ_a (attenuation to due absorption) and $\mu_s = r^{-1}$ (attenuation due to beam spreading). The attenuation of the x-ray beam along a line through the source at angle ϕ is given by Beer's law:

$$\tfrac{d}{dr} I = -[\mu_a(r, \phi) + \tfrac{1}{r}] I.$$

We expose the film by turning on the source for some known period of time. Let us also use I to denote the energy per unit length resulting from this exposure.

Proposition 1. *Let a_ϕ and b_ϕ denote the first and last points of intersection of the X-ray line with the absorbing body, and let $L + h$ be the distance from the source to*

the film. Let $\delta(\phi)$ denote the density of the film at the point where the line meets the film. Then

$$-\int_{a_\phi}^{b_\phi} \mu_a(s,\phi)\,ds = \gamma^{-1}\delta(\phi) - \log\big[\tfrac{I_0\cos^2\phi}{2\pi(L+h)}\big]$$

for some (physical) constant $\gamma > 0$.

The quantities on the right-hand side are determined by measurement. In particular, the measurement is a linear function of the attenuation coefficient.

Proof. Integrating Beer's law from some small $r = r_0 > 0$ to $r = r_\phi$ (the distance to the film along the line) yields

$$\log \tfrac{I(r_\phi,\phi)}{I(r_0,\phi)} = \log \tfrac{r_0}{r_\phi} - \int_{a_\phi}^{b_\phi} \mu_a(s,\phi)\,ds.$$

With ℓ the distance to the front of the body, and L the distance to the back of the body, we have

$$a_\phi = \tfrac{\ell}{\cos\phi}, \quad b_\phi = \tfrac{L}{\cos\phi}, \quad \text{and} \quad r_\phi = \tfrac{L+h}{\cos\phi}.$$

Thus (recalling that $r_0 I(r_0,\phi) = \frac{1}{2\pi} I_0$ due to beam spreading)

$$I(r_\phi,\phi) = I_0 \cdot \frac{\cos\phi}{2\pi(L+h)} \exp\Big\{-\int_{a_\phi}^{b_\phi} \mu_a(s,\phi)\,ds\Big\}.$$

The density of the developed film at a point is proportional to the logarithm of the total energy incident at the point (we take this for granted). We next compute this energy:

We need to determine the flux across the part of the film subtended by a small angle $\Delta\phi$. This is given by

$$\Delta F = \int_{\phi}^{\phi+\Delta\phi} I(r_\phi,\phi)\hat{r}\cdot\hat{n}\,d\sigma,$$

where $\hat{r} = -(\sin\phi,\cos\phi)$, $\hat{n} = (0,-1)$ is the outward unit normal to the film plane, and $d\sigma$ is the arc-length element along the film, given in polar coordinates by

$$d\sigma = \tfrac{L+h}{\cos^2\phi} d\phi$$

(here we are using $r = (L+h)[\cos\phi]^{-1}$ and $d\sigma = \sqrt{r^2 + (\frac{dr}{d\phi})^2}d\phi$).

Now, for $\Delta\phi$ small, we get

$$\begin{aligned}\Delta F &\approx \int_{\phi}^{\phi+\Delta\phi} I(r_\phi,\phi)\hat{r}\cdot\hat{n}\tfrac{L+h}{\cos^2\phi}d\phi \\ &\approx I_0 \frac{\cos^2\phi}{2\pi(L+h)} \exp\Big\{-\int_{a_\phi}^{b_\phi} \mu_a(s,\phi)\,ds\Big\}\frac{L+h}{\cos^2\phi}\Delta\phi.\end{aligned}$$

Using that the length of film subtended by angle $\Delta\phi$ is approximately $\Delta\sigma = \frac{L+h}{\cos^2\phi}\Delta\phi$, we deduce that the energy density at the point P_ϕ (where the line meets the film) is

$$\frac{dF}{d\sigma} = \frac{I_0\cos^2\phi}{2\pi(L+h)} \exp\Big\{-\int_{a_\phi}^{b_\phi} \mu_a(s,\phi)\,ds\Big\}.$$

and so the density of the film is

$$\delta(\phi) = \gamma \log \tfrac{dF}{d\sigma} = \gamma \left[\log \tfrac{I_0 \cos^2 \phi}{2\pi(L+h)} - \int_{a_\phi}^{b_\phi} \mu_a(s,\phi)\, ds \right]$$

for some γ. Rearranging, we get

$$-\int_{a_\phi}^{b_\phi} \mu_a(s,\phi)\, ds = \gamma^{-1}\delta(\phi) - \log\big[\tfrac{I_0 \cos^2 \phi}{2\pi(L+h)}\big],$$

as desired. □

By moving the source and film around a circle enclosing the absorbing body, we can measure the line integrals of the attenuation coefficient for every line intercepting the body. The fact that this allows us to recover the function itself is what makes X-ray CT work.

Before we turn to the mathematical formulation that describes the measurements above, let us conclude this discussion by revisiting some of our physical assumptions above.

The assumption that our X-rays are monochromatic is not realistic. Rather, our X-ray source will be described by some spectral distribution function $S(\lambda)$. The attenuation coefficient of a material is frequency-dependent and typically decreases as the energy increases. This fact leads to the phenomenon of *beam hardening*, which means that the distribution of output energies is skewed towards higher energies. Integrating Beer's law and taking into account the dependence on the spectral parameter, one can derive that the measured output (for an X-ray beam along a line ℓ) is given by

$$\Psi_o = \int_0^\infty S(\lambda) \exp\left\{ -\int_\ell \mu(x,\lambda)\, ds \right\} d\lambda.$$

In particular, the measurements are now modeled by *nonlinear* functions of μ and hence the reconstruction problem becomes much more difficult. We will return to this issue later.

Another issue related to our model arises from the fact that an X-ray beam is not continuous, but rather it is comprised of discrete particles (photons). The physical effect of this is the presence of random noise in the measurements. While this issue could essentially be solved by simply increasing the number of photons, doing so can become dangerous to the patient! We will discuss this tradeoff between patient safety and contrast/resolution in a later section.

Introduction to the Radon transform

In X-ray CT, measurements are modeled by line integrals of a two-dimensional slice of the attenuation coefficient of the object being imaged. The collection of all such line integrals is described mathematically by using an integral transform known as the *Radon transform*.

To define this transform, we begin by introducing a convenient parametrization of lines in the plane. In what follows, we denote the unit circle in $\mathbb{R}^2$ by $\mathbb{S}^1$. That is,

$$\mathbb{S}^1 = \left\{ \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} \in \mathbb{R}^2 : \omega_1^2 + \omega_2^2 = 1 \right\}.$$

Given $t \in \mathbb{R}$ and $\omega \in \mathbb{S}^1$, the set

$$\{(x, y) \in \mathbb{R}^2 : \langle (x, y), \omega \rangle = t\}$$

is a line in the plane, where we recall the notation $\langle \cdot, \cdot \rangle$ denotes the usual inner product. In particular, the vector ω is perpendicular to this line, and the value $|t|$ is the distance from the line to the origin.

We obtain the same set if we use $(-t, -\omega)$ instead of (t, ω). However, if we impose an orientation (i.e. a positive direction along the line), then (t, ω) defines a unique line, which we denote $\ell_{t,\omega}$. To impose an orientation, we observe that $\hat{\omega} := (-\omega_2, \omega_1)^T$ is perpendicular to ω and therefore parallel to the line—we use $\hat{\omega}$ to define the positive direction along the line.

The set of points on $\ell_{t,\omega}$ may be parametrized as follows:

$$\{t\omega + s\hat{\omega} : s \in \mathbb{R}\}.$$

In fact, given (x, y) on the line $\ell_{t,\omega}$, one obtains the corresponding $s \in \mathbb{R}$ by solving

$$\begin{bmatrix} \omega_1 & -\omega_2 \\ \omega_2 & \omega_1 \end{bmatrix} \begin{bmatrix} t \\ s \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

In practice, the CT machine determines a coordinate system (x_1, x_2, x_3) for $\mathbb{R}^3$. Writing μ for the attenuation coefficient, we fix a height c and measure line integrals of the slice $\mu(\cdot, \cdot, c)$. In particular, for the oriented line $\ell_{t,\omega}$, the measurement is described in terms of the integral

$$\int_{-\infty}^{\infty} \mu(t\omega + s\hat{\omega}, c)\, ds.$$

In fact, this is the Radon transform of the function $\mu(\cdot, \cdot, c)$ evaluated at (t, ω).

Definition 1 (Radon transform)**.** *Let $f \in C_b(\mathbb{R}^2)$. The Radon transform of f is the function*

$$\mathcal{R}f : \mathbb{R} \times \mathbb{S}^1 \to \mathbb{R}$$

defined by

$$\mathcal{R}f(t, \omega) = \int_{-\infty}^{\infty} f(s\hat{\omega} + t\omega)\, ds,$$

where

$$\omega = (\omega_1, \omega_2)^T \quad \text{and} \quad \hat{\omega} = (-\omega_2, \omega_1)^T.$$

Remark 1. *To define the Radon transform of f, it is not necessary for f to be continuous, nor is it necessary for f to have bounded support. What is needed*

is enough regularity to restrict the function to a line, as well as enough decay to guarantee that the improper integrals above converge. That is, we need to have

$$\int_{-\infty}^{\infty} |f(t\omega + s\hat{\omega})|\, ds < \infty \quad \textit{for all} \quad (t,\omega) \in \mathbb{R} \times \mathbb{S}^1.$$

We call the set of functions satisfying this property the natural domain *of* $\mathcal{R}$.

Our interest in this section will not be to develop the theory of the Radon transform, but rather to discuss a few relevant properties related to the reconstruction question, and to work towards a formal inversion formula.

First, observe that the Radon transform is a linear transformation. It also preserves nonnegativity. Finally, as $\ell_{t,\omega}$ and $\ell_{-t,-\omega}$ yield the same line, the Radon transform of a function is always even:

$$\mathcal{R}f(t,\omega) = \mathcal{R}f(-t,-\omega) \quad \text{for all} \quad t,\omega.$$

We next compute a simple example.

Example 1. *Let* $B = B(0,1)$ *be the unit ball in* $\mathbb{R}^2$ *and let* χ_B *denote its characteristic function, that is,*

$$\chi_B(x) = \begin{cases} 1 & x \in B \\ 0 & x \notin B. \end{cases}$$

Then for each (t,ω)*, we have that*

$$\mathcal{R}\chi_B(t,\omega) = \textit{length of intersection of } \ell_{t,\omega} \cap \chi_B.$$

Thus

$$\mathcal{R}\chi_B(t,\omega) = \begin{cases} 2\sqrt{1-t^2} & |t| \le 1 \\ 0 & |t| > 1. \end{cases}$$

The next example shows that the Radon transform *cannot* distinguish between two functions that agree almost everywhere (in the sense of Lebesgue measure.)

Example 2. *Define*

$$f(x,y) = \begin{cases} 1 & x \in [-1,1] \quad \textit{and} \quad y = 0, \\ 0 & \textit{otherwise.} \end{cases}$$

Then

$$\mathcal{R}f(t,\omega) = \begin{cases} 2 & \textit{if } t = 0 \quad \textit{and} \quad \omega = (0,\pm 1)^T \\ 0 & \textit{otherwise.} \end{cases}$$

In fact, if we replace f *by any other function that equals 1 on a subset of* $\mathbb{R} \times \{0\}$ *of total length 2, then we will obtain the* same *result for the calculation of the Radon transform.*

Combining the previous example with linearity, we see that the Radon transform may map a nonzero function the identically zero function. This non-uniqueness, or failure of injectivity, seems like it could be an issue, given that we would eventually like to invert the Radon transform!

The resolution to this issue comes from observing that that the non-uniqueness above stems from the fact that the functions in question differ only on sets of Lebesgue measure zero. When we discuss the formal inversion of the Radon transform, we will do so in the setting of function spaces in which such functions are regarded as the same element in the space. In particular, there is a straightforward

mathematical resolution to this issue. At the same time, no physical measurement will be able to detect behavior on sets of measure zero. Thus any practical reconstruction algorithm will necessarily need to be insensitive to this type of issue, anyway.

In practice, we will be imaging objects of finite size; correspondingly, the attenuation coefficient will be modeled by a function of bounded support. This support property is reflected in the Radon transform:

Lemma 1. *Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ has bounded support, i.e. there exists $R > 0$ such that $f(x) = 0$ for $|x| > R$. Then*

$$\mathcal{R}f(t, \omega) = 0 \quad \textit{whenever} \quad |t| > R.$$

Proof. If $\ell_{t,\omega}$ is a line with $|t| > R$, then $\ell_{t,\omega}$ lies outside the support of f. Thus, by definition of the Radon transform, $\mathcal{R}f(t, \omega) = 0$ for any ω. □

The converse is false:

Example 3. *Define the function $f : \mathbb{R}^2 \to \mathbb{R}$ in polar coordinates via*

$$f(r, \theta) = \begin{cases} \frac{1}{r} \cos(\theta) & r > 1 \\ 0 & r \leq 1. \end{cases}$$

Any line $\ell_{t,\omega}$ with $|t| > 1$ lies entirely outside the unit disk, in which case a direct computation (using contour integration, for example) shows that

$$\mathcal{R}f(t, \omega) \equiv 0.$$

In particular, the Radon transform can vanish for $|t| > R$ without the function being zero outside the disk of radius R.

Recall that our main goal is to reconstruct functions from their Radon transforms. Mathematically, this should be equivalent to finding an inversion formula for the Radon transform. However, early reconstruction algorithms did not take this approach. Instead, they relied on direct algebraic techniques. We will discuss this approach in the next section.

Modern reconstruction algorithms are indeed based on inverting the Radon transform. For the rest of this section, we will give a brief and informal introduction into this topic, saving a rigorous treatment for later sections.

One natural approach to inverting the Radon transform is to try *back-projection*, which refers to the following: given a function f and a point $x \in \mathbb{R}^2$, take the average of the values of $\mathcal{R}f$ over all lines that pass through x. Now, for a given direction ω, the value of $t \in \mathbb{R}$ such that $\ell_{t,\omega}$ passes through x is given by $t = \langle x, \omega \rangle$. Thus, parametrizing the set of directions ω by

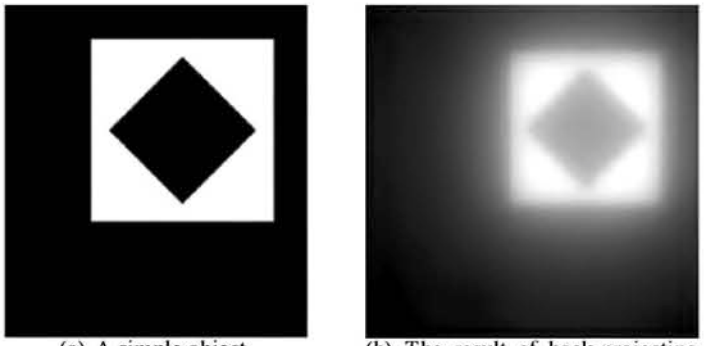
$$\omega = \omega(\theta) = (\cos\theta, \sin\theta)^T, \quad \theta \in [0, 2\pi],$$

we arrive at the following.

Definition 2 (Back-projection)**.** *Given the Radon transform $\mathcal{R}f$ of some function $f : \mathbb{R}^2 \to \mathbb{R}$, the* back-projection *of $\mathcal{R}f$ is the function $\tilde{f} : \mathbb{R}^2 \to \mathbb{R}$ defined by*

$$\tilde{f}(x) = \tfrac{1}{2\pi} \int_0^{2\pi} \mathcal{R}f(\langle x, \omega \rangle, \omega)\, d\theta. \tag{1}$$

The back-projection of $\mathcal{R}f$ does not recover f, but rather a blurred version thereof. (In fact, it is not the inverse of $\mathcal{R}$ but rather a type of transpose operation.) An example is provided in the following figure (imported from [1]).



(a) A simple object.

(b) The result of back-projecting the object's Radon transform.

The actual inversion formula for the Radon transform can instead be described as a 'filtered' back-projection of the Radon transform. To describe what this means, let us briefly introduce two main ingredients, which will be discussed in detail in later sections.

The first ingredient is the *central slice theorem*, which provides a connection between the Radon transform and the Fourier transform (defined and discussed later). This refers to the formula

$$\widetilde{\mathcal{R}f}(r,\omega) := \int_{-\infty}^{\infty} \mathcal{R}f(t,\omega)e^{-itr}\,dt = \hat{f}(r\omega) \tag{2}$$

for $r \in \mathbb{R}$ and $\omega \in \mathbb{S}^1$, expressing the one-dimensional Fourier transform of $\mathcal{R}f$ (in the affine parameter t) in terms of the Fourier transform of f.

The second ingredient is the *Fourier inversion formula*, which expresses a function in terms of its Fourier transform as follows:

$$f(x) = \tfrac{1}{4\pi^2} \int_{\mathbb{R}^2} e^{-ix\cdot\xi}\hat{f}(\xi)\,d\xi. \tag{3}$$

Changing to polar coordinates in (3) and using (2) and evenness of the Radon transform, one can derive the following *filtered back-projection formula*:

$$f(x) = \tfrac{1}{2\pi} \int_0^{\pi} |\nabla|\mathcal{R}f(\langle x,\omega\rangle,\omega)\,d\theta, \tag{4}$$

where $\omega(\theta) = (\cos\theta, \sin\theta)^T$ and

$$|\nabla|\mathcal{R}f(t,\omega) = \tfrac{1}{2\pi} \int_{-\infty}^{\infty} \widetilde{\mathcal{R}f}(r,\omega)e^{irt}|r|\,dr.$$

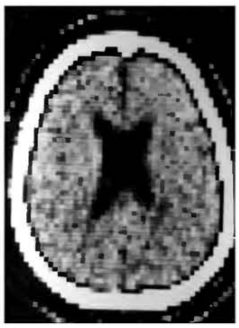
This final integral has the form of a *filter* applied to the Radon transform. The formula (4) then resembles the back-projection formula (1), applied to the filtered Radon transform rather than the Radon transform itself.

The filtered back-projection formula and Fourier inversion formula form the basis for the reconstruction algorithms that we will discuss in later sections. To get

there, we will first need to undertake a rigorous treatment of the Fourier transform and related topics. We will turn to this after our brief detour into algebraic reconstruction techniques in the next section.

Algebraic reconstruction techniques

In this section we will discuss some *algebraic reconstruction techniques*, which seek to produce approximations to a two-dimensional slice of an attenuation coefficient without any explicit reference to inverting the Radon transform. These techniques were used in some of the earliest reconstruction algorithms. While it is possible to obtain very good images with algebraic reconstruction techniques, this approach is typically much slower (in terms of computation techniques) than modern approaches based on exact reconstruction formulas.



The figure above (from [1]) is an example of an early X-ray CT image using the EMI scanner and an algebraic reconstruction technique. The dark edge around the skull is an artifact of the reconstruction algorithm called the *false subarachnoid space*.

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function that we wish to reconstruct. For simplicity, let us assume that f is supported in the square $[-1, 1]^2 \subset \mathbb{R}^2$. As before, we model the X-ray measurements we take as samples of the Radon transform $\mathcal{R}f$ of f. We suppose that we have samples of $\mathcal{R}f$ at the points $\{(t_i, \omega_i)\}_{i=1}^I$, which we denote by

$$p_i = \mathcal{R}f(t_i, \omega_i), \quad i = 1, \ldots, I.$$

We will produce an approximation of f of the form

$$f \approx \sum_{j=1}^{J} x_j b_j \tag{1}$$

for some collection of *basis functions* $\{b_j\}_{j=1}^J$ supported on $[-1, 1]^2$. Using linearity (and continuity) of the Radon transform, we should have

$$\mathcal{R}f \approx \sum_{j=1}^{J} x_j \mathcal{R}b_j.$$

Evaluating this at the points (t_i, ω_i) and denoting

$$r_{ij} = \mathcal{R}b_j(t_i, \omega_i),$$

we can therefore obtain the following linear system for the coefficients x_j:

$$p_i = \sum_{j=1}^{J} x_j r_{ij} \quad \text{for} \quad i = 1, \ldots, I.$$

In vector/matrix notation, we may express this as $p = rx$. We call r the measurement matrix. It has dimensions $I \times J$, where I is the number of measurements and J is the number of basis elements.

The problem now boils down to two main parts:

- The first part concerns finding a good choice of basis functions. For this, one should first of all choose a large enough set of functions that are sufficiently localized that we can faithfully reconstruct the functions f under consideration. In addition, we must be able to compute (or approximate) the measurement matrix r, whose entries consist of the Radon transforms of the basis functions evaluated at the sample points.
- The second main component involves solving the linear system $p = rx$. The challenges here lie in the fact that the system will be large in size and cannot be expected to be consistent in general.

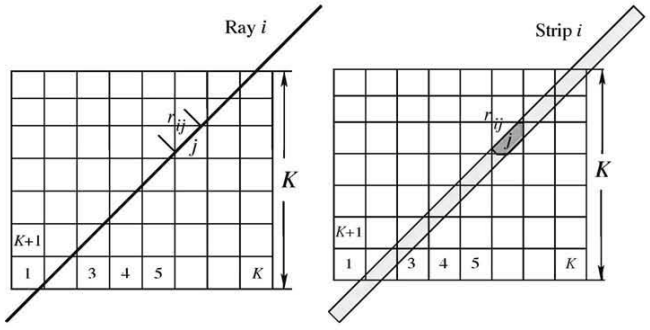
In what follows, we will consider the specific case of the *pixel basis*, defined as follows. Given a positive integer K, we subdivide $[-1, 1]^2$ into a $K \times K$ grid. We label the sub-squares (called *pixels*) S_j^K for $j = 1, \dots K^2$, and let b_j^K be the characteristic function of S_j.

To reconstruct a function f with the pixel basis as in (1), one could take x_j to equal the average of f on S_j^K. Indeed, if f is continuous and supported in $[-1, 1]^2$, then

$$f_K := \sum_{j=1}^{K^2} \left[\frac{1}{|S_j^K|} \int_{S_j^K} f \right] b_j^K \to f$$

uniformly as $K \to \infty$. whenever f is a continuous function supported in $[-1, 1]^2$.

To construct the exact measurement matrix with a one-dimensional X-ray beam, we should take r_{ij} to be the length of the intersection of the i^{th} ray with the j^{th} pixel. If the X-ray has a one-dimensional cross section, one should take the area of the corresponding strip.



(a) Using one-dimensional rays. (b) Using two-dimensional strips.

For an even simpler model, one can take $r_{ij} = 1$ if the center of the j^{th} pixel is contained in the i^{th} strip and 0 otherwise. In fact, this was used in early applications. This approach is computationally simpler (in the sense that you can compute r_{ij} 'as you go'), but it does not give a particularly accurate model for measurements.

Suppose now that we divide the square into 128×128 pixels and take 150 samples of the Radon transform at 128 equally spaced angles. Then the measurement matrix has dimensions $\approx 19000 \times 16000$. This is problematic for several reasons. First,

this matrix is too large to deal with numerically. This is ameliorated somewhat by observing that at least the matrix is *sparse*, in the sense that for each i, there are only about 128 values of j such that $r_{ij} \neq 0$. The second issue is that since the system $p = rx$ has more equations (≈ 19000) than unknowns (≈ 16000), we cannot expect the system to be consistent in general. In this case it is natural to look for the least squares solution to $p = rx$, which entails solving the normal system

$$r^T p = r^T r x, \tag{2}$$

which is guaranteed to be consistent. As discussed in the linear algebra primer, solutions to (2) minimize $\|p - rx\|$ over all choices of x, and if r has full rank, then (2) has a unique solution. Unfortunately, $r^T r$ still has dimensions $\approx 16000 \times 16000$, and this matrix is no longer guaranteed to be sparse. Thus it is not reasonable to try to solve (2) directly.

Our problem is therefore to find a computationally reasonable way to construct an approximate solution to $p = rx$. The method used in algebraic reconstruction techniques is related to the *Kaczmarz method*, or the *method of projections*, which we now describe.

We write the linear system $p = rx$ in the form

$$p_i = r_i \cdot x, \quad i = 1, \ldots, I,$$

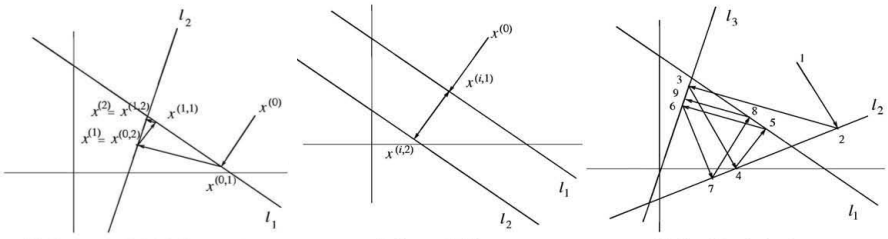
where r_i is the i^{th} row of r. Solving the system is therefore equivalent to simultaneous membership the I hyperplanes determined by the equations above. The Kaczmarz method is an iterative algorithm, described as follows.

1. Choose an initial vector x^0.
2. Project x^0 orthogonally onto the hyperplane $r_1 \cdot x = p_1$, yielding x^1.
   Project x^1 onto the hyperplane $r_2 \cdot x = p_2$, yielding x^2.
   ...
   Project x^{I-1} onto the hyperplane $r_I \cdot x = p_I$, yielding x^I.
3. Repeat Step 2 with x^I instead of x^0.

For the sake of completeness, we recall that the formula for the projection of a vector y onto the hyperplane $r_i \cdot x = p_i$ is given by

$$y - \frac{y \cdot r_i - p_i}{r_i \cdot r_i} r_i. \tag{3}$$

This algorithm is pictured in the figures below (from [1]), depicting some different scenarios that could arise in the case $I = 2$. In the figures, the notation $x^{(j,k)}$ corresponds to our x^{jI+k}.



We have the following result concerning this method.

Theorem 1. *If the system $p = rx$ is consistent, Kaczmarz iteration converges to a solution.*

Proof. Suppose $p = rz$. By construction and the Pythagorean theorem,

$$\|x^{k+1} - z\|^2 + \|x^k - x^{k+1}\|^2 = \|x^k - z\|^2. \tag{4}$$

Indeed, x^{k+1} is the projection of x^k onto some hyperplane $r_i \cdot x = p_i$, of which z is also a member. Thus $x^{k+1} - x^k$ is orthogonal to this hyperplane, and hence to $x^{k+1} - z$.

The equality above implies

$$\|x^{k+1} - z\| \leq \|x^k - z\|,$$

which implies that the sequence $\big\{\|x^k - z\|\big\} \subset [0, \infty)$ converges as $k \to \infty$. In particular, $\{x^k\}$ is a bounded sequence and hence converges along a subsequence (denoted x^{k_j}) to some limit x^*. By construction and the pigeonhole principle, infinitely many terms of the subsequence x^{k_j} must belong to one of the hyperplanes, say $r_\ell \cdot x = p_\ell$. As this further subsequence still converges to x^* and the hyperplane is closed, we deduce $r_\ell \cdot x^* = p_\ell$.

On the other hand, by (4) and convergence of $\big\{\|x^k - z\|\big\}$, we have $\|x^k - x^{k+1}\| \to 0$, and so

$$\|x^{k_j+1} - x^*\| \leq \|x^{k_j+1} - x^{k_j}\| + \|x^{k_j} - x^*\| \to 0.$$

Applying the argument in the preceding paragraph, we then deduce $r_{\ell+1} \cdot x^* = p_{\ell+1}$. Repeating this argument I times implies $r_i \cdot x^* = p_i$ for each $i = 1, \ldots, I$, so that x^* is a solution to $rx = p$.

It remains to show that the full sequence x^k converges to x^*. For this, observe that since $\big\{\|x^k - z\|\big\}$ converges for any solution z, we have in particular that $\big\{\|x^k - x^*\|\big\}$ converges to some limit. However, since the limit along the subsequence x^{k_j} is zero, that limit must be zero. □

In imaging applications, the system is generally *not* consistent. Moreover, because the system is so large ($I \sim 19000$), we cannot practically compute more than a few complete iterations in a reasonable time. In fact, one observes that the quality of the reconstructed image only improves for a few iterates, before becoming worse again (perhaps due to noise in the data and inconsistencies arising from the approximation of the measurement matrix).

This algorithm naturally takes advantage of sparseness in the measurement matrices. Indeed, computing orthogonal projections requires primarily computing inner products with the rows r_i, which becomes simple when r_i has many zeros. Moreover, we can see from (3) that in passing from x^k to x^{k+1}, the only entries that change are those at indices at which the appropriate row has a nonzero entry.

In the case that $rx = p$ has more than one solution, we may use Kaczmarz iteration to find a solution of minimal norm.

Lemma 1. *With $x^0 = 0$, the solution x^* obtained by Kaczmarz iteration has minimal norm.*

Proof. The image $R(r^T)$ is invariant under Kaczmarz iteration. This follows from (3) and the fact that r_i belongs to the image of r^T (it is the image of the i^{th} standard basis vector). Thus if we choose $x^0 = 0 \in R(r^T)$, we deduce that $x^k \in R(r^T)$ for all k and hence $x^* \in R(r^T)$. As $R(r^T) = [\ker(r)]^\perp$, we deduce that

$$\|x^* + v\|^2 = \|x^*\|^2 + \|v\|^2 \geq \|x^*\|^2 \quad \text{for any} \quad v \in \ker(r).$$

As any solution to $rx = p$ is of the form $x^* + v$ for some $v \in \ker(r)$, the result follows. □

We will discuss imaging artifacts in a later section. These frequently appear as rapid oscillations in the image. To minimize such effects, we may wish to construct solutions with the smallest possible variation. To make this precise, define the *average value* of a vector $x \in \mathbb{R}^J$ by

$$\mu_x = \tfrac{1}{J}\langle e, x\rangle, \quad \text{where} \quad e = (1, \dots, 1)^T.$$

We then define the *variance* of $x \in \mathbb{R}^J$ by

$$\sigma_x^2 = \|x - \mu_x e\|^2.$$

Proposition 1. *If $e \in span\{r_i\}$, then finding a minimal variance solution to $rx = p$ is equivalent to finding a minimal norm solution.*

Proof. It suffices to show that if $rx = p$, then $\|x\|^2$ and σ_x^2 differ by a fixed constant. To this end, first observe that if $e = \sum \alpha_i r_i$, then

$$\langle e, x\rangle = \sum \alpha_i \langle r_i, x\rangle = \sum \alpha_i p_i.$$

Thus

$$\|x - \tfrac{1}{J}\langle e, x\rangle e\|^2 = \|x\|^2 - \tfrac{1}{J}\langle e, x\rangle^2 = \|x\|^2 - \tfrac{1}{J}\left(\sum \alpha_i p_i\right)^2,$$

and the result follows. □

We will conclude this section by describing a modification of the basic Kaczmarz algorithm that may reduce the effects of noise and modeling error. In particular, we may introduce some *relaxation parameters* to diminish noise and speed up convergence. This involves introducing factors $\{\lambda_k\} \subset [0, 2]$ so that instead of exactly projecting onto the hyperplane $r_i \cdot x = p_i$ as in (3), we instead use the update

$$y \mapsto y - \lambda_k \frac{y \cdot r_i - p_i}{r_i \cdot r_i} r_i$$

in the k^{th} complete iteration. In particular, $\lambda_k = 0$ means we do not update at all, while $\lambda_k = 1$ recovers the original algorithm. If $0 < \lambda_k < 1$, then we remain on the same side of the hyperplane as the input, while if $1 < \lambda_k < 2$ we end up on the opposite side. The case $\lambda_k = 2$ means we reflect across the hyperplane.

One can prove that as long as $\lambda_k \in (0, 2)$ for all k and the system has a solution, the modified algorithm also converges to a solution. Furthermore, if $x^0 = 0$ then the limit will again be a minimal norm solution. In fact, if we let $\lambda_k \to 0$ as $k \to \infty$, then we may obtain convergence even when the system has no solution.

Using such modified algorithms, one can find approximate solutions that optimize various criteria (e.g. minimizing norm or improving contrast). See, for example, Y. Censor and G. Herman, *On some optimization techniques in image reconstruction from projections*, Appl. Numer. Math. **3** (1987), no. 5, 365–391.

We conclude our discussion with the following observation. In taking the X-ray measurements, adjacent rays will produce measurement vectors r_i and r_{i+1} that are nearly parallel. As a result, when considering subsequent projections of our approximate solutions, our approximation will only change very slightly. In fact, the changes could be lost completely in noise and rounding errors. To improve this situation, one would like to order the hyperplanes so that they are as close to orthogonal as possible. One way this can be done is to order the measurements randomly to minimize the expected correlation between successive measurements.

The Fourier transform

We begin by defining the Fourier transform for the class of L^1 functions, where the definition as an integral transform makes sense. Later we will extend our definition to L^2, on which the Fourier transform has many nice properties. We will focus in this section on the one-dimensional case. We will discuss extensions higher dimensions briefly at the end.

Definition 1. *For $f \in L^1(\mathbb{R})$, we define the* Fourier transform *of f to be the function $\hat{f} : \mathbb{R} \to \mathbb{C}$ defined by*

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-ix\xi}\, dx. \tag{1}$$

Remark 1. *Here $e^{ix\xi}$ is the complex exponential, given by*

$$e^{ix\xi} = \cos(x\xi) + i\sin(x\xi).$$

We view this as an oscillatory state with frequency $\frac{\xi}{2\pi}$. Equation (1) *may be interpreted as computing the component of f in the direction of this state.*

For $f \in L^1(\mathbb{R})$, we can show function $\hat{f} : \mathbb{R} \to \mathbb{C}$ is continuous in ξ and tends to zero as $|\xi| \to \infty$. In particular, we may view the Fourier transform $\mathcal{F}$ as a linear transformation $\mathcal{F} : L^1(\mathbb{R}) \to C_0(\mathbb{R})$.

Proposition 1 (Riemann–Lebesgue lemma)**.** *For $f \in L^1$, we have $\hat{f} \in C_0(\mathbb{R})$.*

Remark 2. *We will prove this below. The Riemann–Lebesgue lemma demonstrates an important fact about the Fourier transform, namely, that it interchanges decay (in this case, the assumption that $f \in L^1$) and regularity (in this case, the fact that $\hat{f}$ is continuous).*

Our first main goal is to prove the *Fourier inversion formula*, which allows us to recover a function from its Fourier transform:

Theorem 1 (Fourier inversion)**.** *Let $f \in L^1(\mathbb{R})$. Suppose in addition that f is uniformly continuous. If $\hat{f} \in L^1(\mathbb{R})$, then*

$$f(x) = \tfrac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi) e^{ix\xi}\, d\xi.$$

Before we can prove Theorem 1, we will need to collect a few lemmas. The first lemma, which consists of working out the Fourier transform in the special case of a Gaussian, will introduce us to some useful algebraic properties of this transformation.

Lemma 1. *Let $f(x) = e^{-ax^2}$ for some $a > 0$. Then*

$$\hat{f}(\xi) = \sqrt{\tfrac{\pi}{a}} e^{-\xi^2/4a}.$$

Proof. We first observe that with $f(x) = e^{-ax^2}$, we have

$$f'(x) = -2axf(x).$$

Now let us take the Fourier transform of both sides of this equality. First, using integration by parts, we have

$$\mathcal{F}(f')(\xi) = \int f'(x) e^{-ix\xi}\, dx = i\xi \int f(x) e^{-ix\xi}\, dx = i\xi \hat{f}(\xi). \tag{2}$$

On the other hand,

$$\int xf(x)e^{-ix\xi}\,dx = i\tfrac{d}{d\xi}\int f(x)e^{-ix\xi}\,dx = i\tfrac{d}{d\xi}\hat{f}(\xi). \tag{3}$$

Combining these, we deduce

$$\tfrac{d}{d\xi}\hat{f}(\xi) = -\tfrac{\xi}{2a}\hat{f}(\xi).$$

This implies

$$\hat{f}(\xi) = \hat{f}(0)e^{-\xi^2/4a} = \left[\int_{\mathbb{R}} e^{-ax^2}\,dx\right]e^{-\xi^2/4a} = \sqrt{\tfrac{\pi}{a}}e^{-\xi^2/4a}, \tag{4}$$

as desired. □

Remark 3. *Equations* (2) *and* (3) *are also related to the fact that the Fourier transform interchanges decay and regularity of functions.*

We next have the following identity.

Lemma 2. *Suppose* $f, g \in L^1(\mathbb{R})$. *Then*

$$\int f(x)\hat{g}(x)\,dx = \int \hat{f}(y)g(y)\,dy.$$

Proof. Both integrals converge absolutely, as the Fourier transform of an L^1 function is bounded. Then, by Fubini's theorem, both of the quantities above are equal to

$$\iint f(x)e^{-ixy}g(y)\,dy\,dx.$$

□

Let us turn to the proof of Theorem 1.

Proof of Theorem 1. Let $f \in L^1(\mathbb{R})$ be uniformly continuous. We will recover the value of f at a point $x \in \mathbb{R}$ by taking the limit of averages of f against rescaled Gaussians. More precisely, let us define the family of functions $K_b(x)$ by

$$K_b(x) = \sqrt{\tfrac{b}{\pi}}e^{-bx^2},$$

which (in light of (4)) obey $\int K_b\,dx = 1$ for all $b > 0$. The integrals

$$\int f(y)K_b(x-y)\,dy \tag{5}$$

represent weighted averages of f around the point x. As $b \to \infty$, these averages are concentrated in smaller and smaller neighborhoods of x, and accordingly we claim the following:

$$f(x) = \lim_{b\to\infty}\int f(y)K_b(x-y)\,dy. \tag{6}$$

To prove (6), we use $\int K_b\,dx \equiv 1$ and a change of variables to write

$$\begin{aligned} f(x) - \int f(y)K_b(x-y)\,dy &= \int [f(x) - f(x-y)]K_b(y)\,dy \\ &= \int [f(x) - f(x - \tfrac{y}{\sqrt{b}})]K_1(y)\,dy. \end{aligned} \tag{7}$$

We now observe that since $f \in L^1$ and f is uniformly continuous, we have $f \in L^\infty$. Furthermore, by continuity of f we have that the integrand tends to zero as $b \to \infty$ for all $y \in \mathbb{R}$. Thus the dominated convergence theorem implies (6).

We now consider the right-hand side of (6). By a change of variables and the fact that K_b is even, we can first rewrite the integral as

$$\int f(y+x)K_b(y)\,dy.$$

Now, by Lemma 1 (with $b = \frac{1}{4a}$), we see that K_b is the Fourier transform of the function $\frac{1}{2\pi}e^{-x^2/4b}$. We are then in a position to apply Lemma 2. This requires that we compute the Fourier transform of the function $y \mapsto f(y+x)$, which (by a change of variables) we find equals $e^{ix\xi}\hat{f}(\xi)$. Thus Lemma 2 now yields

$$\int f(x-y)K_b(y)\,dy = \tfrac{1}{2\pi}\int e^{-\xi^2/4b}e^{ix\xi}\hat{f}(\xi)\,d\xi.$$

To complete the proof of the inversion formula, we recall that $\hat{f} \in L^1$ by assumption, so that by the dominated convergence theorem, this final quantity tends to $\frac{1}{2\pi}\int e^{-ix\xi}\hat{f}(\xi)\,d\xi$ as $b \to \infty$. □

Just as we view the Fourier transform $\mathcal{F}$ as a mapping $\mathcal{F} : L^1(\mathbb{R}) \to C_0(\mathbb{R})$, we may view the operator appearing in the Fourier inversion formula as a linear transformation, denoted $\mathcal{F}^{-1}$, and the Fourier inversion formula reads $f = \mathcal{F}^{-1}\hat{f}$ for suitable f. As the two operators have essentially the same form, the Riemann–Lebesgue lemma implies $\mathcal{F}^{-1} : L^1(\mathbb{R}) \to C_0(\mathbb{R})$. The interpretation of $\mathcal{F}^{-1}$ as an inverse of $\mathcal{F}$ will be clarified after we have introduced the L^2 theory.

The quantity appearing in (5) is an example of a *convolution*, and can be written as

$$K_b * f(x).$$

The argument we used (to show $K_b * f(x) \to f(x)$) is an example of an *approximate identity* argument, which we will discuss in more detail in a later section.

Definition 2. *For $f, g \in L^1(\mathbb{R})$, the* convolution *of f and g is the function $f * g : \mathbb{R} \to \mathbb{C}$ defined by*

$$f * g(x) = \int_{\mathbb{R}} f(x-y)g(y)\,dy.$$

The Fourier transform interacts simply with the convolution product.

Proposition 2. *For $f, g \in L^1(\mathbb{R})$, we have $f * g \in L^1(\mathbb{R})$, and*

$$\mathcal{F}[f * g](\xi) = \hat{f}(\xi)\hat{g}(\xi).$$

Proof. To prove $f * g \in L^1$, we use Fubini's theorem and translation invariance of Lebesgue measure:

$$\begin{aligned}\int |f * g(x)|\,dx &\le \iint |f(x-y)|\,|g(y)|\,dy\,dx \\ &\le \int |g(y)|\left[\int |f(x-y)|\,dx\right]dy \\ &\le \|f\|_{L^1}\int |g(y)|\,dy \le \|f\|_{L^1}\|g\|_{L^1}.\end{aligned}$$

For the identity, we compute directly, again using translation invariance:

$$\begin{aligned}\mathcal{F}[f*g](\xi) &= \iint e^{-ix\xi} f(x-y)g(y)\,dy\,dx \\ &= \int g(y)e^{-iy\xi}\left[\int e^{-i(x-y)\xi} f(x-y)\,dx\right] dy = \hat{f}(\xi)\hat{g}(\xi).\end{aligned}$$

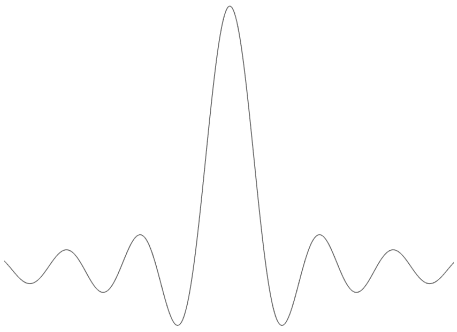
□

With some of the basic theory in place, let us work out a simple illustrative example, which in turn will be used in the proof of the Riemann–Lebesgue lemma.

Example 1. *Let r_1 be the characteristic function of the interval $(-1,1)$. Then*

$$\hat{r}_1(\xi) = \int_{-1}^{1} e^{-ix\xi}\,dx = \left[\tfrac{e^{-ix\xi}}{-i\xi}\right]_{x=-1}^{1} = 2\tfrac{\sin\xi}{\xi}.$$

The function $\xi \mapsto \frac{\sin\xi}{\xi}$ is called the sinc *function, which plays an important role in many signal processing applications. Beware that in many other references, the* sinc *function is instead defined as $\frac{\sin\pi x}{\pi x}$. It is pictured below.*



Observe that $r_1 \in L^1(\mathbb{R})$ and $\hat{r}_1 \in C_0(\mathbb{R})$, but $\hat{r}_1$ is not *absolutely integrable. The discontinuity of r_1 at the endpoints of the interval $(-1,1)$ leads to the characteristic $1/|\xi|$ type decay of the Fourier transform.*

More generally, the Fourier transformation of $\chi_{(a,b)}$ is given by

$$\hat{\chi}_{(a,b)}(\xi) = \tfrac{1}{i\xi}[e^{-ia\xi} - e^{-ib\xi}].$$

The previous example illustrates the Riemann–Lebesgue lemma. In fact, it also provides a *proof* of the Riemann–Lebesgue lemma. In particular, we can see that an arbitrary finite linear combination of characteristic functions of intervals satisfies the conclusions of the Riemann–Lebesgue lemma. Using the general fact in analysis that arbitrary L^1 functions may be approximated in L^1-norm by such linear combinations, we can deduce the result for general L^1 functions.

One can prove a wide range of results demonstrating the interchange of decay/regularity between a function and its Fourier transform. We will not go into too much detail, but let us mention a few representative results, which rely on identities such as those appearing in (2) and (3):

(i) If f has j integrable derivatives, then $\hat{f}$ decays like $(1+|\xi|)^{-j}$.
(ii) The rate of decay of the Fourier transform of a continuous L^1 function may be arbitrarily slow (if the function is very 'noisy').

(iii) If $f \in L^1$ and $[1+|\xi|]^j \hat{f}(\xi)$ is integrable, then f is continuous and has j continuous derivatives.
(iv) If $[1+|x|]^j f$ is integrable, then $\hat{f}$ has j continuous derivatives that all tend to zero as $|\xi| \to \infty$.
(v) The *Paley–Wiener theorem* is a classical result, stating essentially that if the Fourier transform of f decays exponentially, then f can be viewed as the restriction to the line of an analytic function function defined on some strip in the complex plane. This result implies that a function and its Fourier transform cannot both have bounded support.

Let us prove only (i), since we will use it in what follows.

Proof of (i). Let us consider the case $j = 1$; the extension to larger values of j is similar.

We have already seen that $f \in L^1$ implies $\hat{f} \in L^\infty$. Now suppose further that $f' \in L^1$. Then, computing as in (2), we have

$$\hat{f}(\xi) = \tfrac{1}{i\xi}\mathcal{F}(f')(\xi).$$

As $\mathcal{F}(f') \in L^\infty$, this shows that $\hat{f}$ decays like $|\xi|^{-1}$ as $|\xi| \to \infty$. □

♣ ♣ ♣

Our next main goal is the extension of the Fourier transform to L^2. The key fact that underlies this extension is the Parseval formula:

Theorem 2 (Parseval formula)**.** *Suppose $f \in L^1 \cap L^2$. Then $\hat{f} \in L^2$, and*

$$\int |f(x)|^2\,dx = \tfrac{1}{2\pi}\int |\hat{f}(\xi)|^2\,d\xi.$$

Let us first show that this holds for functions with some additional smoothness.

Lemma 3. *Parseval's formula holds $f \in L^1 \cap L^2$ satisfying $f' \in L^1 \cap L^2$ and $f'' \in L^1$.*

Proof. By item (i) above, we have that $\hat{f}$ decays like $(1+|\xi|)^{-2}$, so that $\hat{f} \in L^1$. Note also that $f' \in L^2$ also implies that f is uniformly continuous.

We begin by writing

$$\int |f(x)|^2\,dx = \int f(x)\bar{f}(x)\,dx = f * g(0), \quad \text{where} \quad g(x) = \bar{f}(-x).$$

We claim that the Fourier inversion formula applies to $f * g$. Indeed, $f * g$ is in L^1 by Proposition 2. To see that $f * g$ is uniformly continuous, note that

$$f * g(x_2) - f * g(x_1) = \int [f(x_2 - y) - f(x_1 - y)]g(y)\,dy.$$

Thus uniform continuity follows from the facts that $f \in L^1$ and f is uniformly continuous. Finally, to see that $\mathcal{F}[f * g] \in L^1$, we apply Proposition 2, which shows

$$\mathcal{F}[f * g](\xi) = \hat{f}(\xi)\hat{g}(\xi) = \hat{f}(\xi)\overline{\hat{f}(\xi)} = |\hat{f}(\xi)|^2.$$

Since $\hat{f} \in L^1 \cap L^\infty$, we have $|\hat{f}|^2 \in L^1$ as desired.

We may now apply Theorem 1 and the above identity to $f * g$ to deduce

$$f * g(0) = \tfrac{1}{2\pi}\int \mathcal{F}[f * g](\xi)\,d\xi = \tfrac{1}{2\pi}\int |\hat{f}(\xi)|^2\,d\xi.$$

Combining this with the above identity now yields the result. □

We will recover the Parseval formula for $f \in L^1 \cap L^2$ by approximation.

Lemma 4. *For any $f \in L^1 \cap L^2$, there exists a sequence of functions $f_n \in L^1 \cap L^2$ such that $f_n' \in L^1 \cap L^2$, $f_n'' \in L^1$, and $f_n \to f$ in L^1- and L^2-norms.*

Proof. We can use an approximate identity argument as in the proof of the Fourier inversion formula. Recall the functions

$$K_b(x) = \sqrt{\tfrac{b}{\pi}} e^{-bx^2} \quad \text{for} \quad b > 0.$$

Now consider the sequence

$$f_n(x) = f * K_n(x).$$

By the linearity of the convolution product, we can derive that $f_n' = f * K_n'$ and $f_n'' = f * K_n''$. In particular, using

$$\|f * g\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1},$$

we see that $f_n, f_n', f_n'' \in L^1$. To see that $f_n, f_n' \in L^2$, we first observe

$$\|f * g\|_{L^\infty} \leq \sup_x \int |f(x-y)| |g(y)| \, dy \leq \|f\|_{L^1} \|g\|_{L^\infty}.$$

Therefore (using $K_n', K_n'' \in L^\infty$) we have $f_n, f_n' \in L^1 \cap L^\infty \subset L^2$.

The proof that $f_n \to f$ in L^1- and L^2-norms is similar to the argument in the proof of the Fourier inversion formula (where we showed $f_n(x) \to f(x)$ pointwise). For now, let us take the result for granted. We will discuss such arguments in more detail in the next section. □

We turn to the proof of the Parseval formula.

Proof of Theorem 2. Let $f \in L^1 \cap L^2$ and choose f_n as in Lemma 4. By Lemma 3 and the linearity of the Fourier transform, we have

$$\|f_n - f_m\|_{L^2}^2 = \tfrac{1}{2\pi} \|\hat{f}_n - \hat{f}_m\|_{L^2}^2 \quad \text{for all} \quad m, n \geq 1.$$

As $f_n \to f$ in L^2-norm, we deduce that $\{\hat{f}_n\}$ is a Cauchy sequence in L^2 and hence converges to some limit g. However, because $f_n \to f$ in L^1, we can deduce that $\hat{f}_n \to \hat{f}$ pointwise. Thus, by uniqueness of limits, we have $g = \hat{f}$. Therefore, using Lemma 3 again, we derive

$$\|f\|_{L^2}^2 = \lim_{n \to \infty} \|f_n\|_{L^2}^2 = \lim_{n \to \infty} \tfrac{1}{2\pi} \|\hat{f}_n\|_{L^2}^2 = \tfrac{1}{2\pi} \|\hat{f}\|_{L^2}^2,$$

as desired. □

We now turn to the definition of the Fourier transform on L^2.

Definition 3. *Let $f \in L^2(\mathbb{R})$. We define the* Fourier transform *of f by*

$$\hat{f}(\xi) = \lim_{R \to \infty} \int_{-R}^{R} f(x) e^{-ix\xi} \, dx, \tag{8}$$

where the limit is taken in the L^2 sense.

To make such a definition, we must check that the claimed limit always exists.

Proposition 3. *For $f \in L^2$, the limit in* (8) *exists.*

Proof. Fix $f \in L^2$. Observe that $\chi_{[-R,R]} f \in L^1 \cap L^2$ for any $R > 0$. Writing $g_R = \mathcal{F}(\chi_{[-R,R]} f)$, Parseval's formula implies

$$\|g_{R_2} - g_{R_1}\|_{L^2}^2 = 2\pi \int_{R_1 < |x| < R_2} |f(x)|^2 \, dx$$

for any $R_2 > R_1 > 0$. This, together with the dominated convergence theorem, shows that $\{g_R\}$ is Cauchy in L^2 as $R \to \infty$. Thus g_R has an L^2 limit as $R \to \infty$, as desired. □

We collect the main properties about the Fourier transform on L^2 in the following theorem.

Theorem 3. *The Fourier transform extends to a bounded linear transformation $\mathcal{F} : L^2 \to L^2$. Parseval's formula holds for all $f \in L^2$, and consequently the Fourier transform satisfies the Plancherel formula*

$$\langle f, g \rangle = \tfrac{1}{2\pi} \langle \hat{f}, \hat{g} \rangle \quad \textit{for all} \quad f, g \in L^2.$$

The Fourier inversion formula holds in the following sense:

$$f(x) = \lim_{R \to \infty} \tfrac{1}{2\pi} \int_{-R}^{R} \hat{f}(\xi) e^{ix\xi} \, d\xi, \tag{9}$$

where the limit is taken in the L^2-norm.

Proof. We have described how to extend the Fourier transform to L^2 functions. This transformation is linear by construction. The Parseval formula follows from the definition and the Parseval formula on $L^1 \cap L^2$. Indeed,

$$\tfrac{1}{2\pi} \|\hat{f}\|_{L^2}^2 = \lim_{R \to \infty} \tfrac{1}{2\pi} \|\mathcal{F}\big\{\chi_{[-R,R]} f\big\}\|_{L^2}^2 = \lim_{R \to \infty} \|\chi_{[-R,R]} f\|_{L^2}^2 = \|f\|_{L^2}^2.$$

Boundedness (which is equivalent to continuity by linearity) then follows from Parseval's identity. Parseval's identity also implies Plancherel's identify. To see this, let $f, g \in L^2$ and $t \in \mathbb{R}$. Then

$$\begin{aligned} \|f\|^2 + 2t \operatorname{Re}\langle f, g \rangle + t^2 \|g\|_{L^2}^2 &= \|f + tg\|_{L^2}^2 \\ &= \tfrac{1}{2\pi} \|\hat{f} + t\hat{g}\|_{L^2}^2 \\ &= \tfrac{1}{2\pi} \big[\|\hat{f}\|_{L^2}^2 + 2t \operatorname{Re}\langle \hat{f}, \hat{g} \rangle + t^2 \|\hat{g}\|_{L^2}^2 \big]. \end{aligned}$$

As this holds for all $t \in \mathbb{R}$, we find $\operatorname{Re}\langle f, g \rangle = \frac{1}{2\pi} \operatorname{Re}\langle \hat{f}, \hat{g} \rangle$. The same argument using it instead of t shows that the imaginary parts are equal as well, and so we derive Plancherel's identity.

Finally, we consider the Fourier inversion formula. We let $f \in L^2$, so that $\hat{f} \in L^2$. For any $R > 0$, we have that $\chi_R \hat{f} \in L^1 \cap L^2$, where χ_R denotes the characteristic function of $[-R, R]$. Let us define

$$g_R(x) = \tfrac{1}{2\pi} \int_{-R}^{R} \hat{f}(\xi) e^{ix\xi} \, d\xi.$$

We first wish to show that $\hat{g}_R = \chi_R \hat{f}$. This is 'obvious' by the Fourier inversion formula, but it is subtle here because we do not yet know in which spaces g_R lives. By the Riemann–Lebesgue lemma, we have that $g_R \in C_0(\mathbb{R})$, but we cannot expect $g_R \in L^1$ in general. We can, however, see that $g_R \in L^2$. Indeed,

$$g_R(x) = \tfrac{1}{2\pi} \mathcal{F}\big\{\chi_R \hat{f}\big\}(-x).$$

so that by Parseval's formula we have

$$\|g_R\|_{L^2}^2 = \tfrac{1}{2\pi}\int_{|\xi|\le R} |\hat{f}(\xi)|^2\,d\xi \le \tfrac{1}{2\pi}\|\hat{f}\|_{L^2}^2.$$

Once we establish $\hat{g}_R = \chi_R\hat{f}$, we can complete the proof of (9) by using Parseval's theorem. In particular, we have

$$\|g_R - f\|_{L^2}^2 = \tfrac{1}{2\pi}\|\hat{g}_R - \hat{f}\|_{L^2}^2 = \tfrac{1}{2\pi}\int_{|\xi|>R} |\hat{f}(\xi)|^2\,d\xi \to 0 \quad \text{as} \quad R\to\infty.$$

We turn to the evaluation of $\hat{g}_R$. *This proof may be skipped in lecture.* Recalling the definition of the Fourier transform on L^2, we fix $T>0$ and compute

$$\begin{aligned}\int_{-T}^{T} e^{-ix\xi} g_R(x)\,dx &= \tfrac{1}{2\pi}\int_{-R}^{R}\left[\int_{-T}^{T} e^{-ix(\xi-\eta)}\,dx\right]\hat{f}(\eta)\,d\eta \\ &= \tfrac{1}{2\pi}\int \{\chi_R\hat{f}\}(\eta)\cdot\mathcal{F}[\chi_T](\xi-\eta)\,d\eta \\ &= \tfrac{1}{2\pi}\chi_R\hat{f} * \mathcal{F}[\chi_T](\xi).\end{aligned}$$

Writing

$$K_T(\xi) = \tfrac{1}{2\pi}\mathcal{F}[\chi_T](\xi) = \tfrac{1}{\pi}\cdot T\,\mathrm{sinc}(T\xi),$$

the problem has now been reduced to another approximate identity type argument, in the sense that we would like to prove

$$\chi_R\hat{f} * K_T \to \chi_R\hat{f} \quad \text{as} \quad T\to\infty. \tag{10}$$

This is a bit subtle, however, because the sinc function fails to be absolutely integrable. To resolve this point, we observe that we already know (by the definition of the Fourier transform on L^2) that convergence in L^2 holds as $T\to\infty$:

$$\chi_R\hat{f} * K_T \to \hat{g}_R \quad \text{in} \quad L^2 \quad \text{as} \quad T\to\infty.$$

Thus it suffices to identify the limit as $\chi_R\hat{f}$, for which almost everywhere convergence in (10) will be enough. For this, we argue as follows.

Our first goal is the following: there exists a sequence $T_n\to\infty$ such that for almost every x, we have

$$\lim_{n\to\infty}\tfrac{1}{\pi}\int_{|y|\le M}\mathrm{sinc}(y)[\chi_R\hat{f}](x-\tfrac{y}{T_n})\,dy = \tfrac{1}{\pi}[\chi_R\hat{f}(x)]\int_{|y|\le M}\mathrm{sinc}(y)\,dy \tag{11}$$

for all $M\in\mathbb{N}$. To this end, we note that for a given $M\in\mathbb{N}$, we may use Minkowski's integral inequality and continuity of translatons in L^2 to deduce

$$\begin{aligned}&\left\|\int_{|y|\le M}\mathrm{sinc}(y)\big\{[\chi_R\hat{f}](x-\tfrac{y}{T}) - [\chi_R\hat{f}](x)\big\}\,dy\right\|_{L^2_x} \\ &\le \big\|\chi_M(y)\,\mathrm{sinc}(y)\|\chi_R\hat{f}(x-\tfrac{y}{T}) - \chi_R\hat{f}(x)\|_{L^2_x}\big\|_{L^1_y} \\ &\le C\log M \sup_{y\in[-M,M]}\|\chi_R\hat{f}(\cdot-\tfrac{y}{T}) - \chi_R\hat{f}(\cdot)\|_{L^2_x} \\ &\to 0 \quad \text{as} \quad T\to\infty.\end{aligned}$$

We now rely on the fact that L^2 convergence implies almost everywhere convergence along a subsequence. In particular, we may find a sequence $T_n^1\to\infty$ and a full measure set S_1 so that (11) holds along this sequence for $x\in S_1$ and with $M=1$. Passing to a subsequence T_n^2 of of T_n^1, we may find another full measure set S_2 so

that (11) holds along this subsequence for $x \in S_2$ and with $M = 2$. Proceeding in this way, we then take the sequence $T_n = T_n^n$ and the full measure set $S = \cap_M S_M$ and obtain that (11) holds for every $M \in \mathbb{N}$ and for every $x \in S$. Without loss of generality, we may also assume that $\hat{f}$ is finite on S.

Now let $x \in S$ and take $\varepsilon > 0$. As $\int \operatorname{sinc}(y)\, dy = \pi$ (a fact proven by contour integration), we may choose $M \in \mathbb{N}$ sufficiently large to guarantee

$$\left| 1 - \tfrac{1}{\pi} \int_{|y| \leq M} \operatorname{sinc}(y)\, dy \right| < \frac{\varepsilon}{1 + |[\chi_R \hat{f}](x)|}.$$

On the other hand, as $\operatorname{sinc} \in L^2$ we may use Cauchy–Schwarz and the dominated convergence theorem and choose M possibly even larger to guarantee

$$\sup_{T>0} \left| \int_{|y|>M} \operatorname{sinc}(y) [\chi_R \hat{f}](x - \tfrac{y}{T})\, dy \right| \leq \| \operatorname{sinc} \|_{L^2(|y|>M)} \| \chi_R \hat{f} \|_{L^2(\mathbb{R})} < \varepsilon.$$

Finally, choosing $x \in S$, we obtain

$$\begin{aligned}
&\left| \tfrac{1}{\pi} \int \operatorname{sinc}(y) \chi_R \hat{f}(x - \tfrac{y}{T_n}) - [\chi_R \hat{f}](x) \right| \\
&\leq \left| \tfrac{1}{\pi} \int_{|y| \leq M} \operatorname{sinc}(y) [\chi_R \hat{f}](x - \tfrac{y}{T_n})\, dy - \tfrac{1}{\pi} [\chi_R \hat{f}](x) \int_{|y| \leq M} \operatorname{sinc}(y)\, dy \right| \\
&\quad + \left| [\chi_R \hat{f}](x) \left(\tfrac{1}{\pi} \int_{|y| \leq M} \operatorname{sinc}(y)\, dy - 1 \right) \right| \\
&\quad + \left| \int_{|y|>M} \operatorname{sinc}(y) [\chi_R \hat{f}](x - \tfrac{y}{T_n})\, dy \right| \\
&< 2\varepsilon + o(1) \quad \text{as} \quad n \to \infty.
\end{aligned}$$

We conclude that $\chi_R \hat{f} * K_{T_n} \to \chi_R \hat{f}$ almost everywhere as $n \to \infty$. As $\chi_R \hat{f} * K_T \to \hat{g}_R$ in L^2, we conclude that $\chi_R \hat{f} = \hat{g}_R$, as was needed to show. □

We have now constructed the Fourier transform $\mathcal{F}$ as a bounded linear transformation from L^2 to L^2. Furthermore, we have the inverse map $\mathcal{F}^{-1}$, e.g. by the formula (9).

To conclude this section, let us briefly discuss the extension to higher dimensions. The definition of the Fourier transform in higher dimensions is similar to the $1d$ version, namely,

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-ix \cdot \xi} f(x)\, dx,$$

defined initially for $f \in L^1$. The Fourier inversion formula is similar, as well. It takes the form

$$f(x) = \tfrac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix \cdot \xi} \hat{f}(\xi)\, d\xi$$

for suitable f. Similar phenomena occur regarding the interchange of decay and regularity.

The Parseval formula in higher dimensions takes the form

$$\|f\|_{L^2}^2 = \tfrac{1}{(2\pi)^n} \|\hat{f}\|_{L^2}^2,$$

which again plays a key role in the development of the L^2 theory.

The convolution product

The *convolution product* was introduced in the previous section. In particular, we defined

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy$$

for $f, g \in L^1$ and obtained $f * g \in L^1$. In fact, the integral converges in other settings, e.g. if $f \in L^1$ and g is bounded. These facts may be expressed by using the estimates

$$\|f * g\|_{L^1} \leq \|f\|_{L^1}\|g\|_{L^1} \quad \text{and} \quad \|f * g\|_{L^\infty} \leq \|f\|_{L^1}\|g\|_{L^\infty},$$

both of which we already proved in the previous section. More generally, one has the estimate

$$\|f * g\|_{L^p} \leq \|f\|_{L^1}\|g\|_{L^p} \quad \text{for any} \quad 1 \leq p \leq \infty,$$

which is an instance of *Young's convolution inequality.*

Example 1. *Let $g = \frac{1}{|B_r(0)|}\chi_{B_r(0)}$, where $B_r(0)$ is the ball of radius $r > 0$ centered at $x = 0$ and $|\cdot|$ denotes Lebesgue measure. Then*

$$f * g(x) = \frac{1}{|B_r(0)|}\int_{B_r(0)} f(x-y)\,dy = \frac{1}{|B_r(x)|}\int_{B_r(x)} f(y)\,dy.$$

That is, $f * g(x)$ is the average value of f on the ball $B_r(x)$.

More generally, convolutions have the interpretation of weighted averages, where $f * g$ represents the weighted average of f against g (or vice versa).

The convolution product gives rise to an important class of linear transformations. In particular, for a fixed $f \in L^1$, the mapping

$$g \mapsto C_f(g) := f * g$$

defines a linear transformation from L^p to itself, for any $1 \leq p \leq \infty$. In the language of engineering, we say that convolution with a fixed L^1 function defines a *linear filter.*

An essential property of filters defined via convolution is that they are *shift invariant.*

Definition 1. *Let $f : \mathbb{R}^n \to \mathbb{C}$. For $\tau \in \mathbb{R}^n$, the* shift *of f by τ is the function*

$$f_\tau(x) = f(x-\tau).$$

A filter $\mathcal{A} : L^\infty \to L^\infty$ is shift invariant *if*

$$\mathcal{A}[f_\tau] = [\mathcal{A}f]_\tau \quad \textit{for all} \quad \tau \in \mathbb{R}^n \quad \textit{and} \quad f \in L^\infty. \tag{1}$$

Remark 1. *In mathematical settings one often speaks of 'translations' rather than 'shifts'. We can also describe translation invariance by saying that an operator 'commutes with translations'.*

Proposition 1. *Filters defined by convolution are shift invariant.*

Proof. Let $\psi \in L^1$. Then by a change of variables we have

$$\begin{aligned} C_\psi(f_\tau)(x) &= \int \psi(x-y)f(y-\tau)\,dy \\ &= \int \psi(x-y-\tau)f(y)\,dy = [C_\psi(f)](x-\tau). \end{aligned}$$

□

Remark 2. *The converse is essentially true as well. That is, any shift invariant linear filter can be represented by convolution. However, to make this precise one must introduce the notion of convolution with distributions (also called 'generalized functions'), which we will not pursue here.*

Another essential property of filters defined via convolution is that they are *Fourier multiplier operators*.

Definition 2. *Let $T : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. We say that T is a* Fourier multiplier operator *if there exists a bounded function $m : \mathbb{R}^n \to \mathbb{C}$ such that*

$$\mathcal{F}[T(f)](\xi) = m(\xi)\hat{f}(\xi)$$

(as an equality of L^2 functions). We may write $T = \mathcal{F}^{-1}m\mathcal{F}$, and we call m the symbol *or* multiplier *for T.*

Proposition 2. *Filters defined by convolution are Fourier multiplier operators.*

Proof. Suppose $\psi \in L^1(\mathbb{R}^n)$ and define $C_\psi(g) = \psi * g$ as above. Proposition 2 from the previous section implies

$$\mathcal{F}[C_\psi(g)](\xi) = \mathcal{F}[\psi * g](\xi) = \hat{\psi}(\xi)\hat{g}(\xi) \tag{2}$$

for all $g \in L^1$. This identity shows that $C_\psi : L^2 \to L^2$ is a Fourier multiplier operator with symbol $\hat{\psi}$.

The following may be skipped in lecture: Strictly speaking, we actually need prove (2) for $g \in L^2$. We can achieve this as follows: For $g \in L^2$, we may define $g_k = \chi_{[-k,k]}g$, which satisfy $g_k \in L^1 \cap L^2$ and $g_k \to g$ in L^2-norm. This implies $\psi * g_k \to \psi * g$ in L^2, as

$$\|\psi * g_k - \psi * g\|_{L^2} = \|\psi * [g_k - g]\|_{L^2} \leq \|\psi\|_{L^1}\|g_k - g\|_{L^2}.$$

On the other hand, by Parseval's theorem, we also have $\hat{g}_k \to \hat{g}$ in L^2-norm. Thus, as $\hat{\psi} \in L^\infty$, we have $\hat{\psi}\hat{g}_k \to \hat{\psi}\hat{g}$ in L^2-norm. Therefore the identity (2) also holds for $g \in L^2$, provided it is interpreted as an equality of L^2 functions. □

Remark 3. *The preceding result has real practical importance. It shows that we can compute filters arising from convolution using the Fourier transform and its inverse. Indeed, we have*

$$C_\psi(f) = \mathcal{F}^{-1}(\hat{\psi}\hat{f}).$$

This is important because the Fourier transform has efficient and accurate numerical implementations.

So far, we have primarily considered convolution operators C_ψ with $\psi \in L^1$. If we additionally assume that ψ is smooth, then the convolution $f * \psi$ is smooth as well.

Proposition 3. *Suppose f is a locally integrable function and that ψ is a function with bounded support and k continuous derivatives. Then $f*\psi$ also has k continuous derivatives. In fact, for any multiindex α with $|\alpha| \leq k$, we have*

$$\partial_x^\alpha(f * \psi) = f * (\partial_x^\alpha \psi).$$

Remark 4. *A* multiindex *$\alpha = (\alpha_1, \ldots, \alpha_n)$ is an element of $[\mathbb{N} \cup \{0\}]^n$. We write*

$$\partial_x^\alpha = [\tfrac{\partial}{\partial x_1}]^{\alpha_1} \cdots [\tfrac{\partial}{\partial x_n}]^{\alpha_n}.$$

Remark 5. *The integrals defining the convolutions in Proposition 3 converge due to the fact that f is locally integrable and ψ (along with all of its derivatives) have bounded support.*

Proof. Let us prove something simpler, namely, that if f is locally integrable and $\psi \in C_b$, then $f*\psi$ is continuous. Repeating the argument using difference quotients will then yield the more general result.

Given any $x_1, x_2 \in \mathbb{R}^n$, we may write

$$f * \psi(x_2) - f * \psi(x_1) = \int f(y)[\psi(x_2 - y) - \psi(x_1 - y)]\, dy.$$

Now suppose ψ is supported on $B_R(0)$ and let $\varepsilon > 0$. By uniform continuity of ψ, we may choose $0 < \delta < 1$ so that $|z_2 - z_1| < \delta$ implies $|\psi(z_2) - \psi(z_1)| \leq \frac{\varepsilon}{M}$, where we take

$$M = \int_{B_R(x_1)} |f(y)|\, dy + \int_{B_R(x_2)} |f(y)|\, dy,$$

which is finite by the assumption of local integrability. Then for $|x_2 - x_1| < \delta$, we obtain

$$|f * \psi(x_2) - f * \psi(x_1)| < \varepsilon.$$

This implies continuity. □

Remark 6. *The proof above also shows that if $f \in L^1$ and $\psi \in C_b$, then $f * \psi$ is* uniformly *continuous.*

Remark 7. *We can understand the smoothing effect as follows. As we saw in the previous section, smoothness is connected to decay of the Fourier transform. If we take the Fourier transform of $f * \psi$, we obtain $\hat{f}\hat{\psi}$. As $f \in L^1 \implies \hat{f} \in L^\infty$, we see that $\hat{f}\hat{\psi}$ decays just as rapidly as $\hat{\psi}$. Therefore $f * \psi$ is just as smooth as ψ.*

Combining the previous proposition with the 'approximate identity' arguments mentioned in the last section, we can now describe a general method for the smooth approximation of functions.

Proposition 4 (Approximation to the identity)**.** *Suppose $\varphi \in L^1$ and $\int \varphi = 1$. For each $\varepsilon > 0$, define*

$$\varphi_\varepsilon(x) = \tfrac{1}{\varepsilon}\varphi(\tfrac{x}{\varepsilon}).$$

Then:

(i) *If $f \in L^2$, then $\varphi_\varepsilon * f \to f$ in L^2 as $\varepsilon \to 0$.*
(ii) *If $f \in L^1$, then $\varphi_\varepsilon * f \to f$ in L^1 as $\varepsilon \to 0$.*
(iii) *If f is locally integrable and continuous at x and φ has bounded support, then $\varphi_\varepsilon * f(x) \to f(x)$ as $\varepsilon \to 0$.*

Remark 8.

(a) While we have focused on L^1 and L^2, parts (i)–(ii) actually hold on L^p for any $1 \leq p < \infty$.

(b) If φ is smooth, then by Proposition 3 we see that $\varphi_\varepsilon * f$ give smooth approximations to f in suitable norms.

(c) When proving the Fourier inversion formula, we essentially proved (iii) under stronger assumptions on f (namely, $f \in L^1$ and f uniformly *continuous) but weaker conditions on the kernel (namely, we did not use a kernel with bounded support).*

Proof. These proofs may be skipped in lecture.

As (i) involves the L^2 topology, we can give a simplified proof using Parseval's formula. In particular, it is equivalent to prove that

$$\|[\hat{\varphi}_\varepsilon - 1]\hat{f}\|_{L^2} \to 0 \quad \text{as} \quad \varepsilon \to 0. \tag{3}$$

Now a change of variables shows

$$\widehat{\varphi_\varepsilon}(\xi) = \hat{\varphi}(\varepsilon\xi).$$

Now observe that for every ξ, we have $\hat{\varphi}(\varepsilon\xi) \to \hat{\varphi}(0) = \int \varphi = 1$ as $\varepsilon \to 0$. Thus (3) holds by the dominated convergence theorem.

We turn to (ii). This proof generalizes to treat the L^p case for $1 \leq p < \infty$ by incorporating Minkowski's integral inequality (i.e. $\|F(x,y)\|_{L^pL^1} \leq \|F(x,y)\|_{L^1L^p}$ for $1 \leq p \leq \infty$) and continuity of translations in L^p.

We let $\eta > 0$ and use the continuity of translations in L^1 to find $\delta > 0$ so that

$$\|f(\cdot - y) - f(\cdot)\|_{L^1} < \eta \quad \text{for all} \quad |y| < \delta.$$

We now use $\int \varphi = 1$ to write

$$\left\| \int f(x-y)\varphi_\varepsilon(y)\,dy - f(x) \right\|_{L^1} = \left\| \int [f(x-y) - f(x)]\varphi_\varepsilon(y)\,dy \right\|_{L^1}.$$

We split the dy integral into two regions, namely, $|y| < \delta$ and $|y| \geq \delta$. We first estimate

$$\begin{aligned} &\left\| \int_{|y|\leq\delta} [f(x-y) - f(x)]\varphi_\varepsilon(y)\,dy \right\|_{L^1_x} \\ &\qquad \leq \big\| \varphi_\varepsilon(y) \|f(x-y) - f(x)\|_{L^1_x} \big\|_{L^1_y(|y|\leq\delta)} < \varepsilon \|\varphi\|_{L^1}. \end{aligned}$$

On the other hand, by the triangle inequality,

$$\left\| \int_{|y|>\delta} [f(x-y) - f(x)]\varphi_\varepsilon(y)\,dy \right\|_{L^1_x} \leq 2\|f\|_{L^1}\|\varphi_\varepsilon\|_{L^1(|y|>\delta)}.$$

Now observe that by a change of variables and the dominated convergence theorem,

$$\|\varphi_\varepsilon\|_{L^1(|y|>\delta)} = \|\varphi\|_{L^1(|y|>\delta\eta^{-1})} \to 0 \quad \text{as} \quad \eta \to 0.$$

Combining the estimates for the two regions, we complete the proof.

Finally, consider (iii). We let $\eta > 0$ and choose $\delta > 0$ so that

$$|f(x) - f(x-y)| < \eta \quad \text{for all} \quad |y| < \delta.$$

As ϕ has bounded support, we have that $\phi_\varepsilon(y) = \frac{1}{\varepsilon}\phi(\frac{y}{\varepsilon})$ is supported in $\{|y| \leq \delta\}$ for all ε sufficiently small. Thus

$$\begin{aligned} |\phi_\varepsilon * f(x) - f(x)| &= \left| \int_{|y|\leq\delta} \phi_\varepsilon(y)[f(x-y) - f(x)]\,dy \right| \\ &\leq \int_{|y|\leq\delta} |f(x-y) - f(y)|\,|\phi_\varepsilon(y)|\,dy \leq \eta\|\phi\|_{L^1}. \end{aligned}$$

This completes the proof. □

As an application, let us remove the assumption of uniform continuity from the Fourier inversion formula.

Corollary 1. *The following Fourier inversion formula holds provided $f \in L^1$ and $\hat{f} \in L^1$:*

$$f(x) = \tfrac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} \hat{f}(\xi)\, d\xi \quad \textit{for almost every} \quad x \in \mathbb{R}. \tag{4}$$

Proof. This proof may be skipped in lecture.

Let $f \in L^1$ with $\hat{f} \in L^1$ and take a family of approximate identities $\varphi_\varepsilon(x) = \varepsilon^{-1}\varphi(\frac{x}{\varepsilon})$, where φ is smooth and has bounded support. Then for every $\varepsilon > 0$, we have $\varphi_\varepsilon * f \in L^1$ and (by Remark 6) is uniformly continuous. Moreover, its Fourier transform is given by $\hat{\varphi}(\varepsilon\xi)\hat{f}(\xi) \in L^1$. Therefore the Fourier inversion formula proved in the previous section implies

$$\varphi_\varepsilon * f(x) = \tfrac{1}{2\pi} \int \hat{\varphi}(\varepsilon\xi)\hat{f}(\xi) e^{ix\xi}\, d\xi \tag{5}$$

for all $x \in \mathbb{R}$ and $\varepsilon > 0$. By the dominated convergence theorem (and the fact that $\hat{\varphi}(0) = \int \varphi = 1$), the right-hand side of (5) converges to

$$\tfrac{1}{2\pi} \int \hat{f}(\xi) e^{ix\xi}\, d\xi$$

pointwise for $x \in \mathbb{R}$. On the other hand, the left-hand side of (5) converges to f in L^1-norm as $\varepsilon \to 0$, and hence (by uniqueness of limits) we derive (4). □

In each of the items in Proposition 4, we see that the functions φ_ε converge to some kind of identity element for the convolution product. However, the limiting object is *not* a function. In fact, the convolution product on L^1 does not have an identity element.

Proposition 5 (No identity element)**.** *There is no function $g \in L^1$ such that $f * g = f$ for all $f \in L^1$.*

Proof. For simplicity, let us work in the one-dimensional case.

Suppose such $g \in L^1$ exists and let $f(x) = e^{-x^2/2}$. Then we have $\hat{f}(\xi) = \sqrt{2\pi} e^{-\xi^2/2}$, and so by the convolution identity $\mathcal{F}[f * g] = \hat{f}\hat{g}$ we derive

$$e^{-\xi^2/2}\hat{g}(\xi) = e^{-\xi^2/2} \quad \text{for all} \quad \xi \in \mathbb{R}.$$

This implies $\hat{g}(\xi) = 1$ for all $\xi \in \mathbb{R}$, which contradicts the Riemann–Lebesgue lemma. Indeed, if $g \in L^1$ then we must have $g \to 0$ as $|\xi| \to \infty$. □

Although the convolution product has no identity element in L^1, there is in fact a perfectly natural notion of an identity element for the convolution product, namely, the *Dirac delta distribution*. This object arises frequently in mathematics, physics, and engineering. It is not a function, but rather a *linear functional* or *distribution*. It may also be called the *δ-function* (a misnomer) or a *unit impulse*.

Definition 3 (Dirac delta distribution)**.** *The Dirac delta distribution is the functional $\delta : C(\mathbb{R}^n) \to \mathbb{C}$ defined by*

$$\delta(f) = f(0) \quad \textit{for any} \quad f \in C(\mathbb{R}^n).$$

As mentioned above, δ is not itself a function, but rather a 'distribution' or 'generalized function'.

The term 'generalized function' is reasonable, in the sense that functions themselves may be identified as distributions. For example, any function $g \in L^1$ can be identified with a functional $T_g : L^\infty(\mathbb{R}^n) \to \mathbb{C}$ by setting

$$T_g(f) = \int f(y)g(y)\,dy \quad \text{for any} \quad f \in L^\infty(\mathbb{R}^n). \tag{6}$$

This defines a continuous linear transformation with norm equal to $\|g\|_{L^1}$. Moreover, the mapping $g \mapsto T_g$ is injective.

If we (formally) suppose that δ arises as in (6), so that

$$\text{“}f(0) = \int \delta(y)f(y)\,dy\text{”},$$

then we obtain

$$\text{“}f * \delta(x) = \int \delta(y)f(x-y)\,dy = f(x)\text{”},$$

demonstrating that δ is indeed the identity element for the convolution product. In fact, by extending the definition of convolution to accommodate distributions, one can make this rigorous. Furthermore, one can extend the Fourier transform to the setting of distributions, and in this setting one obtains $\mathcal{F}\delta \equiv 1$, which is precisely what we expect in light of the proof of Proposition 5.

The results of Proposition 4 may be summarized succinctly by saying that for $\varphi \in L^1$ with $\int \varphi = 1$, the functions $\varphi_\varepsilon(x) = \frac{1}{\varepsilon}\varphi(\frac{x}{\varepsilon})$ converge to δ 'in the sense of distributions'.

In many applications one needs to design suitable approximations to δ. For example, if we have some 'noisy' data, then (as we have seen above) convolution with an approximate identity provides a way to smooth out the data without losing too much information. In practice, one can build approximate identities either by using functions like φ_ε above or by approximating $\mathcal{F}\delta$ (which is just the constant function 1) in Fourier space.

Example 2. *Suppose we want to construct a function φ such that $\hat{\varphi}$ is approximately one in a given interval $[-B, B]$. If we choose*

$$\hat{\varphi}(\xi) = \chi_{[-B,B]}(\xi),$$

then we obtain

$$\varphi(x) = \psi_B(\xi) := \tfrac{1}{\pi} B\,\mathrm{sinc}(Bx)$$

(called a sinc pulse *in this context). This function is not a good choice for an approximate identity because it is not absolutely integrable. On the other hand, if we approximate $\hat{\delta}$ by*

$$\hat{\varphi}(\xi) = \tfrac{1}{2B}\chi_{[-B,B]} * \chi_{[-B,B]}(\xi),$$

then we obtain the sinc^2 pulse

$$\varphi(x) = \tfrac{1}{2B}\psi_B^2(x),$$

which is positive, absolutely integrable, and more sharply peaked at $x = 0$. Neither sinc *nor* sinc^2 *have bounded support. They both have oscillatory tails as $|x| \to \infty$, known in engineering literature as* side lobes.

We will close this section by discussing briefly the notion of *resolution* in this setting. We have previously seen that convolution has a smoothing effect on functions, which can be helpful in suppressing 'noise'. On the other hand, convolution makes

the function 'blurrier'. To make this more precise, we begin with the following result.

Proposition 6. *The support of $f * g$ (i.e. the closure of the set where $f * g$ is non-zero) is contained in the algebraic sum of* $\operatorname{supp} f$ *and* $\operatorname{supp} g$*, i.e. the set*

$$\operatorname{supp} f + \operatorname{supp} g = \{x + y : x \in \operatorname{supp} f \quad \textit{and} \quad y \in \operatorname{supp} g\}.$$

Proof. If x is not in $\operatorname{supp} f + \operatorname{supp} g$, then for every y we have $f(y)g(x-y) = 0$ and hence $f * g(x) = 0$. In fact, we can apply this argument on some ball containing x that is disjoint from $\operatorname{supp} f + \operatorname{supp} g$. □

Using this support property, we see that if ψ is a nonnegative approximate identity supported in the ball of radius ε, we have that the support of $\psi * f$ is contained in the set of points within a distance ε of the support of f. The value $f * \psi(x)$ is a weighted average of the values of f over $[x - \varepsilon, x + \varepsilon]$.

Now consider what happens if we take the j-fold convolution of ψ with itself, denoted $\psi *_j \psi$. This will still give an approximate identity, as its total integral is one:

$$\int \psi *_j \psi(x)\, dx = \mathcal{F}[\psi *_j \psi](0) = [\hat{\psi}(0)]^j = \left[\int \psi(x)\, dx\right]^j = 1.$$

By the above result, $\psi *_j \psi$ will be supported in $[-j\varepsilon, j\varepsilon]$, and so $f * [\psi *_j \psi](x)$ will be a weighted average over $[x - j\varepsilon, x + j\varepsilon]$. On the other hand, the Fourier transform of $\psi *_j \psi$ decays j times faster than $\hat{\psi}$. Thus $f * [\psi *_j \psi]$ will be a *blurrier*, *smoother* version of f than $f * \psi$.

We can balance the two effects (smoothing/blurring) by considering the scaled j-fold convolution $j[\psi *_j \psi](jx)$. This function is again supported in $[-\varepsilon, \varepsilon]$, so that it blurs f no more than ψ itself. On the other hand, it is far more smoothing than ψ, as its Fourier transform is given by $[\hat{\psi}(\frac{\xi}{j})]^j$, which decays j times faster than $\hat{\psi}$.

More on the Radon transform

With the theory of the Fourier transform and convolutions in place, we return to our study of the Radon transform. Our first main goal will be an inversion formula.

We begin by recalling the definition of the Radon transform: for a function $f \in C_b(\mathbb{R}^2)$, the Radon transform is the function

$$\mathcal{R}f : \mathbb{R} \times \mathbb{S}^1 \to \mathbb{R}$$

given by

$$\mathcal{R}f(t,\omega) = \int_{-\infty}^{\infty} f(s\hat{\omega} + t\omega)\, ds,$$

where $\omega = (\omega_1, \omega_2)^T$ and $\hat{\omega} = (-\omega_2, \omega_1)^T$. In particular, $\mathcal{R}f(t,\omega)$ is the integral of f over the oriented line $\ell_{t,\omega}$, where ω is the vector perpendicular to this line and $|t|$ is the distance from the line to the origin. Continuity and bounded support are not necessary conditions for the Radon transform to be well-defined. What is really needed is

$$\int_{-\infty}^{\infty} |f(t\omega + s\hat{\omega})|\, ds < \infty \quad \text{for all} \quad (t,\omega) \in \mathbb{R} \times \mathbb{S}^1.$$

The set of functions satisfying this property is called the *natural domain* of the Radon transform. This includes, for example, any piecewise continuous function decaying like $|x|^{-1-\varepsilon}$ for some $\varepsilon > 0$ as $|x| \to \infty$.

When integrating with respect to ω, it is useful to use the parametrization

$$\omega(\theta) = (\cos\theta, \sin(\theta))^T, \quad \theta \in [0, 2\pi).$$

Then integration over $\mathbb{R} \times \mathbb{S}^1$ is given by

$$\int_0^{2\pi} \int_{-\infty}^{\infty} h(t, \omega(\theta))\, dt\, d\theta, \quad \text{denoted by} \quad \int_0^{2\pi} \int_{-\infty}^{\infty} h(t,\omega)\, dt\, d\omega,$$

and the $L^2(\mathbb{R} \times \mathbb{S}^1)$-norm is defined via

$$\|h\|^2_{L^2(\mathbb{R}\times\mathbb{S}^1)} = \int_0^{2\pi} \int_{-\infty}^{\infty} |h(t,\omega)|^2\, dt\, d\omega.$$

We say $h : \mathbb{R} \times \mathbb{S}^1 \to \mathbb{R}$ is continuous if $h(t,\theta) := h(t, \omega(\theta))$ is 2π-periodic in θ and continuous as a function on $\mathbb{R} \times [0, 2\pi]$. Differentiability is defined similarly.

Our first main result connects the Radon transform and the Fourier transform, which will allow us to leverage the Fourier inversion formula into a Radon inversion formula.

Theorem 1 (Central slice theorem)**.** *Let $f \in L^1(\mathbb{R}^2)$ be in the natural domain of $\mathcal{R}$. For any $r \in \mathbb{R}$ and $\omega \in \mathbb{S}^1$,*

$$\int_{\mathbb{R}} \mathcal{R}f(t,\omega) e^{-itr}\, dt = \hat{f}(r\omega),$$

where $\hat{f}$ is the (two-dimensional) Fourier transform of f.

Remark 1. *The expression on the left gives the one-dimensional Fourier transform of $t \mapsto \mathcal{R}f(t,\omega)$ (for fixed $\omega \in \mathbb{S}^1$).*

Proof. We first use the definition of $\mathcal{R}$ to write

$$\int_{\mathbb{R}} \mathcal{R}f(t,\omega)e^{-itr}\,dt = \iint_{\mathbb{R}^2} f(t\omega + s\hat{\omega})e^{-itr}\,ds\,dt.$$

Now let us change variables and write $x = x(t,s) = t\omega + s\hat{\omega}$. The Jacobian is

$$\det\begin{bmatrix} \omega_1 & -\omega_2 \\ \omega_2 & \omega_1 \end{bmatrix} = |\omega|^2 = 1.$$

Moreover, $x \cdot \omega = t$, and hence the integral becomes

$$\int_{\mathbb{R}^2} f(x)e^{-ix\cdot(r\omega)}\,dx = \hat{f}(r\omega).$$

This completes the proof. □

In light of the central slice theorem, we introduce the following notation:

Definition 1. *For a function $h = h(t,\omega)$ on $\mathbb{R}\times\mathbb{S}^1$, we define*

$$\tilde{h}(r,\omega) := \int_{-\infty}^{\infty} h(t,\omega)e^{-irt}\,dt.$$

Then the central slice theorem is the identity

$$\hat{f}(r\omega) = \widetilde{\mathcal{R}f}(r,\omega).$$

Before moving on to the Radon inversion formula, let us first point out that in contrast to the Fourier transform, the Radon transform does *not* extend to a continuous map from $L^2(\mathbb{R}^2) \to L^2(\mathbb{R}\times\mathbb{S}^1)$, nor can $\mathcal{R}^{-1}$ be a continuous map from $L^2(\mathbb{R}\times\mathbb{S}^1)$ to $L^2(\mathbb{R}^2)$. This will be a consequence of the following Parseval formula for the Radon transform.

Theorem 2 (Parseval formula for the Radon transform)**.** *Suppose $f \in L^2$ and f is in the natural domain of $\mathcal{R}$. Then*

$$\int_{\mathbb{R}^2} |f(x)|^2\,dx = \tfrac{1}{4\pi^2}\int_0^{\pi}\int_{-\infty}^{\infty} |\widetilde{\mathcal{R}f}(r,\omega)|^2|r|\,dr\,d\omega.$$

Proof. We assume additionally that $f \in L^1$. This may be removed with an approximate identity argument.

We first recall that since $\ell_{t,\omega} = \ell_{-t,-\omega}$ are the same line, the Radon transform satisfies $\mathcal{R}f(t,\omega) = \mathcal{R}f(-t,-\omega)$. This implies

$$\widetilde{\mathcal{R}f}(-r,-\omega) = \int_{-\infty}^{\infty} \mathcal{R}f(t,-\omega)e^{irt}\,dt = \int_{-\infty}^{\infty} \mathcal{R}f(-t,\omega)e^{irt}\,dt = \widetilde{\mathcal{R}f}(r,\omega),$$

where we have changed variables $t \mapsto -t$ in the final step.

Changing to polar coordinates and applying the central slice theorem, we can therefore obtain

$$\begin{aligned}\int_{\mathbb{R}^2} |f(x)|^2\,dx &= \tfrac{1}{4\pi^2}\int_0^{2\pi}\int_0^{\infty} |\hat{f}(r\omega)|^2 r\,dr\,d\omega \\ &= \tfrac{1}{4\pi^2}\int_0^{\pi}\int_{-\infty}^{\infty} |\widetilde{\mathcal{R}f}(r,\omega)|^2|r|\,dr\,d\omega,\end{aligned}$$

as desired. □

Using this Parseval formula, we may now preclude the possibility that $\mathcal{R}$ or $\mathcal{R}^{-1}$ be bounded between the appropriate L^2 spaces.

Corollary 1. *There are no $m > 0$, $M > 0$ such that*

$$\|\mathcal{R}f\|_{L^2(\mathbb{R}\times\mathbb{S}^1)} \leq M\|f\|_{L^2(\mathbb{R}^2)} \quad \textit{and} \quad \|\mathcal{R}f\|_{L^2(\mathbb{R}\times\mathbb{S}^1)} \geq m\|f\|_{L^2(\mathbb{R}^2)} \tag{1}$$

for all $f \in L^2(\mathbb{R}^2)$.

Proof. This proof may be skipped in lecture. By the $1d$ Parseval formula for the Fourier transform, we have

$$\|\mathcal{R}f\|^2_{L^2(\mathbb{R}\times\mathbb{S}^1)} = \tfrac{1}{2\pi}\int_0^{2\pi}\int_{-\infty}^{\infty} |\widetilde{\mathcal{R}f}(r,\omega)|^2\, dr\, d\omega.$$

Thus, if the first bound in (1) holds for some $M > 0$, we have by the Parseval theorem for the Radon transform that

$$\tfrac{1}{2\pi}\int_0^{2\pi}\int_{-\infty}^{\infty} |\widetilde{\mathcal{R}f}(r,\omega)|^2\, dr\, d\omega \leq \tfrac{M}{4\pi^2}\int_0^{\pi}\int_{-\infty}^{\infty} |\widetilde{\mathcal{R}f}(r,\omega)|^2 |r|\, dr\, d\omega \tag{2}$$

for all $f \in L^2(\mathbb{R}^2)$. Now let us consider the family of functions $f_a(t) = a^{1/4}e^{-at^2}$. Then we have $\tilde{f}_a(r) = \sqrt{\pi}a^{-1/4}e^{-r^2/4a}$. Thus, after a change of variables $y = r/\sqrt{a}$, we find that for the left-hand side of (2) we obtain

$$\text{LHS}(2) = \pi\int_{-\infty}^{\infty} e^{-y^4/16}\, dy \quad \text{for all} \quad a > 0.$$

On the other hand, by the same change of variables,

$$\text{RHS}(2) = \tfrac{M}{4}a^{\frac{1}{2}}\int_{-\infty}^{\infty} e^{-y^4/16}|y|\, dy \to 0 \quad \text{as} \quad a \to 0.$$

Thus an inequality of the form (2) (with M fixed) that holds for all $f \in L^2$ is impossible. A similar argument shows that the second inequality in (1) is impossible, as well. □

Remark 2. *The quantity appearing in the Parseval formula, namely, $|r|^{\frac{1}{2}}\widetilde{\mathcal{R}f}(r,\omega)$, is the Fourier transform of the half-derivative operator in the affine parameter. Thus the Parseval formula implies that the Radon transform of an L^2-function must have some smoothness in the affine parameter.*

Using the central slice theorem and Fourier inversion, we can now give an inversion formula for the Radon transform.

Theorem 3 (Radon inversion formula)**.** *Let $f \in L^1$ be in the natural domain of $\mathcal{R}$, and suppose $\hat{f} \in L^1$. Then*

$$f(x) = \tfrac{1}{4\pi^2}\int_0^{\pi}\int_{-\infty}^{\infty} e^{irx\cdot\omega}\widetilde{\mathcal{R}f}(r,\omega)|r|\, dr\, d\omega.$$

Proof. We apply the Fourier inversion formula and change to polar coordinates to write

$$f(x) = \tfrac{1}{4\pi^2}\int_{\mathbb{R}^2} e^{ix\xi}\hat{f}(\xi)\, d\xi = \tfrac{1}{4\pi^2}\int_0^{2\pi}\int_0^{\infty} e^{irx\cdot\omega}\hat{f}(r\omega) r\, dr\, d\omega.$$

We now apply the central slice theorem. This yields

$$f(x) = \tfrac{1}{4\pi^2}\int_0^{2\pi}\int_0^{\infty} e^{ix\cdot(r\omega)}\widetilde{\mathcal{R}f}(r,\omega) r\, dr\, d\omega. \tag{3}$$

Recalling that $\widetilde{\mathcal{R}f}(-r,-\omega) = \widetilde{\mathcal{R}f}(r,\omega)$, we may rewrite (3) to obtain the stated inversion formula. □

Remark 3. *In imaging applications, it is typical to work with piecewise continuous functions f of bounded support. This does not guarantee that $\hat{f} \in L^1$, and so the Fourier inversion formula needs to be interpreted appropriately. In particular, using the L^2 theory, we may write*

$$f(x) = \lim_{R\to\infty} \tfrac{1}{4\pi^2} \int_0^\pi \int_{-R}^R e^{ix\cdot(r\omega)} \widetilde{\mathcal{R}f}(r,\omega)|r|\, dr\, d\omega$$

as an L^2 limit. Typically, if f is continuous at x, then the integral on the right-hand side will exist as an improper Riemann integral and converge to $f(x)$.

Remark 4. *We now have completed a very idealized mathematical model for X-ray CT imaging:*

- *The attenuation coefficient f of a two-dimensional slice of material determines the attenuation of the intensity I of X-rays of a given energy traveling along a line $\ell_{t,\omega}$ via Beer's law: $\frac{d}{ds} I = -fI$.*
- *By comparing the intensity of the incident beam to the emitted beam, we measure the Radon transform of f:*

$$\mathcal{R}f(t,\omega) = -\log\left[\tfrac{I_o(t,\omega)}{I_i(t,\omega)}\right].$$

- *Using the Radon inversion formula, we reconstruct f from the measurements of $\mathcal{R}f$.*

This is clearly an overly simplified model. One obvious issue is the fact that we will have only finitely many samples of $\mathcal{R}f$ to work with. Nonetheless, this basic model will already be good starting point for developing some practical reconstruction algorithms.

The Radon inversion formula can be viewed as two steps. In particular, one first performs the radial integral, and then takes the angular integral.

- The radial integral has the form of a *filter* applied to the Radon transform. Indeed, we may write this integral as

$$|\nabla|\mathcal{R}f(t,\omega) = \tfrac{1}{2\pi} \int_{-\infty}^{\infty} \widetilde{\mathcal{R}f}(r,\omega) e^{irt} |r|\, dr, \quad t = \langle x, \omega\rangle.$$

  This has the form of a Fourier multiplier operator $\mathcal{F}^{-1} m \mathcal{F}$ applied to $\mathcal{R}f$, where $\mathcal{F}$ is the Fourier transform in the affine parameter and the symbol m is given by $m(r) = |r|$.
- The angular integral is essentially just the back-projection formula, but applied to the *filtered* Radon transform rather than directly to the Radon transform:

$$f(x) = \tfrac{1}{2\pi} \int_0^\pi |\nabla|\mathcal{R}f(\langle x,\omega\rangle, \omega)\, d\omega.$$

For this reason, we call the Radon inversion formula the *filtered back-projection formula*. To better understand the filtered back-projection formula (and how we might build approximate versions of it), we will need to take a closer look at the filtering step, which is quite a bit more subtle than the back-projection step.

The filter $|\nabla|$ appearing above is the composition of two filters. This is most readily seen on the Fourier side, where the symbol $|r|$ is given by the product of r

and $\operatorname{sign}(r)$, where

$$\operatorname{sign}(r) = \begin{cases} 1 & r > 0 \\ 0 & r = 0 \\ -1 & r < 0. \end{cases}$$

We first consider both of these operators separately. To remain consistent with the above notation, we write functions as $f = f(t)$ and write the Fourier transform as $\hat{f} = \hat{f}(r)$.

Example 1. *The Fourier multiplier operator with symbol $m(r) = r$ is $\frac{1}{i}\partial_t$. Indeed, integrating by parts we have*

$$\mathcal{F}[\tfrac{1}{i}\partial_t f](r) = \tfrac{1}{i}\int_{\mathbb{R}} e^{-irt}\partial_t f(t)\,dt = r\int_{\mathbb{R}} e^{-irt} f(t)\,dt = r\hat{f}(r).$$

Example 2. *The Fourier multiplier operator with symbol $m(r) = \operatorname{sign}(r)$ is called the* Hilbert transform*, denoted by $\mathcal{H}$. Thus for $f \in L^2$ such that $\hat{f} \in L^1$, we have*

$$\mathcal{H}f(t) = \tfrac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(r)\operatorname{sign}(r) e^{itr}\,dr.$$

Differentiation is simple in the sense that it is a local operator. To evaluate the derivative of f at a point only requires the values of f nearby that point. On the other hand, real data will always contain some noise, which means that the data will need to be smoothed out before it can be differentiated. On the other hand, the Hilbert transform is a *nonlocal* operator. That is, to compute $\mathcal{H}f$ at a single point requires knowledge of f at all points. This poses a different challenge, as in practice we can only ever compute integrals over finite intervals.

With the above definitions in place, the filtered back-projection formula can be rewritten

$$f(x) = \tfrac{1}{2\pi i}\int_0^{\pi} \mathcal{H}[\partial_t \mathcal{R}f](\langle x, \omega\rangle, \omega)\,d\omega.$$

In particular, we reconstruct f by back-projecting the Hilbert transform of $\frac{1}{i}\partial_t \mathcal{R}f$.

Remark 5. *If f has bounded support, then so does $\partial_t \mathcal{R}f$. Using the theory of analytic functions, one can then show that $\mathcal{H}\partial_t \mathcal{R}f$ does* not *have bounded support. Thus the integrand in the filtered back-projection formula is typically non-zero, and outside of the support of f the integral only vanishes due to cancellation between positive and negative parts.*

Example 3 (Example of Radon inversion)**.** *Consider the simple example of $f = \chi_{B_1}$, where B_1 is the closed unit ball. Then, as we have previously computed,*

$$\mathcal{R}f(t,\omega) = \begin{cases} 2\sqrt{1-t^2} & |t| \le 1, \\ 0 & |t| > 1. \end{cases}$$

Let us apply the filtered back-projection formula. This requires that we compute $\mathcal{H}\partial_t \mathcal{R}f$, which we observe equals $\partial_t \mathcal{H}\mathcal{R}f$. This relies on the fact that all Fourier multiplier operators commute.

A computation using methods from complex analysis shows that

$$\tfrac{1}{i}\mathcal{H}\mathcal{R}f(t,\omega) = \begin{cases} 2t & |t| < 1 \\ 2(t + \sqrt{t^2-1}) & t < -1 \\ 2(t - \sqrt{t^2-1}) & t > 1. \end{cases}$$

Now, this function fails to be differentiable at $t = \pm 1$. Nonetheless, we can interpret $\frac{1}{i}\partial_t \mathcal{H}\mathcal{R}f$ as a 'weak derivative' in L^1, given by

$$\tfrac{1}{i}\partial_t \mathcal{H}\mathcal{R}f(t,\omega) = \begin{cases} 2 - \frac{2|t|}{\sqrt{t^2-1}} & |t| > 1, \\ 2 & |t| < 1. \end{cases}$$

We now need to do the back-projection step. For $x \in B_1$, we have $|\langle x, \omega\rangle| \le 1$, and so

$$\tfrac{1}{2\pi i}\int_0^\pi \mathcal{H}\partial_t \mathcal{R}f(\langle x,\omega\rangle,\omega)\,d\omega = \tfrac{1}{2\pi}\int_0^\pi 2\,d\omega = 1 = \chi_{B_1}(x),$$

as desired. Next consider $x \in \mathbb{R}^2$ with $|x| > 1$. Using the rotation symmetry, it suffices to evaluate $f(x,0)$ for some $x \in \mathbb{R}$ with $x > 1$. In this case $\langle x, \omega\rangle = x\cos\theta$. We now choose the angle $\theta_x \in (0, \frac{\pi}{2})$ so that $x\cos\theta_x = 1$. Then the inversion formula becomes

$$f(x,0) = \tfrac{1}{2\pi}\left[4\int_0^{\theta_x}\left(1 - \tfrac{|x\cos\theta|}{\sqrt{x^2\cos^2\theta - 1}}\right)d\theta - 2\int_{\theta_x}^{\pi-\theta_x} d\theta\right] = 0.$$

This is a much more complicated integral, involving an unbounded integrand. In particular, this would be a more challenging integral to compute numerically.

We turn to some practical issues related to the computation of the filtered back-projection formula. We begin with a further discussion of the Hilbert transform.

One effective approach to approximating the Hilbert transform is to use the Fourier representation.

Proposition 1. *Suppose $\{\phi_\varepsilon\}$ is a uniformly bounded family of locally integrable functions that satisfies $\phi_\varepsilon(\xi) \to \operatorname{sign}(\xi)$ as $\varepsilon \to 0$ for all $\xi \in \mathbb{R}$. If $f \in L^2$, then*

$$\mathcal{H}f = \lim_{\varepsilon\to 0} \mathcal{F}^{-1}[\phi_\varepsilon \hat{f}],$$

where the limit is taken in L^2-norm.

Proof. By Parseval's formula, we have

$$\|\mathcal{H}f - \mathcal{F}^{-1}[\phi_\varepsilon\hat{f}]\|_{L^2}^2 = \tfrac{1}{2\pi}\int_{\mathbb{R}} |\operatorname{sign}(\xi) - \phi_\varepsilon(\xi)|^2 |\hat{f}(\xi)|^2\,d\xi \to 0 \quad \text{as} \quad \varepsilon \to 0$$

by the dominated convergence theorem. □

Example 4. *If f is smooth enough that $\hat{f}$ decays as $|\xi| \to \infty$, then we can write*

$$\mathcal{H}f(t) = \lim_{\varepsilon\to 0} \tfrac{1}{2\pi}\int_{\mathbb{R}} \phi_\varepsilon(\xi)\hat{f}(\xi)e^{it\xi}\,d\xi$$

as a pointwise limit. If we take

$$\phi_\varepsilon(\xi) = \hat{h}_\varepsilon(\xi) := \operatorname{sign}(\xi)e^{-\varepsilon|\xi|},$$

then one can compute the inverse Fourier transform of $\hat{h}_\varepsilon$ to be

$$h_\varepsilon(t) = \tfrac{i}{\pi}\tfrac{t}{t^2+\varepsilon^2}.$$

This is a smooth function that decays at infinity (although not fast enough to belong to L^1). Using these functions, we arrive at approximations to the Hilbert transform expressed as convolutions:

$$\mathcal{H}_\varepsilon f = \mathcal{F}^{-1}[\hat{f}\,\hat{h}_\varepsilon] = f * h_\varepsilon.$$

Note that as $h_\varepsilon \in L^2$, these convolutions are well-defined pointwise for $f \in L^2$

If we send $\varepsilon \to 0$ in the previous example, we arrive at a guess for an expression of the Hilbert transform as a convolution operator. This is expected, as any Fourier multiplier operator should be given by convolution with the inverse Fourier transform of the symbol.

As $h_\varepsilon(t) \to \frac{i}{t\pi}$ pointwise, we arrive formally at the identity

$$\text{“} \quad \mathcal{H}f(t) = \frac{i}{\pi}\int_{-\infty}^{\infty} \tfrac{f(s)}{t-s}\, ds \quad \text{”}.$$

As $\frac{1}{t}$ is not integrable near $t = 0$, this expression cannot be expected to converge absolutely and hence requires a suitable interpretation. For this, we use the *Cauchy principal value* interpretation. It requires some degree of smoothness for f.

Theorem 4 (Hilbert transform as convolution)**.** *Suppose f has bounded support and is Hölder continuous of some order $\alpha > 0$, i.e.*

$$\text{there exists} \quad M > 0 \quad \text{such that} \quad |f(t) - f(s)| \leq M|t-s|^\alpha \tag{4}$$

for all s, t. Then

$$\mathcal{H}f = \tfrac{i}{\pi}\,\mathrm{PV}[f * \tfrac{1}{s}],$$

where PV *denotes the Cauchy principal value, i.e.*

$$\mathrm{PV}[f * \tfrac{1}{s}](t) = \lim_{\varepsilon \to 0} \int_{|s|>\varepsilon} \tfrac{f(t-s)}{s}\, ds. \tag{5}$$

Proof. This proof may be skipped in lecture.

Let us first show that the limit defining the principal value exists for all $t \in \mathbb{R}$. For $t \notin [-\frac{R}{2}, \frac{R}{2}]$, we have

$$|s| \geq |t| - |t-s| > \tfrac{R}{2} - \tfrac{R}{2} \implies |s| > 0,$$

while

$$|s| \leq |t-s| + |t| \leq \tfrac{R}{2} + |t|.$$

Thus the limit exists and equals

$$\int_{\mathbb{R}} \tfrac{f(t-s)}{s}\, ds$$

in this regime.

For $t \in [-\frac{R}{2}, \frac{R}{2}]$, we first observe that since

$$\int_{\varepsilon<|s|<R} \tfrac{ds}{s} = 0,$$

we have by the triangle inequality and (4)

$$\begin{aligned}\left|\int_{|s|>\varepsilon} \tfrac{f(t-s)}{s}\, ds\right| &= \left|\int_{\varepsilon<|s|<R} \tfrac{f(t-s)-f(t)}{s}\, ds\right| \\ &\leq M\int_{\varepsilon<|s|<R} s^{-1+\alpha}\, ds.\end{aligned}$$

This implies that we the limit as $\varepsilon \to 0$ exists.

To conclude the proof, we recall the functions h_ε from the example above and endeavor to prove

$$\lim_{\varepsilon\to 0} h_\varepsilon * f(t) = \tfrac{i}{\pi}\,\mathrm{PV}(f * \tfrac{1}{s})(t)$$

pointwise for $t \in \mathbb{R}$. For $t \notin [-\frac{R}{2}, \frac{R}{2}]$, we use that $h_\varepsilon(s) \to \frac{i}{\pi s}$ pointwise and argue as above to deduce

$$f * h_\varepsilon(t) \to \tfrac{i}{\pi} \int_{\mathbb{R}} \tfrac{f(t-s)}{s}\, ds,$$

yielding the correct value in this regime. For $t \in [-\frac{R}{2}, \frac{R}{2}]$, we use that h_ε is odd to write

$$\begin{aligned} &\int_{\varepsilon<|s|<R} f(t-s)\big[\tfrac{1}{s} - \tfrac{s}{s^2+\varepsilon^2}\big]\, ds + \int_{|s|<\varepsilon} \tfrac{sf(t-s)}{s^2+\varepsilon^2}\, ds \\ &\qquad = \int_{\varepsilon<|s|<R} \tfrac{\varepsilon^2[f(t-s)-f(t)]}{s(s^2+\varepsilon^2)}\, ds + \int_{|s|<\varepsilon} \tfrac{s[f(t-s)-f(t)]}{s^2+\varepsilon^2}\, ds \end{aligned}$$

We rewrite the first integral with a change of variables to obtain

$$\int_{1<|s|<\varepsilon^{-1}R} \tfrac{f(t-s\varepsilon)-f(t)}{\sigma(\sigma^2+1)}\, ds,$$

which tends to zero as $\varepsilon \to 0$ by dominated convergence. For the second integral, we can utilize the Hölder bound and again use dominated convergence, bounding the integrand by $M|s|^{-1+\alpha}$ (which is integrable on $(-1,1)$, say). This completes the proof. □

Remark 6. *The preceding proof shows that for a Hölder continuous function supported in* $[-\frac{R}{2}, \frac{R}{2}]$*, we have*

$$\mathcal{H}f(t) = \tfrac{i}{\pi} \int_{-R}^{R} \frac{f(t-s)-f(t)}{s}\, ds \quad \textit{for} \quad t \in [-\tfrac{R}{2}, \tfrac{R}{2}].$$

♣ ♣ ♣

In the rest of this section, we will briefly discuss issues related to approximating the Radon transform and its inverse.

Recall that the Radon inversion formula requires that we take the Hilbert transform of the derivative of some measured data (the Radon transform of the function we are trying to recover). We also saw (from the Parseval formula) that unless this measured data has some regularity in the affine parameter, it is not the Radon transform of an L^2 function. (In fact, the more general problem of characterizing the range of the Radon transform is a relevant and important problem, although we do not pursue these questions here.) In applications, our measured data will not be so regular due to measurement noise, and more generally we should not expect that the data will belong to the range of the Radon transform.

One approach is to work with regularized inverses as follows. Recall that we previously approximated the Hilbert transform $\mathcal{H}$ by the family of convolution operators

$$\mathcal{H}_\varepsilon f = f * h_\varepsilon, \quad h_\varepsilon(t) = \tfrac{i}{\pi} \tfrac{t}{t^2+\varepsilon^2}.$$

Using this and recalling $f * \partial_t g = \partial_t f * g$, we may derive the following approximate Radon inversion formula:

$$f(x) \approx \tfrac{1}{2\pi i} \int_0^\pi [\partial_t h_\varepsilon * \mathcal{R}f](\langle x, \omega\rangle, \omega)\, d\omega.$$

A key advantage of this approximation is that we no longer need to approximate derivatives of $\mathcal{R}f$.

Another approach to approximation is to use the Fourier representation. For this, we first fix $N > 0$. Suppose $\psi_N \in L^1$ satisfies $\int \psi_N = 1$ and $\psi_N' \in L^1$. Suppose further that $\hat{\psi}_N$ is bounded, even $\hat{\psi}_N(r) = 0$ for $|r| > N$. We can then approximate the Fourier multiplier operator $|\nabla|$ (i.e. the filter in the Radon inversion formula) by

$$P_{\leq N}|\nabla| := \mathcal{F}^{-1}[\hat{\psi}_N(r)|r|]\mathcal{F}.$$

Proposition 2. *Then the approximation of f given by*

$$f_N(x) := \tfrac{1}{2\pi} \int_0^\pi P_{\leq N}|\nabla|\mathcal{R}f(\langle x, \omega\rangle, \omega)\, d\omega \tag{6}$$

is given by the convolution

$$f_N = K_N * f, \quad \text{where} \quad K_N(\rho) = -\tfrac{1}{\pi} \int_\rho^\infty \frac{\psi_N'(t)}{\sqrt{t^2 - \rho^2}}\, dt. \tag{7}$$

Proof. This may be skipped in lecture.

We express $K_N * f$ using the filtered back-projection formula:

$$K_N * f(x) = \tfrac{1}{4\pi^2} \int_0^\pi \int_{-\infty}^\infty \widetilde{\mathcal{R}[K_N * f]}(r, \omega) e^{irx\cdot\omega}|r|\, dr\, d\omega.$$

In the lemma below, we will show that

$$\mathcal{R}[K_N * f](t, \omega) = \int_{-\infty}^\infty \mathcal{R}K_N(s, \omega)\mathcal{R}f(t - s, \omega)\, ds.$$

Thus, using the convolution identity for the $1d$ Fourier transform, we obtain

$$K_N * f(x) = \tfrac{1}{4\pi^2} \int_0^\pi \int_{-\infty}^\infty \widetilde{\mathcal{R}f}(r, \omega)\widetilde{\mathcal{R}K_N}(r, \omega) e^{irx\cdot\omega}|r|\, dr\, d\omega,$$

and so the result follows provided we can establish

$$\mathcal{R}K_N(r, \omega) = \psi_N(r) \tag{8}$$

(as this yields $\widetilde{\mathcal{R}K_N} = \hat{\psi}_N$). The verification of (8) is a direct computation, which also fits into the theory of the Radon transform restricted to radial functions. We will not pursue this topic in these notes. □

In the preceding proof, we needed the following lemma, which states that the Radon transform converts convolution in the plane to convolution in the affine parameter.

Lemma 1. *For f, g piecewise continuous of bounded support,*

$$\mathcal{R}[f * g](t, \omega) = \int_{-\infty}^\infty \mathcal{R}f(s, \omega)\mathcal{R}g(t - s, \omega)\, ds.$$

Proof. This proof may be skipped in lecture.

We first write

$$\mathcal{R}(f * g)(t, \omega) = \int f * g(t\omega + s\hat{\omega})\, ds.$$

Now we compute, using the coordinates $(a, b) \mapsto (a\hat{\omega} + b\omega)$ (which has area element $da\, db$):

$$f * g(t\omega + s\hat{\omega}) = \int_{\mathbb{R}^2} f(y)g(t\omega + s\hat{\omega} - y)\, dy$$

$$= \iint f(a\hat{\omega} + b\omega) g((s-a)\hat{\omega} + (t-b)\omega)\, da\, db.$$

Thus by changing the order of integration and changing variables in the ds integral, we obtain

$$\begin{aligned}
\mathcal{R}(f * g)(t, \omega) &= \iiint f(a\hat{\omega} + b\omega) g((s-a)\hat{\omega} + (t-b)\omega)\, da\, db\, ds \\
&= \int \left[\int f(a\hat{\omega} + b\omega) \left[\int g((s-a)\hat{\omega} + (t-b)\omega)\, ds \right] da \right] db \\
&= \int \mathcal{R}f(b, \omega) \mathcal{R}g(t-b, \omega)\, db,
\end{aligned}$$

as desired. □

Proposition 2 shows that the approximation f_N obtained is a blurred version of the real image f. As N increases, we decrease the amount of blur but also reduce the suppression of noise in the data, which is a type of tradeoff we have encountered before.

Remark 7. *Unfortunately, this result excludes the choice $\hat{\psi} = \chi_{[-N,N]}$, as then $\psi = N \operatorname{sinc}(N\cdot) \notin L^1$. Nonetheless, in this case one can compute the integral defining K_N explicitly (as a convergent improper integral) by using the theory of Bessel functions.*

We finally turn to some continuity (or, equivalently, boundedness) properties for the Radon transform and its inverse. Continuity properties for the Radon transform are important to guarantee that the measurement process is stable. Continuity properties for the inverse map are important to guarantee that the reconstruction process is stable. We will not attempt a comprehensive treatment of these issues. Instead we will try to present just a few representative results. A related issue is characterizing the range of the Radon transform, which can yield consistency conditions for our measured data; however, we will not discuss this topic here.

For the Radon transform itself, let us prove the following straightforward estimate. We will consider functions supported where $\{|x| \leq L\}$ for some $L > 0$, in which case $\mathcal{R}f(t, \omega) = 0$ for any $|t| > L$.

Proposition 3. *Suppose $f \in L^2(\mathbb{R}^2)$ and f is supported in $\{|x| \leq L\}$. Then for all $\omega \in \mathbb{S}^1$, we have*

$$\int_{-L}^{L} |\mathcal{R}f(t, \omega)|^2\, dt \leq 2L \|f\|_{L^2}^2.$$

Proof. Using the support properties of f and the Cauchy–Schwarz inequality we estimate

$$\begin{aligned}
\int_{-L}^{L} |\mathcal{R}f(t, \omega)|^2\, dt &= \int_{-L}^{L} \left| \int_{-L}^{L} f(t\omega + s\hat{\omega})\, ds \right|^2 dt \\
&\leq 2L \int_{-L}^{L} \int_{-L}^{L} |f(t\omega + s\hat{\omega})|^2\, ds\, dt = 2L \|f\|_{L^2}^2.
\end{aligned}$$

□

Remark 8. *A more careful argument implies the slightly stronger bound*

$$\int_{-L}^{L} \frac{|\mathcal{R}f(t, \omega)|^2}{\sqrt{L^2 - t^2}}\, dt \leq 2 \|f\|_{L^2}^2.$$

We turn to the inverse transform. In order to avoid several subtle mathematical issues, we utilize the *approximate* inverses introduced above. In particular, we fix $N \gg 1$ and ψ_N as above, and we define corresponding operators

$$\mathcal{R}_N^{-1} F(x) := \tfrac{1}{2\pi} \int_0^\pi P_{\leq N} |\nabla| F(\langle x, \omega \rangle, \omega)\, d\omega,$$

for $F \in L^1(\mathbb{R} \times \mathbb{S}^1)$, where $P_{\leq N}|\nabla| := \mathcal{F}^{-1}[\hat{\psi}_N(r)|r|]\mathcal{F}$. The limit $N = \infty$ and $\hat{\psi} \equiv 1$ formally recovers the true inverse transform.

Suppose we make a measurement Rf_m, which approximates the Radon transform $\mathcal{R}f$ of the unknown image f. Our reconstructed image is given by

$$f_{app} := \mathcal{R}_N^{-1} R f_m,$$

and our interest is in estimating the difference between f and f_{app} in terms of the difference between $\mathcal{R}f$ and Rf_m. As we will see, this relies on continuity properties for $\mathcal{R}_N^{-1}$.

Proposition 4. *We have the following pointwise bound:*

$$|f - f_{app}| \leq |f - K_N * f| + C_N \|\mathcal{R}f - Rf_m\|_{L^1(\mathbb{R}\times\mathbb{S}^1)}, \tag{9}$$

where K_N is as in (7) *and*

$$C_N := \tfrac{1}{2\pi} \|\mathcal{F}^{-1}[\hat{\psi}_N(r)|r|]\|_{L^\infty}.$$

Proof. This proof may be skipped in lecture.

Recall from Proposition 2 that $\mathcal{R}_N^{-1}\mathcal{R}f = K_N * f$. Thus

$$f - f_{app} = f - K_N * f + \mathcal{R}_N^{-1}[\mathcal{R}f - Rf_m],$$

and we are faced with estimating the second term.

In particular, we need to prove an $L^1 \to L^\infty$ bound for $\mathcal{R}_N^{-1}$ with constant given by C_N. To this end, we observe that

$$P_{\leq N}|\nabla| f = g_N * f, \quad \text{where} \quad g_N = \mathcal{F}^{-1}[\hat{\psi}_N(r)|r|],$$

so that

$$\begin{aligned} \mathcal{R}_N^{-1} F(x) &= \tfrac{1}{2\pi} \int_0^\pi (g_N *_t F)(\langle x, \omega \rangle, \omega)\, d\omega \\ &= \tfrac{1}{2\pi} \int_0^\pi \int_{\mathbb{R}} g_N(\langle x, \omega \rangle - s) F(s, \omega)\, ds\, d\omega. \end{aligned}$$

Thus

$$\|\mathcal{R}_N^{-1} F\|_{L^\infty} \leq \tfrac{1}{2\pi} \int_0^\pi \int_{\mathbb{R}} \|g_N\|_{L^\infty} |F(s,\omega)|\, ds\, d\omega \leq C_N \|F\|_{L^1(\mathbb{R}\times\mathbb{S}^1)},$$

yielding the desired bound. □

The first term on the right-hand side of (9) can be made arbitrarily small by choosing N sufficiently large, at least at points of continuity of f. As for estimating C_N, we have (by the support properties of $\hat{\psi}_N$)

$$\|\mathcal{F}^{-1}[\hat{\psi}_N(r)|r|]\|_{L^\infty} \leq \|\hat{\psi}_N(r)|r|\|_{L^1} \leq N^2 \|\hat{\psi}_N\|_{L^\infty} \leq N^2 \|\psi_N\|_{L^1}.$$

Thus we expect this component of the error to *increase* as N increases. Nonetheless, Proposition 4 does achieve our stated goal, in that it allows us to estimate the difference between the true image and the reconstructed image in terms of the measurement error.

Fourier series

So far, we have been working with a *continuous* model for measurement and reconstruction in X-ray CT. In practice, however, our measurement data will be a finite, discrete set of points, and so it will be important to build a discrete model. Our first step in this direction is to introduce the theory of Fourier series.

To simplify formulas, we will initially work with functions defined on the interval $[0,1] \subset \mathbb{R}$. We will discuss the extension to general bounded intervals and higher dimensions at the end of the section.

The general problem we consider is that of representing functions by Fourier series, i.e. as a linear combination of complex exponentials $e^{2\pi inx}$, where $n \in \mathbb{Z}$. Such a function automatically admits a 1-periodic extension to all of $\mathbb{R}$, as $e^{2\pi in(x+1)} \equiv e^{2\pi inx}$ for all $n \in \mathbb{Z}$ and $x \in [0,1]$. Accordingly, it is natural to consider the *1-periodic extension* of $f : [0,1) \to \mathbb{C}$, defined by $f(x+n) = f(x)$ for $x \in [0,1)$ and $n \in \mathbb{Z}$. We note that if f is a continuous function on $[0,1]$, then its 1-periodic extension is continuous if and only if $f(0) = f(1)$. Similar assertions (using limits from the left/right) may be made concerning differentiability.

Example 1. *The function $f(x) = x$ is continuous on $[0,1]$ but its 1-periodic extension is not.*

The function $g(x) = x(1-x)$ is continuous on $[0,1]$ with a 1-periodic extension; however, its derivative $g'(x) = 1 - 2x$ is not continuous as a 1-periodic function.

The function $h(x) = e^{2\pi ix}$, along with all of its derivatives, is continuous as a 1-periodic function.

Given a function $f : [0,1] \to \mathbb{C}$, we seek to represent f in the form

$$f(x) = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{2\pi inx} \tag{1}$$

for some complex coefficients $\hat{f}(n)$ (called the *Fourier coefficients* of f). Observing that

$$\int_0^1 e^{2\pi inx} e^{-2\pi imx}\, dx = \delta_{nm} := \begin{cases} 1 & n = m \\ 0 & n \neq m, \end{cases}$$

we derive that if such a decomposition holds, the Fourier coefficients should be defined by

$$\hat{f}(n) := \int_0^1 f(x) e^{-2\pi inx}\, dx.$$

We observe that for $f \in L^1([0,1])$, the Fourier coefficients form a bounded sequence of complex numbers. Indeed,

$$\sup_{n \in \mathbb{Z}} |\hat{f}(n)| \leq \|f\|_{L^1}.$$

Example 2. *If $f(x) = \cos(2\pi n_0 x)$, then*

$$\hat{f}(n) = \begin{cases} \frac{1}{2} & n = \pm n_0 \\ 0 & n \neq \pm n_0 \end{cases}$$

In this case, (1) *reduces to the well-known formula*

$$\cos(2\pi n_0 x) = \tfrac{1}{2}[e^{2\pi in_0 x} + e^{-2\pi in_0 x}].$$

Example 3. *Let $f(x) = \chi_{[a,b](x)}$ for some $0 \leq a < b < 1$. Then*

$$\hat{f}(n) = \begin{cases} b-a & n = 0, \\ \frac{1}{2\pi i n}[e^{-2\pi i n a} - e^{2\pi i n b}] & n \neq 0. \end{cases}$$

In general, because $\hat{f}(n)$ decays like $\frac{1}{n}$, the series (1) will not converge absolutely (e.g. take $b - a = \frac{1}{4}$). Thus we will need to be careful about how we interpret a statement like (1).

Using the definition of the Fourier coefficients, one can readily verify a few symmetry properties. For example, if f is real-valued, then

$$\hat{f}(-n) = \overline{\hat{f}(n)}.$$

Moreover, real-valued even functions have real Fourier coefficients, and real-valued odd functions have purely imaginary coefficients, where we say $f : [0,1] \to \mathbb{R}$ is *even* if $f(x) = f(1-x)$ and *odd* if $f(x) = -f(1-x)$.

To study the convergence properties of Fourier series as in (1), we introduce the following partial sum operators. For $N \in \mathbb{N}$, we define the partial sum operator $S_N : L^1([0,1]) \to C([0,1])$ by

$$S_N(f;x) = \sum_{|n| \leq N} \hat{f}(n) e^{2\pi i n x}.$$

For each fixed N, this defines a continuous linear operator.

Our goal is to establish the convergence of $S_N(f)$ to f in suitable topologies as $N \to \infty$. In general, *pointwise* convergence of $S_N(f;x)$ to $f(x)$ may fail, even at points of continuity of f. For smooth enough functions, however, we will obtain decay for the Fourier coefficients, which will in turn allow us to establish convergence. Our first result in this direction is the following.

Theorem 1 (Fourier inversion formula)**.** *Suppose $f \in C([0,1])$ and $\{\hat{f}(n)\} \in \ell^1$. Then*

$$f(x) = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{2\pi i n x} \quad \textit{for all} \quad x \in [0,1].$$

Proof of Theorem 1. We will use an approximate identity argument.

For $r \in (0,1)$, let us define the convergent series

$$f_r(x) = \sum_{n \in \mathbb{Z}} \hat{f}(n) r^{|n|} e^{2\pi i n x}.$$

Using absolute convergence and the fact that $\lim_{r \to 1} r^n = 1$ for each $n \in \mathbb{Z}$, we may derive that

$$\lim_{r \to 1} f_r(x) = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{2\pi i n x}$$

uniformly for $x \in [0,1]$. Thus to prove the theorem it suffices to establish $f_r(x) \to f(x)$ as $r \to 1$.

For this, we compute

$$f_r(x) = \sum_{n \in \mathbb{Z}} \hat{f}(n) r^{|n|} e^{2\pi i n x} = \int_0^1 f(y) \left[\sum_{n \in \mathbb{Z}} r^{|n|} e^{2\pi i n (x-y)} \right] dy,$$

where we have used absolute convergence to interchange summation and integration. We may therefore write

$$f_r(x) = \int_0^1 f(y) P_r(x-y)\, dy,$$

where

$$P_r(x) = \sum_{n\in\mathbb{Z}} r^{|n|} e^{2\pi i n x} = 1 + 2\,\mathrm{Re} \sum_{n=1}^{\infty} r^n e^{2\pi i n x}.$$

Using the formula for the sum of a geometric series ($\sum_{n=1}^{\infty} \rho^n = \frac{\rho}{1-\rho}$) and some algebra, we may obtain the following:

$$P_r(x) = \frac{1-r^2}{1-2r\cos(2\pi x)+r^2}.$$

We now observe that P_r are positive functions obeying

$$\int_0^1 P_r(x)\, dx = 1 \quad \text{for all} \quad r \in (0,1)$$

and $\lim_{r\to 1} P_r(x) = 0$ for all $x \neq 0$. Thus, using approximate identity arguments (as in previous sections), we may establish that for $f \in C([0,1])$,

$$f_r(x) = \int_0^1 P_r(x-y) f(y)\, dy \to f(x) \quad \text{as} \quad r \to 1,$$

as desired. □

The previous result assumed a decay condition on the Fourier coefficients, namely, $\{\hat{f}(n)\} \in \ell^1$. While we cannot expect such decay for arbitrary $f \in L^1$, we can establish the following.

Theorem 2 (Riemann–Lebesgue lemma)**.** *If $f \in L^1([0,1])$, we have $\hat{f}(n) \to 0$ as $|n| \to \infty$.*

Proof of Theorem 2. We can prove this result by using the Riemann–Lebesgue lemma for the Fourier transform. Indeed, if we set $F = f\chi_{[0,1]} \in L^1(\mathbb{R})$, then we have $\hat{f}(n) = \hat{F}(2\pi n)$, where $\hat{F}$ is the Fourier transform of F. □

The Riemann–Lebesgue does *not* give any quantitative decay rate for the Fourier coefficients. Such a result would be impossible, since for any sequence $\{a_n\}$ with $\lim_{|n|\to\infty} a_n = 0$, one may find $f \in L^1$ such that $|\hat{f}(n)| \geq |a_n|$ for all n.

On the other hand, we can relate smoothness properties of f to decay properties for $\hat{f}(n)$ (and vice versa), just as in the case of the Fourier transform. We record some results of this type in the following proposition.

Proposition 1 (Exchange of smoothness/decay)**.** *If $f \in C^k([0,1])$ and $f_+^{(j)}(0) = f_-^{(j)}(1)$ for $j = 0, \ldots, k-1$ (where $\pm$ denote limits from the right/left), then*

$$\hat{f}(n) = \frac{1}{(2\pi i n)^k} \widehat{f^{(k)}}(n) \quad \text{for} \quad n \neq 0,$$

so that $|\hat{f}(n)| = o(|n|^{-k})$ as $|n| \to \infty$.

Conversely, if $f \in L^1([0,1])$ and there exist $C > 0$ and $\varepsilon > 0$ such that

$$|\hat{f}(n)| \leq C[1+|n|]^{-(k+\varepsilon)},$$

then $f \in C^{k-1}([0,1])$ *with* $f_+^{(j)}(0) = f_-^{(j)}(1)$ *for* $j = 1, \dots, k-1$.

We will not prove Proposition 1 in its entirety. Let us only demonstrate the main idea of the proof.

Partial proof. We use the identity

$$e^{-2\pi inx} = -\tfrac{1}{2\pi in}\tfrac{d}{dx}e^{-2\pi inx}$$

to integrate by parts, yielding

$$\begin{aligned}\hat{f}(n) &= \int_0^1 f(x)e^{-2\pi inx}\,dx \\ &= \tfrac{1}{2\pi in}[f(0) - f(1)] + \tfrac{1}{2\pi in}\widehat{f'}(n).\end{aligned}$$

In particular, if f is periodic then the first term vanishes. The extension to higher order derivatives follows from further integration by parts. □

Analogous to the case of the Fourier transform, functions with jump discontinuities exhibit a typical $\frac{1}{|n|}$ decay rate in their Fourier coefficients.

Example 4. *Let* $f(x) = x(1-x)$ *for* $x \in [0,1]$. *As we saw above, this function is periodic but its derivative* $f'(x) = 1 - 2x$ *is not. We have*

$$\hat{f}(n) = \begin{cases} \frac{1}{6} & n = 0 \\ -\frac{1}{2\pi^2 n^2} & n \neq 0 \end{cases} \quad \textit{and} \quad \widehat{f'}(n) = \begin{cases} 0 & n = 0 \\ \frac{1}{\pi in} & n \neq 0, \end{cases}$$

showing that $\widehat{f'}(n) = (2\pi in)\hat{f}(n)$. *Note also that while* $\hat{f}(n) = \mathcal{O}(|n|^{-2})$, *we have that* f *fails to be* C^1 *as a periodic function, showing that the* $\varepsilon > 0$ *appearing in the preceding proposition is indeed necessary.*

As was the case of the Fourier transform, there is a very natural L^2 theory for Fourier series. Note that since $L^2([0,1]) \subset L^1([0,1])$, we can already define the Fourier coefficients of an L^2-function as a bounded sequence that decays as $n \to \infty$. We also have the following Parseval formula:

Theorem 3 (Parseval formula)**.** *If* $f \in L^2([0,1])$, *then*

$$\int_0^1 |f(x)|^2\,dx = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2.$$

To prove Theorem 3, we will need some preliminary lemmas.

The first result we need is Bessel's inequality, which is equivalent to the statement that the partial Fourier series represent the orthogonal projection onto the subspace of trigonometric polynomials. To make this precise, let us define

$$V_N = \left\{ \sum_{|n|\leq N} a_n e^{2\pi inx} : a_n \in \mathbb{C} \right\} \quad \text{for} \quad N \in \mathbb{N}.$$

Then each V_N is a finite-dimensional closed subspace of the Hilbert space $L^2([0,1])$, and we have the following result:

Lemma 1 (Bessel's inequality)**.** *For $N \in \mathbb{N}$, the partial sum operator S_N is the orthogonal projection of L^2 onto V_N. In particular, for any $\{a_n\}_{|n|\le N} \subset \mathbb{C}$, we have*

$$\left\| f - \sum_{|n|\le N} \hat{f}(n) e^{2\pi i n x} \right\|_{L^2}^2 \le \left\| f - \sum_{|n|\le N} a_n e^{2\pi i n x} \right\|_{L^2}^2, \tag{2}$$

with equality if and only if $a_n = \hat{f}(n)$ for all $|n| \le N$.

Consequently, $\{\hat{f}(n)\}_{n\in\mathbb{Z}} \in \ell^2$, with $\|\hat{f}(n)\|_{\ell^2} \le \|f\|_{L^2}$.

Proof. By definition of $\hat{f}(n)$ and orthonormality of $\{e^{2\pi i x k}\}_{k=-N}^{N}$, we have that

$$\langle e^{2\pi i x k}, f - \sum_{|n|\le N} \hat{f}(n) e^{2\pi i n x} \rangle = 0 \quad \text{for all} \quad |k| \le N.$$

This implies that $S_N f$ is the orthogonal projection onto V_N, which in turn implies (2). In fact, by orthogonality, we have that to obtain equality in (2), we should add

$$\left\| \sum_{|n|\le N} [a_n - \hat{f}(n)] e^{2\pi i n x} \right\|_{L^2}^2$$

to the left-hand side. Specializing to $a_n \equiv 0$, we obtain

$$\sum_{|n|\le N} |\hat{f}(n)|^2 = \left\| \sum_{|n|\le N} \hat{f}(n) e^{2\pi i n x} \right\|_{L^2}^2 \le \|f\|_{L^2}^2$$

uniformly in N, which implies the desired ℓ^2 bound on the Fourier coefficients. □

The other key ingredient in the proof of the Parseval formula is the fact that trigonometric polynomials are dense in L^2. To derive this fact, we will actually use a quantity related to the partial Fourier sums, namely, the *Fejér means* of the function.

We begin by taking a closer look at the partial Fourier series. Fix a periodic function $f : [0,1] \to \mathbb{C}$. Proceeding as in the proof of the inversion formula, we may write

$$S_N f(x) = \int_0^1 f(y) \left[\sum_{|n|\le N} e^{2\pi i n (x-y)} \right] dy, \tag{3}$$

which we may express as the convolution

$$S_N f(x) = f * D_N(x), \quad \text{where} \quad D_N(x) := \sum_{|n|\le N} e^{2\pi i n x}. \tag{4}$$

Remark 1. *If f and g are periodic functions on $[0,1]$, we may define the convolution*

$$f * g(x) = \int_0^1 f(y) g(x-y)\, dy$$

as a periodic function on $[0,1]$. The periodic convolution shares many properties with the convolution on $\mathbb{R}$. We also have the following product formula for the coefficients:

$$\widehat{f * g}(n) = \hat{f}(n)\hat{g}(n), \tag{5}$$

which can be checked by direct computation.

We call D_N the *Dirichlet kernel*. By using the formula for the sum of a geometric series, we may also write

$$D_N(x) = \frac{\sin[2\pi(N+\frac{1}{2})x]}{\sin(\pi x)}.$$

Observe that by the representation in (4), we have that

$$\hat{D}_N(n) = \begin{cases} 1 & |n| \leq N \\ 0 & |n| > N. \end{cases}$$

The kernels D_N share some properties with approximate identities as $N \to \infty$. For example,

$$\int D_N(x)\,dx = \hat{D}_N(0) = 1 \quad \text{for all} \quad N,$$

and the kernels become increasingly concentrated around $x = 0$ as $N \to \infty$. However, one can also prove that

$$\int_0^1 |D_N(x)|\,dx \approx \log N \to \infty \quad \text{as} \quad N \to \infty,$$

and so we cannot apply the usual approximate identity arguments with D_N. Instead of studying the Dirichlet kernels directly, we take an average in N. As we will see, this leads to a well-behaved family of kernels.

Definition 1. *The* Fejér kernels *are defined by*

$$F_N(x) = \tfrac{1}{N+1}\sum_{j=0}^{N} D_j(x).$$

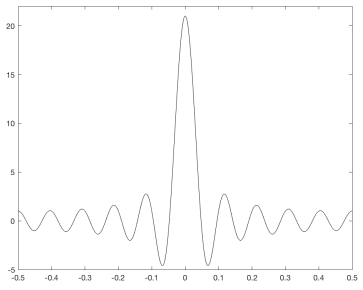
These are given explicitly by

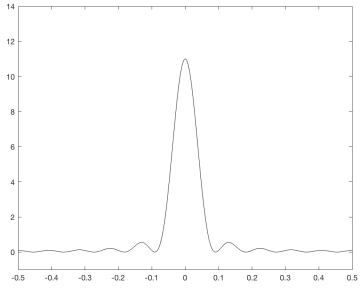
$$F_N(x) = \tfrac{1}{N+1}\left[\frac{\sin[\pi(N+1)x]}{\sin(\pi x)}\right]^2.$$

The Fejér means *of a function $f : [0,1] \to \mathbb{C}$ are defined by*

$$\sigma_N(f;x) = \tfrac{1}{N}\sum_{j=0}^{N} S_j(f;x) = f * F_N(x).$$

The following shows a comparison between the Dirichlet and Fejér kernels (with $N = 20$):





The Fejér kernels are amenable to approximate identity arguments, as one can verify that the L^1-norms of the F_N are uniformly bounded in N. In particular, we can prove the following:

Theorem 4 (Fejér's Theorem)**.** *Let $f \in L^1([0,1]) \cap C([0,1])$. Then*

$$\lim_{N\to\infty} \sigma_N(f;x) = f(x)$$

uniformly over $x \in [0,1]$.

Before proving this theorem, let us record an important corollary, which is the second main ingredient in the proof of Parseval's formula.

Corollary 1. *Trigonometric polynomials are dense in L^2. That is, for $f \in L^2([0,1])$ and $\varepsilon > 0$, there exists a trigonometric polynomial*

$$g(x) = \sum_{|n|\le N} a_n e^{2\pi i n x}$$

(for some $N \in \mathbb{N}$ and $\{a_n\} \subset \mathbb{C}$) such that

$$\|f - g\|_{L^2} < \varepsilon.$$

Proof. Using convolution with approximate identities, we may first find $\tilde{f} \in C([0,1])$ such that

$$\|f - \tilde{f}\|_{L^2} < \tfrac{1}{2}\varepsilon.$$

By Fejér's theorem, there exists N such that $g := \tilde{f} * F_N$ obeys

$$\sup_{x\in[0,1]} |\tilde{f}(x) - g(x)| < \tfrac{1}{2}\varepsilon,$$

which implies

$$\|\tilde{f} - g\|_{L^2} < \tfrac{1}{2}\varepsilon.$$

The result will now follow once we verify that g is a trigonometric polynomial. For this, we use (3) to write

$$\begin{aligned} g(x) = \tilde{f} * F_N(x) &= \tfrac{1}{N+1} \sum_{j=0}^{N} f * D_j(x) \\ &= \tfrac{1}{N+1} \sum_{j=0}^{N} \int_0^1 f(y) \sum_{|n|\le j} e^{2\pi i n(x-y)}\, dy \\ &= \tfrac{1}{N+1} \sum_{j=0}^{N} \sum_{|n|\le j} \left[\int_0^1 f(y) e^{-2\pi i n y} \right] e^{2\pi i n x}, \end{aligned}$$

which shows that g is indeed a linear combination of complex exponentials. □

We turn to the proof of Fejér's theorem.

Proof of Fejér's theorem. We use the usual approximate identity argument. We write

$$\sigma_N(f;x) - f(x) = \int_0^1 F_N(y)[f(x-y) - f(x)]\, dy.$$

The contribution near $y = 0$ is small using the continuity of f and the uniform L^1-bounds on the F_N. For y away from zero, we use boundedness of f and the fact that for any $\delta > 0$, $F_N(y) \to 0$ uniformly on $\{\delta < |y| < 1\}$ as $N \to \infty$. □

We are now in a position to prove the Parseval formula.

Proof of the Parseval formula. We first show that the Parseval formula holds for an arbitrary trigonometric polynomial. Indeed, if

$$f(x) = \sum_{|n|\leq N} a_n e^{2\pi inx}$$

(so that $\hat{f}(n) = a_n$), then (using orthonormality of $\{e^{2\pi inx}\}$) we have

$$\int_0^1 |f(x)|^2\,dx = \sum_{n,m} \bar{a}_n a_m \int_0^1 e^{2\pi i(n-m)}\,dx = \sum_{|n|\leq N} |a_n|^2,$$

giving the result in this case.

Next, we let $f \in L^2$. By Bessel's inequality, we have

$$\sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2 \leq \|f\|_{L^2}^2.$$

Thus it suffices to show that for any $\varepsilon > 0$, we have

$$\|f\|_{L^2} \leq \left[\sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2\right]^{\frac{1}{2}} + \varepsilon.$$

To this end, we let $\varepsilon > 0$ and choose a trigonometric polynomial

$$g(x) = \sum_{|n|\leq N} a_n e^{2\pi inx}$$

obeying

$$\|g - f\|_{L^2} < \varepsilon.$$

Now, using the triangle inequality and Bessel's inequality, we have

$$\begin{aligned}\|f\|_{L^2} &\leq \|f - S_N(f)\|_{L^2} + \|S_N(f)\|_{L^2}\\ &\leq \|f - g\|_{L^2} + \left[\sum_{|n|\leq N} |\hat{f}(n)|^2\right]^{\frac{1}{2}}\\ &\leq \varepsilon + \left[\sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2\right]^{\frac{1}{2}},\end{aligned}$$

as desired. □

Let us now return to the question of convergence of Fourier series. For a typical L^2 function, we only know that $\{\hat{f}(n)\}$ belongs to ℓ^2, and hence the Fourier series cannot be expected to converge pointwise in general. Nonetheless, we do have an L^2 inversion formula.

Proposition 2 (Fourier inversion in L^2)**.** *For $f \in L^2([0,1])$, we have*

$$\lim_{M,N\to\infty} \left\| f - \sum_{j=-M}^{N} \hat{f}(j) e^{2\pi ijx} \right\|_{L^2} = 0.$$

Proof. This proof may be skipped in lecture.

Using the orthogonal projection property for partial Fourier series, we first observe that

$$\left\| f - \sum_{j=-M}^{N} \hat{f}(j) e^{2\pi i j x} \right\|_{L^2}^2 = \|f\|_{L^2}^2 - \sum_{j=-M}^{N} |\hat{f}(j)|^2.$$

By the Parseval formula, we may deduce that the right-hand side tends to zero as $N, M \to \infty$. □

While L^2-norm convergence seems less simple than pointwise convergence, it is actually reasonable to use in problems where measurement is a serious concern. In particular, any $g \in L^2$ can be used to define a 'measurement' $m_g(f) = \langle f, g \rangle$ (and any continuous linear 'measurement' is of this form). Then L^2 convergence guarantees that $m_g(S_N(f)) \to m_g(f)$ for any $g \in L^2$.

Example 5. *If we let*

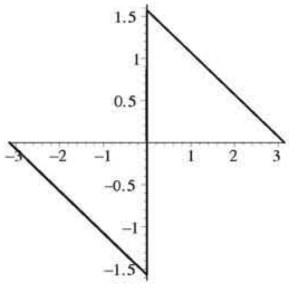
$$g_\varepsilon(y) = \begin{cases} 0 & |x-y| > \varepsilon \\ \frac{1}{2\varepsilon} & |x-y| \le \varepsilon, \end{cases}$$

then we have $\int g_\varepsilon = 1$*, and we use* g_ε *to model a measurement near* x *with resolution of size* ε*. In place of the pointwise difference* $S_N(f;x) - f(x)$*, we may consider the average*

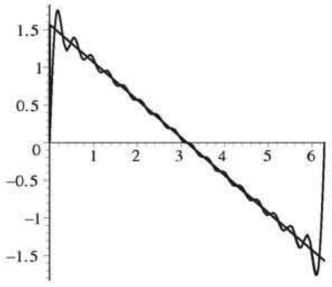
$$\int_0^1 |f(y) - S_N(f;y)|\, g_\varepsilon(y)\, dy \le \|f - S_N(f)\|_{L^2} \|g_\varepsilon\|_{L^2} = \tfrac{1}{\sqrt{2\varepsilon}} \|f - S_N(f)\|_{L^2}.$$

Thus for a fixed resolution, the measured error tends to zero as $N \to \infty$*.*

The failure of pointwise convergence of Fourier series is related to an important phenomenon known as the *Gibbs phenomenon*. In particular, near jump discontinuities the Fourier series exhibit oscillatory behavior that consistently overshoots the jump in a 'universal' way. An example is shown in the following figure from [1].



(a) A periodic function with a jump discontinuity.



(b) Fourier reconstruction with $N = 16$.

A precise formulation is the following:

Theorem 5 (Gibbs phenomenon)**.** *Suppose* f *is a piecewise continuously differentiable function with a jump discontinuity at* $x = x_0$ *of size* $h > 0$*, i.e.*

$$\lim_{x \to x_0^+} f(x) = \lim_{x \to x_0^-} f(x) + h.$$

Then

$$\lim_{N\to\infty} S_N(f;x_0) = \lim_{x\to x_0^-} f(x) + \tfrac{1}{2}h,$$

while for any small $\varepsilon > 0$, we have

$$\lim_{N\to\infty} \max_{0<x-x_0<\varepsilon} \big[S_N(f;x) - f(x)\big] = (G-1)\tfrac{h}{2},$$

where G is the universal constant given by

$$G := \tfrac{2}{\pi}\int_0^1 \tfrac{\sin \pi t}{t}\, dt = 1.17898...$$

Remark 2. *This shows that near a jump discontinuity, the Fourier series overshoot the jump by around 9% of the height of the jump. In particular, the size of the overshoot does not decrease as $N \to \infty$ (although it does concentrate in an interval of size $\sim \frac{1}{N}$ around the discontinuity).*

We will not prove this theorem. Instead, let us just consider one example that demonstrates this phenomenon. In fact, this example already contains the main idea of the proof of the more general result.

Example 6. *Consider the 2π-periodic function*

$$g(x) = \begin{cases} \frac{\pi}{2} - \frac{x}{2} & 0 \le x \le \pi \\ -\frac{\pi}{2} - \frac{x}{2} & -\pi \le x, < 0. \end{cases}$$

which has a jump of height $h = \pi$ at $x = 0$. Here we use the theory of Fourier series on $(-\pi, \pi)$, which will be discussed below. The Fourier series for g reduces to

$$g(x) = \sum_{k=1}^{\infty} \frac{\sin kx}{k}.$$

We would like to find the maximum of $S_N(g;x) - g(x)$, which we do by setting the derivative equal to zero. By explicit calculation, this reduces to the condition

$$\sin\big[(N+\tfrac{1}{2})x\big] = 0,$$

for which the smallest solution is

$$x_N := \tfrac{\pi}{N+\frac{1}{2}}.$$

We now note that

$$S_N(g;x_N) = \sum_{k=1}^{N} \frac{\sin\big[k(\frac{\pi}{N+\frac{1}{2}})\big]}{\frac{k}{N+\frac{1}{2}}} \frac{1}{N+\frac{1}{2}}$$

is a Riemann sum for the integral

$$\int_0^1 \frac{\sin(\pi x)}{x}\, dx = \tfrac{\pi}{2}G.$$

Thus (noting that $g(x_N) \to \frac{\pi}{2}$) we observe

$$S_N(g;x_N) - g(x_N) \to \tfrac{\pi}{2}[G-1] = \tfrac{h}{2}[G-1]$$

as $N \to \infty$.

One way to reduce the Gibbs effect is to use the Fejér kernels instead of the Dirichlet kernels. However, this comes at the expense of reducing the overall resolution.

♣ ♣ ♣

In this section, let us briefly discuss the extension of Fourier series to more general intervals, as well as to higher dimensions.

For a P-periodic function, the Fourier coefficients are given by

$$\hat{f}(n) = \int_0^P f(x) e^{-2\pi i n x/P}\, dx.$$

One then has the Fourier inversion formula

$$f(x) = \tfrac{1}{P} \sum_{n\in\mathbb{Z}} \hat{f}(n) e^{2\pi i n x/P}$$

and the Parseval formula

$$\int_0^P |f(x)|^2\, dx = \tfrac{1}{P} \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2.$$

One can derive explicit formulas for the Dirichlet and Fejér kernels, as well.

To extend the theory to higher dimensions, we define

$$\hat{f}(k) = \int_{[0,1]^n} f(x) e^{-2\pi i k\cdot x}\, dx,$$

where now $f : [0,1]^n \to \mathbb{C}$ and $k \in \mathbb{Z}^n$. The Riemann–Lebesgue lemma also holds in higher dimensions, and we have the same Fourier inversion formula provided the Fourier coefficients decay sufficiently. We also have the Parseval formula and a natural L^2-based theory.

Sampling theory

In this section, we discuss the topic of sampling functions of continuous variables.

A basic model for measurement is the evaluation of a continuous function f at a discrete set of points. A set $\{x_j\}$ is *discrete* if it has no convergent subsequence, and the values $\{f(x_j)\}$ are called *samples* of f. We frequently consider equally spaced samples at points, so that

$$x_j = x_0 + j\ell \quad \text{for} \quad j \in \mathbb{Z} \quad \text{and some} \quad \ell > 0.$$

In this case we call ℓ^{-1} the *sampling rate*.

A more realistic model for measurement of a function f is the evaluation of a convolution $f * \varphi$, where $\varphi \in L^1$. As we have discussed before, convolution is a smoothing operator that suppresses the high frequency content of f. In fact, in many applications we work with the assumption that f has no high-frequency content at all.

Definition 1. *A function $f : \mathbb{R} \to \mathbb{R}$ is* bandlimited *(or has* finite bandwidth*) if $\hat{f}$ is supported in a finite interval.*

Bandlimited functions are always smooth, as their Fourier transforms have arbitrarily good decay. In particular, for an L-bandlimited function $f \in L^2$ we have the Fourier inversion formula

$$f(x) = \tfrac{1}{2\pi} \int_{-L}^{L} \hat{f}(\xi) e^{ix\xi} \, d\xi. \tag{1}$$

The first main result in this section is Nyquist's theorem, which states that bandlimited functions are uniquely by a discrete set of uniformly spaced samples of sufficiently large sampling rate.

Theorem 1 (Nyquist's theorem)**.** *If $f \in L^2(\mathbb{R})$ and*

$$\hat{f}(\xi) = 0 \quad \textit{for} \quad |\xi| > L,$$

then f is uniquely determined by the samples $\{f(\frac{n\pi}{L}) : n \in \mathbb{Z}\}$.

If $\{f(\frac{n\pi}{L})\}_{n\in\mathbb{Z}} \in \ell^1$, then

$$f(x) = \sum_{n\in\mathbb{Z}} f(\tfrac{n\pi}{L}) \operatorname{sinc}(Lx - n\pi). \tag{2}$$

Proof. By (1), the numbers $\{2\pi f(-\frac{n\pi}{L})\}$ are the Fourier coefficients of $\hat{f}$, and so the Fourier inversion formula implies

$$\hat{f}(\xi) = \tfrac{\pi}{L} \sum_{n\in\mathbb{Z}} f(-\tfrac{n\pi}{L}) e^{in\pi\xi/L} = \tfrac{\pi}{L} \sum_{n\in\mathbb{Z}} f(\tfrac{n\pi}{L}) e^{-in\pi\xi/L} \quad \text{for} \quad |\xi| \le L,$$

with convergence in the L^2-sense. In particular,

$$\hat{f}(\xi) = \tfrac{\pi}{L} \left[\sum_{n\in\mathbb{Z}} f(\tfrac{n\pi}{L}) e^{-in\pi\xi/L} \right] \chi_{[-L,L]}(\xi).$$

Thus we see that the samples of f determine its Fourier transform, and hence f itself.

If the samples form an ℓ^1 sequence, then we can compute

$$f(x) = \tfrac{1}{2\pi} \int_{-L}^{L} \hat{f}(\xi) e^{ix\xi} \, d\xi$$

$$= \sum_{n\in\mathbb{Z}} f(\tfrac{n\pi}{L}) \tfrac{1}{2L} \int_{-L}^{L} e^{i\xi[x-n\pi/L]}\, d\xi$$
$$= \sum_{n\in\mathbb{Z}} f(\tfrac{n\pi}{L}) \operatorname{sinc}(Lx - n\pi).$$

□

Remark 1. *The exponentials $e^{\pm iLx}$ have frequency $\frac{L}{2\pi}$. Nyquist's theorem therefore states that if $\frac{L}{2\pi}$ is the highest frequency appearing in a function, then we must sample a function at the rate $\frac{L}{\pi}$ (the* Nyquist rate*) to completely determine f. Sampling at a lower/higher rate is called* under/oversampling*.*

The reconstruction formula in (2) is sometimes called the *Shannon–Whittaker interpolation formula*. To use this formula in practice, we would truncate the sum after finitely many terms. However, because

$$|\operatorname{sinc}(Lx - n\pi)| \approx \tfrac{1}{n},$$

the partial sums may converge to f very slowly. This situation can be improved by oversampling, as we now explain:

Oversampling. Suppose f is an $(L-\eta)$-bandlimited function for some $\eta > 0$. Let us choose a *window function* $\hat{\varphi}$ such that

$$\hat{\varphi} = \begin{cases} 1 & |\xi| \le L - \eta \\ 0 & |\xi| > L. \end{cases}$$

Repeating the computation in the proof of Theorem 1 and using the fact that $\hat{f}(\xi) = \hat{f}(\xi)\hat{\varphi}(\xi)$ for $|\xi| \le L$, we derive

$$\begin{aligned} f(x) &= \tfrac{1}{2L} \sum_{n\in\mathbb{Z}} f(\tfrac{n\pi}{L}) \int_{-L}^{L} \hat{\varphi}(\xi) e^{i\xi[x-n\pi/L]}\, d\xi \\ &= \tfrac{1}{2L} \sum_{n\in\mathbb{Z}} f(\tfrac{n\pi}{L}) \varphi(x - \tfrac{n\pi}{L}). \end{aligned} \tag{3}$$

This generalizes (2), which corresponds to choosing $\hat{\varphi} = \chi_{[-L,L]}$.

By choosing a smoother window function, we may obtain faster convergence of the series above (as a smoother Fourier transform corresponds to better decay of φ). On the other hand, φ will not be as simple as the explicit sinc function. Furthermore, we need to sample f above the Nyquist rate. Indeed, f is an $(L-\eta)$-bandlimited function, but we use the sample spacing $\frac{\pi}{L} < \frac{\pi}{L-\eta}$.

In applications, we will be sampling functions of bounded of support, which *cannot* be bandlimited. Thus, we will always be undersampling. To understand the effects of undersampling, we introduce a result known as the *Poisson summation formula.*

Theorem 2 (Poisson summation formula)**.** *If $f \in L^1$ is continuous and $\{\hat{f}(2\pi n)\} \in \ell^1(\mathbb{Z})$, then*

$$\sum_{n\in\mathbb{Z}} f(x+n) = \sum_{n\in\mathbb{Z}} \hat{f}(2\pi n) e^{2\pi i n x}. \tag{4}$$

Proof. The function on the left-hand side of (4) is 1-periodic and integrable on $[0,1]$. Thus it suffices to prove that the Fourier coefficients of this function are

given by $\hat{f}(2\pi m)$ (i.e. samples of the Fourier transform of f). We compute

$$\begin{aligned}\int_0^1 \sum_{n\in\mathbb{Z}} f(x+n)e^{-2\pi imx}\,dx &= \sum_{n\in\mathbb{Z}} \int_n^{n+1} f(x)e^{-2\pi imx} \\ &= \int_{\mathbb{R}} f(x)e^{-2\pi imx}\,dx = \hat{f}(2\pi m).\end{aligned}$$

□

Remark 2. *More generally, for a $2L$-periodic function we obtain*

$$\sum_{n\in\mathbb{Z}} f(x+2nL) = \tfrac{1}{2L}\sum_{n\in\mathbb{Z}} \hat{f}(\tfrac{n\pi}{L})e^{in\pi x/L}.$$

Using this, we observe that if f is supported in $[-L,L]$, then

$$f(x) = \tfrac{1}{2L}\sum_{n\in\mathbb{Z}} \hat{f}(\tfrac{n\pi}{L})e^{in\pi x/L} \quad \textit{for} \quad x\in[-L,L],$$

giving a dual version of the Nyquist theorem. This is relevant in MRI, where one actually measures samples of the Fourier transform of the image function.

Arguing similarly to the above, we can obtain the following 'dual Poisson summation formula', which will play a key role in our understanding of undersampling.

Theorem 3 (Dual Poisson summation)**.** *Suppose $f, \hat{f} \in L^1$ and $\{f(\frac{n\pi}{L})\} \in \ell^1$. Then*

$$\sum_{n\in\mathbb{Z}} \hat{f}(\xi+2nL) = \tfrac{\pi}{L}\sum_{n\in\mathbb{Z}} f(\tfrac{n\pi}{L})e^{-in\pi\xi/L}.$$

Undersampling. Given a function f and $L>0$, we may use the Shannon–Whittaker formula to define an L-bandlimited approximation to f, namely

$$F_L(x) := \sum_{n\in\mathbb{Z}} f(\tfrac{n\pi}{L})\,\mathrm{sinc}(Lx-n\pi), \tag{5}$$

which agress with f at the sample points $\frac{n\pi}{L}$. Then, by the same computations that led to the Shannon–Whittaker formula and the dual Poisson summation formula, we obtain

$$\hat{F}_L(\xi) = \sum_{n\in\mathbb{Z}} \hat{f}(\xi+2nL)\chi_{[-L,L]}(\xi).$$

Thus

$$\hat{f}(\xi) - \hat{F}_L(\xi) = \begin{cases} \hat{f}(\xi) & |\xi|>L \\ -\sum_{n\neq 0}\hat{f}(\xi+2nL) & |\xi|\le L. \end{cases} \tag{6}$$

In particular, we have that $\hat{f}\equiv\hat{F}_L$ if and only if f is L-bandlimited. Otherwise, we see two sources of error arising in the approximation F_L:

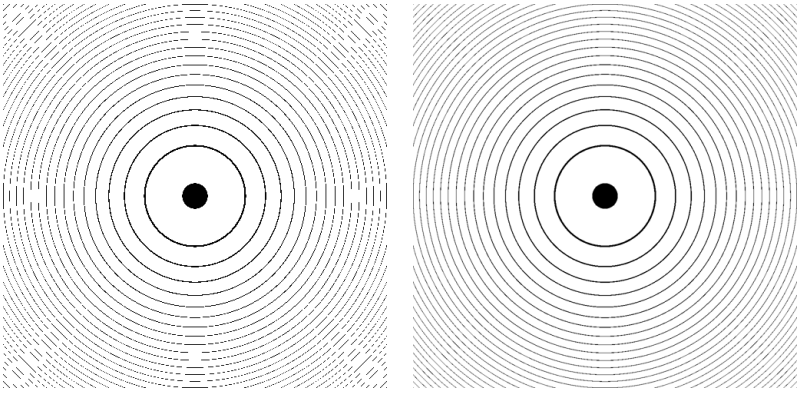
- First, we have the *truncation error*, which refers to the fact that the high-frequency information in f is not available in F_L.
- Next, we have the *aliasing effect*, which refers to the fact that the high-frequency information in f reappears in the low frequencies of F_L.

Example 1. *Movie images are sampled at some fixed rate. The tires of a driving car may rotate at a higher frequency, leading to undersampling and aliasing effects.*

Example 2. *The following images show two reconstructions of the function* $f : [0,1] \times [0,1] \to \mathbb{R}$ *defined by*

$$f(x) = \begin{cases} 1 & \left|15|x|^2 - \lfloor 15|x|^2 \rfloor\right| < \frac{1}{15} \\ 0 & \text{otherwise} \end{cases}$$

using a 500×500 *sample grid versus a* 1000×1000 *sample grid. Aliasing effects are clearly present in the reconstruction using the larger sample spacing.*



Continuing from above, we may express the approximation F_L using the inverse Fourier transform:

$$F_L(x) = \tfrac{1}{2\pi} \int_{-L}^{L} \hat{f}(\xi) e^{ix\xi}\, d\xi + \tfrac{1}{2\pi} \int_{-L}^{L} \sum_{n \neq 0} \hat{f}(\xi + 2nL) e^{ix\xi}\, d\xi.$$

Thus F_L is the sum of the partial Fourier inverse of f (which produces Gibbs oscillations if f has jump discontinuities) and the aliasing error.

As mentioned above, nontrivial functions with bounded support cannot be bandlimited. Nonetheless, if the function we sample is smooth enough, then its Fourier transform decays rapidly and so we may view the function as 'effectively bandlimited'. In applications, it is not enough that the Fourier transform itself becomes small; instead, we need to choose a high enough sampling rate that the entire aliasing error

$$\sum_{n \neq 0} \hat{f}(\xi + 2nL)$$

becomes small.

To diminish aliasing effects in general, we may first apply a *low-pass filter* to the signal being sampled. An *ideal* low-pass filter replaces f with f_L obeying

$$\hat{f}_L(\xi) = \begin{cases} \hat{f}(\xi) & |\xi| \leq L \\ 0 & |\xi| > L \end{cases}.$$

In this case, the Shannon–Whittaker reconstruction formula is just the partial Fourier inverse, i.e.

$$f_L(x) = \tfrac{1}{2\pi} \int_{-L}^{L} \hat{f}(\xi) e^{ix\xi}\, d\xi.$$

Thus the aliasing effect is removed, while any Gibbs oscillation effects will remain.

The measurement process itself can attenuate the high-frequency content of the signal. We may model measurement of a signal f as a convolution $\varphi * f$, where $\int \varphi = 1$. If φ is supported in $[-\eta, \eta]$, then the measurement of f at $\frac{n\pi}{L}$ is given by

$$\varphi * f(\tfrac{n\pi}{L}) = \int_{-\eta}^{\eta} f(\tfrac{n\pi}{L} - x)\varphi(x)\, dx.$$

As $\mathcal{F}[\varphi * f] = \hat{\varphi}\hat{f} \to 0$ when $|\xi| \to \infty$, we see that measurement attenuates the high frequencies. On the other hand, as $\hat{\varphi}(0) = 1$, we expect that $\mathcal{F}[\varphi * f]$ resembles $\hat{f}$ at low frequencies. The aliasing error in the measured samples is

$$\sum_{n \neq 0} \hat{\varphi}(\xi + 2nL)\hat{f}(\xi + 2nL),$$

which can be made small if φ is smooth.

Remark 3. *Signals are often sent along a* carrier *frequency (e.g. in FM radio or MRI). For example, we have a bandlimited signal f with Fourier support in $[\omega - B, \omega + B]$, where for simplicity we assume $\omega = NB$ for some $N \in \mathbb{N}$.*

Suppose we sample f at $\{\frac{n\pi}{B} : n \in \mathbb{Z}\}$ and build the bandlimited approximation F_B as in (5)*. By the dual Poisson summation formula,*

$$\hat{F}_B(\xi) = \hat{f}(\omega + \xi).$$

Thus $f(x) = e^{-i\omega x} F_B(x)$. We call F_B the demodulated *version of f.*

So far, we have discussed the problem of sampling continuous functions defined on the line. We turn to the problem of sampling functions defined on intervals (or their periodic extensions). This involves the theory of the *finite Fourier transform.*

Definition 2. *Let $\{x_j\}_{j=0}^{m-1} \subset \mathbb{C}$. The* finite Fourier transform *of $\{x_j\}$ is the sequence $\{\hat{x}_j\}_{j=0}^{m-1} \subset \mathbb{C}$ defined by*

$$\hat{x}_k = \tfrac{1}{m} \sum_{j=0}^{m-1} x_j e^{-2\pi i j k/m}.$$

We may also write $\{\hat{x}_j\} = \mathcal{F}_m(\{x_j\})$.

By summing the geometric series, we can obtain

$$\sum_{j=0}^{m-1} \exp\{\tfrac{2\pi i j}{m}[k - \ell]) = m\delta_{k\ell},$$

which implies that the set

$$\{(1, \exp\{\tfrac{2\pi i k}{m}\}, \exp\{\tfrac{4\pi i k}{m}\}, \ldots, \exp\{\tfrac{2\pi i (m-1) k}{m}\}) : k = 0, \ldots, m-1\}$$

forms an orthogonal basis for $\mathbb{C}^m$. These vectors consist of samples of the functions $e^{2\pi i k x}$ at the points $\frac{j}{m}$, where $j = 0, \ldots, m-1$.

The identity above also implies that the inverse of the finite Fourier transform is given by

$$x_j = \sum_{k=0}^{m-1} \hat{x}_k e^{2\pi i j k/m}.$$

Just as we extended functions periodically, we extend finite sequences periodically via

$$x_{j+m} = x_j \quad \text{and} \quad \hat{x}_{k+m} = \hat{x}_k.$$

Remark 4. *When $m \in 2^{\mathbb{N}}$, the* fast Fourier transform (FFT) *provides a way to compute the Fourier transform (or inverse Fourier transform) using $\mathcal{O}(m \log m)$ computations, compared to the $\mathcal{O}(m^2)$ computations usually needed to multiply an $m \times m$ matrix and a vector in $\mathbb{C}^m$.*

Now suppose f is an L-periodic function with Fourier coefficients $\{\hat{f}(n)\}$. We say f is *N-bandlimited* if $\hat{f}(n)$ for all $|n| \geq N$, in which case

$$f(x) = \tfrac{1}{L} \sum_{|n|<N} \hat{f}(n) e^{2\pi i n x/L}.$$

Lemma 1. *If f is L-periodic and N-bandlimited, then the Fourier coefficients of f correspond to the finite Fourier transform of the sequence of samples $f(\frac{jL}{2N-1})$. In particular, if*

$$x_j = f(\tfrac{jL}{2N-1}), \quad j = 0, \ldots, 2N-2,$$

then

$$\mathcal{F}_{2N-1}(\{x_j\}) = \tfrac{1}{L}\{\hat{f}(0), \ldots, \hat{f}(N-1), \hat{f}(1-N), \ldots, \hat{f}(-1)\}.$$

Proof. This proof may be skipped in lecture.

Extend the finite Fourier transform $\hat{x}_k$ periodically as above. By direct computation, for $|k| \leq N-1$,

$$\begin{aligned}
(2N-1)\hat{x}_k &= \sum_{j=0}^{2N-2} x_j \exp\{-\tfrac{2\pi i k j}{2N-1}\} \\
&= \sum_{j=0}^{2N-2} f(\tfrac{jL}{2N-1}) \exp\{-\tfrac{2\pi i j k}{2N-1}\} \\
&= \tfrac{1}{L} \sum_{j=0}^{2N-2} \sum_{|n|<N} \hat{f}(n) \exp\{\tfrac{2\pi i n j L}{(2N-1)L} - \tfrac{2\pi i j k}{2N-1}\} \\
&= \tfrac{1}{L} \sum_{|n|<N} \hat{f}(n) \sum_{j=0}^{2N-2} \exp\{\tfrac{2\pi i j}{2N-1}[n-k]\} = \tfrac{2N-1}{L} \hat{f}(k).
\end{aligned}$$

This implies the result. □

Using the inversion of the finite Fourier transform, we can obtain a Nyquist theorem for bandlimited periodic functions.

Theorem 4 (Nyquist for periodic functions)**.** *If f is L-periodic and N-bandlimited, then f can be reconstructed from the samples*

$$\{f(\tfrac{jL}{2N-1}) : j = 0, \ldots, 2N-2\}.$$

In particular,

$$f(x) = \tfrac{1}{2N-1} \sum_{j=0}^{2N-2} f(\tfrac{jL}{2N-1}) \tfrac{\sin((2N-1)y_j(x))}{\sin(y_j(x))}, \quad \textit{where} \quad y_j(x) = \pi[\tfrac{x}{L} - \tfrac{j}{2N-1}].$$

Proof. This proof may be skipped in lecture.

Using the Fourier inversion formula and the preceding result:

$$\begin{aligned} f(x) &= \tfrac{1}{L} \sum_{|n|<N} \hat{f}(n) e^{2\pi i n x/L} \\ &= \tfrac{1}{2N-1} \sum_{j=0}^{2N-2} \sum_{|n|<N} f(\tfrac{jL}{2N-1}) \exp\{-\tfrac{2\pi i n j}{2N-1} + \tfrac{2\pi i n x}{L}\} \\ &= \tfrac{1}{2N-1} \sum_{j=0}^{2N-2} f(\tfrac{jL}{2N-1}) \tfrac{\sin((2N-1)y)}{\sin y}, \quad \text{where} \quad y = \pi[\tfrac{x}{L} - \tfrac{j}{2N-1}]. \end{aligned}$$

□

Similar to the continuous case, the function given in the Nyquist theorem always produces an N-bandlimited function

$$F_N(x) = \tfrac{1}{2N-1} \sum_{j=0}^{2N-2} f(\tfrac{jL}{2N-1}) \tfrac{\sin((2N-1)y_j(x))}{\sin(y_j(x))}$$

that agrees with f at the sample points $\{\frac{jL}{2N-1}\}$, where $j = 0, \dots, 2N-2$. The Fourier coefficients of F_N are given by

$$\hat{F}_N(k) = \hat{f}(k) + \sum_{n \neq 0} \hat{f}(k + n(2N-1)),$$

and so we again see the aliasing effect if f is itself not N-bandlimited.

♣ ♣ ♣

We turn to a discussion of sampling in higher dimensions. In this setting, the notion of bandlimited data refers to having bounded Fourier support. For $B \in \mathbb{R}^n_+$, we say $f : \mathbb{R}^n \to \mathbb{C}$ is B-bandlimited if

$$\hat{f}(\xi_1, \dots, \xi_n) = 0 \quad \text{if} \quad |\xi_j| > B_j \quad \text{for} \quad j = 1, \dots, n.$$

For $R > 0$, we may also define $f : \mathbb{R}^n \to \mathbb{C}$ to be R-bandlimited if

$$\hat{f}(\xi) = 0 \quad \text{for} \quad |\xi| > R.$$

Theorem 5 (Nyquist in higher dimensions)**.** *Suppose $B \in \mathbb{R}^n_+$ and $f \in L^2$ is B-bandlimited. Then f can be reconstructed from the samples*

$$\big\{ f(\tfrac{j_1 \pi}{B_1}, \dots, \tfrac{j_n \pi}{B_n}) : (j_1, \dots, j_n) \in \mathbb{Z}^n \big\}.$$

Remark 5. *To apply this to an R-bandlimited function would lead to oversampling, as $\hat{f}$ vanishes in a large subset of $[-R, R]^n$.*

The Poisson summation formula also extends to higher dimensions via

$$\sum_{j \in \mathbb{Z}^n} f(x + j) = \sum_{k \in \mathbb{Z}^n} \hat{f}(2\pi i k) e^{2\pi i \langle x, k \rangle}.$$

In general, many general results and computational methods related to sampling theory in higher dimensions rely on the assumption of uniformly spaced sample points, say $\{x_j\}$. In practice, we may only be able to sample on some points $\{y_k\}$ that do not belong to a uniform grid. To arrive at an approximation to samples of a function f on a uniform grid, one may try to interpolate the values. For

example, suppose that x_j is a point on the uniform grid and $\{y_{k_1}, \ldots, y_{k_\ell}\}$ are the non-uniform sample points closest to x_j. If we can write

$$x_j = \sum_{i=1}^{\ell} \lambda_i y_{k_i}$$

for some $\lambda_i \in (0, 1)$, then we may approximate

$$f(x_j) \approx \sum_{i=1}^{\ell} \lambda_i f(y_{k_i}).$$

If the function f is reasonably smooth, this may work reasonably well. In later sections, we discuss other approaches to interpolation, as well as computational schemes adapted to nonuniform samples.

To conclude, let us point out one other important topic related to sampling and computation, namely, quantization error. This is related to the inherently finite representation of numbers on the computer and can lead to truncation/rounding errors. We will not discuss this topic in these notes.

Theory of filters

In this section we will discuss the theory of filters in general. We focus largely on shift-invariant linear filters, which are widely used in applications. Much of what we discuss will be restricted to one-dimension, although there are natural extensions to higher dimensions.

The word filter is basically synonymous with function, mapping, operator, transformation, and so on. Whenever some input is mapped to some output, we can say we have applied a filter to the input. Our focus will be on *linear* filters. We keep in mind two canonical classes of linear filters. At this stage, we will not concern ourselves with specifying the domain of each filter.

- First, we have *multiplication operators*. Given a function ψ, we may define the filter M_ψ on a function $x(t)$ via
$$M_\psi x(t) = \psi(t)x(t).$$
Given two functions ψ_1 and ψ_2, the filters M_ψ *commute*: $M_{\psi_1}M_{\psi_2} = M_{\psi_2}M_{\psi_1}$.
- Next, we have *integral transforms*. Given a function $a = a(t,s)$, we may define the filter A by
$$Ax(t) = \int a(t,s)x(s)\,ds,$$
which may only be well-defined under suitable assumptions on a and x. If $a(t,s) = \psi(t-s)$ for some function ψ, then we call A a *convolution operator* and write $Ax = \psi * x$.

An important class of filters that we have already encountered is the class of *shift-invariant* integral transforms. Writing $x_\tau(t) = x(t-\tau)$, these filters obey
$$A(x_\tau) = (Ax)_\tau.$$
We have the following fundamental result concerning such filters.

Proposition 1. *An integral transform is shift-invariant if and only if is a convolution operator, which holds if and only if it is a multiplication operator in the Fourier representation.*

Proof. The equivalence between convolution operators and Fourier multiplier operators follows from the identity
$$\mathcal{F}[f * g] = \hat{f}\,\hat{g}.$$
It is also straightforward to check that convolution operators are shift invariant. On the other hand, if an integral transform is shift-invariant, then combining the identities
$$(Ax)_\tau(t) = \int a(t-\tau,\sigma)x(\sigma)\,d\sigma$$
and
$$Ax_\tau = \int a(t,\sigma+\tau)x(\sigma)\,d\sigma$$
for arbitrary inputs x shows that
$$a(t,\sigma+\tau) = a(t-\tau,\sigma) \quad \text{for all} \quad t,\sigma,\tau.$$
Evaluating at $\sigma = 0$ shows $a(t,\tau) = a(t-\tau,0)$, and the result follows by defining the convolution kernel ψ via $\psi(t) = a(t,0)$. □

Remark 1. *We can now observe two key properties of linear shift-invariant filters. First, because of their representation as Fourier multiplier operators, they can be implemented efficiently. Second, all linear shift-invariant filters commute, as can be seen either from the convolution formulation or the Fourier multiplier formulation.*

We next recall the Dirac δ distribution $\delta f = f(0)$, which (under a suitable interpretation of the convolution) satisfies $f * \delta \equiv f$.

Definition 1. *Given a linear, shift-invariant filter A, we call the underlying convolution kernel ψ the* impulse response *of A. This refers to the fact that*

$$\psi(t) = \Psi * \delta(t) = A\delta(t).$$

We may also call ψ the point spread function *or* PSF *of A.*

If we are studying an unknown physical system that we believe to be linear and shift-invariant, then we may derive the underlying convolution kernel by measuring the response to an approximation to the δ distribution (e.g. some localized energetic input).

Any linear, shift-invariant filter is also a Fourier multiplier operator, as

$$\mathcal{F}[Ax] = \mathcal{F}[\psi * x] = \hat{\psi}\hat{x}.$$

Definition 2. *Let A be a linear, shift-invariant filter with impulse response given by ψ. Then $\hat{\psi}$ is called the* transfer function *of A (or* modulation transfer function [MTF]*).*

Linear, shift-invariant filters are therefore completely determined by their impulse response or transfer function.

We may formally understand the transfer function as describing the effect of the filter on a pure oscillatory state, that is,

$$A[e^{it\xi}] = \int e^{i(t-s)\xi}\psi(s)\,ds = \hat{\psi}(\xi)e^{it\xi}.$$

For a general input $x(t)$, we often express the Fourier transform $\hat{x}(\xi)$ in polar coordinates:

$$\hat{x}(\xi) = |\hat{x}(\xi)|e^{i\phi(\xi)},$$

where $|\hat{x}(\xi)|$ is the *amplitude* at frequency $\frac{\xi}{2\pi}$ and $\phi(\xi) \in \mathbb{R}$ is the *phase*. We call these quantities the *harmonic components* of the input x.

Example 1. *Suppose $x(t) = \cos(\omega t)$ for $t \in [t_1, t_2]$. We would informally say that x has frequency $\frac{\omega}{2\pi}$ on the interval. In fact, for the idealized signal $x_i(t) = \cos(\omega t)$ for all $t \in \mathbb{R}$, one has*

$$\hat{x}(t) = \pi[\delta(\xi - \omega) + \delta(\xi + \omega)].$$

(To make this precise, we need to use the extension of the Fourier transform to distributions.) A more realistic model is

$$x_r(t) = \psi(t)\cos(\omega t),$$

where ψ is a smooth cutoff to $[t_1, t_2]$. Then

$$\hat{x}_r(t) = \tfrac{1}{2}[\hat{\psi}(\xi - \omega) + \hat{\psi}(\xi + \omega)].$$

If $\psi = 1$ over and decays smoothly to zero, then $|\hat{\psi}|$ will be sharply peaked at $\xi = 0$, so that $|\hat{x}_r|$ describes which frequencies are present are in the signal. If ψ is shifted by some t_0, then its Fourier transform is multiplied by $e^{it_0\xi}$; accordingly, the phase indicates where the signal is nonzero.

If we write the transfer function $\hat{\psi}$ for a linear, shift-invariant filter in its harmonic components, i.e.

$$\hat{\psi}(\xi) = a(\xi)e^{i\theta(\xi)},$$

then we have

$$\mathcal{F}(\psi * x) = a(\xi)|\hat{x}(\xi)|e^{i[\theta(\xi)+\phi(\xi)]},$$

where $\hat{x}$ is as above. This gives the frequency-space description of the filter, namely, the amplification or attenuation of each coefficient and corresponding phase shift.

Before turning to some concrete examples, let us introduce a few important types of filters.

- Given a collection $\{A_1, \dots, A_k\}$ of filters, we may form the *cascade of filters* by the composition

  $$x \mapsto A_k \circ \cdots \circ A_1 x.$$

  In particular, more complicated filters may be built out of simpler ones. In general, the order matters. In theory, this is not the case for linear, shift-invariant filters; in practice, however, the order in which filters are computed numerically can make a big difference.
- A one-dimensional integral filter with kernel $a(t, s)$ is *causal* if $a(t, s) = 0$ for $s > t$. We may write

  $$Ax(t) = \int_{-\infty}^{t} a(t, s)x(s)\, ds.$$

  For linear, shift-invariant filters, this corresponds to saying that the convolution kernel ψ vanishes for $t < 0$.
- A filter is a *bandpass filter* if it restricts the range of the Fourier transform (to some *passband*).
- A filter is *isotropic* if it commutes with rotations. A linear, shift-invariant filter is isotropic if and only if its impulse response or transfer function is radial.

Let us consider several examples to illustrate the ideas introduced above.

Example 2 (Identity)**.** *Let $Af = f$ (the identity map). This is a linear, shift-invariant filter. Its impulse response is the δ distribution, and its transfer function is $\hat{\delta} \equiv 1$.*

Example 3 (Differentiation)**.** *Let $Af = f'$ (the derivative of f). This is a linear, shift-invariant filter. Its impulse response should be δ', whatever that means. In fact, δ' is a well-defined distribution, given by*

$$\delta' f = -f'(0).$$

(This definition is obtained by performing a formal integration by parts in the integral $\int \delta' f$.) The transfer function is once again a bit more mundane, given by $\mathcal{F}[\delta'](\xi) = i\xi$.

Example 4 (Ideal bandpass filters)**.** *For $\alpha < \beta$, we define the bandpass filter $B_{[\alpha,\beta]}$ via*

$$B_{[\alpha,\beta]}x = \mathcal{F}^{-1}[\chi_{[\alpha,\beta]}(|\xi|)\hat{x}(\xi)].$$

This is a linear, shift-invariant filter that is defined to have transfer function equal to $\chi_{[\alpha,\beta]}(|\xi|)$. *The impulse response function is*

$$b_{[\alpha,\beta]}(t) = 2\,\mathrm{Re}\big[e^{-\frac{it(\alpha+\beta)}{2}}\,\tfrac{\sin(t(\beta-\alpha)/2)}{\pi t}\big].$$

We call $[\alpha,\beta]$ *the* passband. *If* $\alpha = 0$, *the filter is called an* ideal lowpass filter, *while an* ideal highpass filter *has transfer function* $1-\chi_{[0,\beta]}(|\xi|)$.

An ideal lowpass filter coincides with a partial inverse, with point spread function given by

$$h_\beta(t) = \frac{\beta\,\mathrm{sinc}(\beta t)}{\pi}.$$

Because such functions fail to be absolutely integrable, the filtered outputs will be susceptible to Gibbs oscillations.

Instead of exact lowpass filters, one can use approximate lowpass filters (known as *apodizing filters*), as in the following two examples.

Example 5 (Tent function)**.** *Given* $\beta > 0$, *define the transfer function*

$$\hat{t}_\beta(\xi) = \tfrac{1}{\beta}\chi_{[-\frac{\beta}{2},\frac{\beta}{2}]} * \chi_{[-\frac{\beta}{2},\frac{\beta}{2}]}(\xi),$$

which obeys $\hat{t}_\beta(0) = 1$ *and* $\hat{t}_\beta(\xi) = 0$ *for* $|\xi| \geq \beta$. *The corresponding point spread function is*

$$t_\beta(s) = \tfrac{1}{\beta}\big[\tfrac{\sin(\beta s/2)}{\pi s}\big]^2.$$

Example 6 (Hanning window)**.** *For* $\beta > 0$, *we may define the transfer function*

$$\hat{h}_\beta(\xi) = \begin{cases} \cos^2(\frac{\pi\xi}{2\beta}) & |\xi| < \beta \\ 0 & |\xi| > \beta, \end{cases}$$

which is even smoother than the tent function. The point spread function is given by

$$h_\beta(t) = \tfrac{\pi}{2\beta^2}\,\tfrac{\sin(t\beta)}{t[(\pi/\beta)^2 - t^2]}.$$

Example 7 (Hilbert transform)**.** *The Hilbert transform* H *is a linear, shift-invariant filter with transfer function* $\mathrm{sign}(\xi)$. *The impulse response function is the distribution* $\mathrm{PV}[\frac{i}{\pi t}]$.

Example 8 (Ramp filter)**.** *The ramp filter* $|\nabla|$ *is defined via the transfer function* $|\xi|$. *It can be written as* $\frac{1}{i}H\partial_x$, *where* H *is the Hilbert transform. More general* fractional derivatives $|\nabla|^s$ *can be defined via the transfer function* $|\xi|^s$.

Example 9 (Commutativity)**.** *Suppose* $A_1 x = \partial_t x$ *and* $A_2 x = \varphi * x$ *for some differentiable* φ *of bounded support. In theory,* $A_1 \circ A_2 = A_2 \circ A_1$. *In practice, however, it is much simpler to compute* $A_1 \circ A_2$ *than* $A_2 \circ A_1$. *Indeed,*

$$A_1 \circ A_2 x = \partial_t[\varphi * x] = [\partial_t \varphi] * x,$$

whereas computing $A_2 \circ A_1$ *would require that we numerically approximate the derivative of an arbitrary input before convolving with* φ.

In many instances, it is natural to try to invert a given filter. As a simple example, suppose our model for measurement of a signal x is given by a convolution $Ax = k * x$. To recover x, we would then therefore need to invert this filter.

Of course, this is not always possible. Consider, for example, a bandpass filter, which explicitly throws away part of the function.

Nonetheless, as a starting point for this problem, we observe the formal identity

$$x = \mathcal{F}^{-1}\left[\frac{\mathcal{F}[k * x]}{\hat{k}}\right],$$

where $k * x$ represents the measured data. In some cases, $\hat{k}$ may vanish, and so this formula may not be useful. In general, if $k \in L^1$, then we know that $\hat{f}(k) \to 0$ as $|\xi| \to \infty$. Therefore the inversion process (i.e. the filter with transfer function given by $1/\hat{k}$) will amplify high frequencies. In realistic settings, the measurement may be modeled by $k * x + n$, where n represents some noise. We then have

$$\frac{\mathcal{F}[k * x + n]}{\hat{k}} = \hat{x} + \frac{\hat{n}}{\hat{k}},$$

which shows that the high frequencies in the noise will be amplified by this approach to inversion.

A modification to this approach involves truncating the support of the transfer function and using the following approximate inverse:

$$y \mapsto \mathcal{F}^{-1}\left[\frac{\chi_{[-a,a]}(\xi)}{\hat{k}(\xi)} y\right] \quad \text{for some} \quad a > 0,$$

which would give perfect reconstruction for suitably bandlimited data. While real signals will not typically be bandlimited, they may be '*effectively bandlimited*' in the sense that most of the content of the signal is contained in a bounded interval of frequencies.

This modification still does not address the possibility that $\hat{k}$ could vanish at some finite frequencies. In this setting, we can 'repair' the transfer function near its zeros. That is, near a zero ξ_0 of $\hat{k}$, we could replace $\hat{k}(\xi)$ with $\pm\varepsilon$ in a neighborhood of size ε.

In practice, designing inverse filters is really a problem in engineering. We will conclude our discussion by considering just one example.

Example 10. *Suppose x is L-bandlimited. By Nyquist's theorem, we should sample x at equally spaced points with a mesh size at most $\frac{\pi}{L}$.*

To model the measurement of x, fix $\varepsilon > 0$ and define

$$h_\varepsilon(t) = \tfrac{1}{2\varepsilon}\chi_{[-1,1]}(\tfrac{t}{\varepsilon}).$$

Observe that $h_\varepsilon * x$ is still L-bandlimited.

Samples of $h_\varepsilon * x$ at $\{\frac{n\pi}{L}\}_{n\in\mathbb{Z}}$ allow for recovery of $\mathcal{F}[h_\varepsilon * x]$ via

$$\mathcal{F}[h_\varepsilon * x](\xi) = \chi_{[-L,L]}(\xi) \sum_{n\in\mathbb{Z}} h_\varepsilon * x(\tfrac{n\pi}{L}) e^{-in\pi\xi/L}.$$

Now, $\hat{h}_\varepsilon(\xi) = \operatorname{sinc}(\varepsilon\xi/2)$ has its first zero at $2\pi/\varepsilon$. If ε is sufficiently small (so that $\frac{\varepsilon}{2} < \frac{\pi}{L}$), then we may safely divide $\mathcal{F}[h_\varepsilon * x]$ by $\hat{h}_\varepsilon$ and hence exactly reconstruct $\hat{x}$.

In fact, if ε is small enough (e.g. $\varepsilon = \frac{\pi}{L}$) then the consecutive sampling intervals do not overlap and we may obtain a uniform lower bound on $\hat{h}_\varepsilon$ over $[-L, L]$.

This example shows that a bandlimited signal can be exactly and stably reconstructed from fairly realistic measurements, provided the 'resolution' of the measuring device is high enough (relative to the bandwidth of the signal).

The preceding example leads us naturally to the notion of *resolution*. In fact, resolution is not exactly a well-defined mathematical concept and is mostly useful for purposes of comparison of different filters.

For a filter A, the resolution R_A in the output of A is given as a length, which may be interpreted as the size of the smallest discernible feature in the output; the minimum separation between discernible features; or the extent to which a pointlike object is spread out. The resolution increases as R_A decreases.

We will focus on giving just a few possible definitions for resolution in the output of a linear, shift-invariant filter $Ax = k * x$.

- Assume k attains its maximum M at zero. For $\kappa \in (0, 1)$, let
$$\Delta(A, \kappa) = t_+(\kappa) - t_-(\kappa),$$
where $t_\pm(\kappa)$ denote the largest negative and smallest positive values where $k(t) = \kappa M$. This is called the *full-width κ-maximum* of A. If k has zeros, then the special case of $\kappa = 0$ just gives the largest negative and smallest positive zeros. A special case is given by $\kappa = \frac{1}{2}$, which corresponds to the FWHM (full width half maximum).
- Suppose $|\hat{k}|$ attains it maximum at $\xi = 0$. For $\varepsilon \in (0, 1)$, we let $\xi_\pm$ denote the largest negative value and smallest positive values of ξ such that $|\hat{k}(\xi)| \leq \varepsilon|\hat{k}(0)|$. Then the *$\varepsilon$-Nyquist width* is defined by
$$\Delta(A, \mathrm{Ny}, \varepsilon) = \frac{\pi}{\min\{|\xi_+|, |\xi_-|\}}.$$
This is connected to the fact that a broader transfer function (corresponding to a smaller Δ) corresponds to keeping a larger range of frequencies, and hence retaining a higher degree of resolution in the output of the filter.

The notions of resolution introduced above carry over naturally to higher-dimensional filters, provided the filters are isotropic. Otherwise, the situation becomes more complicated.

The resolution generally decreases as we apply cascades of filters; however, the precise statements depend on (i) the types of filters appearing in the cascades and (ii) the precise notion of resolution being considered. We will not pursue this topic in detail here.

In the final topic of this section, we will consider the filtering of periodic inputs (or periodic extensions of inputs defined on intervals).

Viewing inputs as defined on $\mathbb{R}$, we say $x(t)$ is *L-periodic* if $x(t + L) = x(t)$ for all $t \in \mathbb{R}$. A filter A is *L-periodic* if it maps L-periodic functions to L-periodic functions.

The L-periodic unit impulse δ_L is given formally by the sum
$$\delta_L(t) = \sum_{j \in \mathbb{Z}} \delta(t + jL).$$

In this setting, an L-periodic, linear, shift-invariant filter is given by convolution with the *impulse response* $k(t)$, defined by

$$k(t) = A(\delta_L).$$

The *transfer function* in this case becomes the sequence $\{\hat{k}(n)\}_{n\in\mathbb{Z}}$, obtained by applying the filter to the complex exponentials $e^{2\pi int/L}$:

$$A(e^{2\pi int/L}) = \hat{k}(n)e^{2\pi int/L}.$$

As the Fourier coefficients of $k * x$ are given by $\hat{k}(n)\hat{x}(n)$, we have the following Fourier representation:

$$Ax(t) = \tfrac{1}{L}\sum_{n\in\mathbb{Z}} \hat{k}(n)\hat{x}(n)e^{2\pi int/L}.$$

Example 11. *The bandpass filter with passband* $[a, b]$ *has transfer function*

$$\hat{k}(n) = \begin{cases} 1 & |n| \in [a, b] \\ 0 & \textit{otherwise.} \end{cases}$$

Various notions of resolution carry over to this setting as well.

Implementation of filters

In this section, we discuss (briefly) the actual 'implementation' of shift-invariant filters. This refers to the transition from the continuous filters introduced in the previous section to their approximate realizations on finitely sampled data. The main tools needed are the finite Fourier transform and Riemann sums.

We suppose x is some input to a linear, shift-invariant filter $\mathcal{H}$ with impulse response h and transfer function $\hat{h}$. Fixing some $t_0 \in \mathbb{R}$, $\tau > 0$, and $N \in \mathbb{N}$, our goal is to derive practical approximations to the values

$$h * x(t_0 + k\tau), \quad k = 0, \dots, N-1$$

using finitely many samples of x and h (or $\hat{h}$). We will assume uniform sampling of x and h. In situations where this is not feasible, one may interpolate the available data to produce equally spaced samples. To simplify matters, we will fix $t_0 = 0$ and consider the case $\tau = \frac{1}{N}$, which corresponds to the sample points $\{k/N\}$ for a signal supported on $[0, 1]$.

We begin with the Riemann sum approximation

$$h * x(\tfrac{k}{N}) = \int_0^1 h(\tfrac{k}{N} - y)x(y)\,dy \approx \tfrac{1}{N} \sum_{j=0}^{N-1} h(\tfrac{k-j}{N})x(\tfrac{j}{N}). \tag{1}$$

This approximation requires that we sample the impulse response h at the $2N-1$ points

$$-\tfrac{N-1}{N}, \dots, -\tfrac{1}{N}, 0, \tfrac{1}{N}, \dots, \tfrac{N-1}{N}.$$

We therefore define the *sample sequences* $h_s \in \mathbb{R}^{2N-1}$ and $x_s \in \mathbb{R}^{2N-1}$ as follows. First, we set

$$h_s = \{h(0), \dots, h(\tfrac{N-1}{N}), h(-\tfrac{N-1}{N}), \dots, h(-\tfrac{1}{N})\}.$$

Next, we apply *zero padding* and define x_s as follows:

$$x_s = \{x(0), \dots, x(\tfrac{N-1}{N}), 0, \dots, 0\}.$$

We extend both h_s and x_s as $(2N-1)$-periodic sequences.

Continuing from (1), we therefore have the initial approximation

$$h * x(\tfrac{k}{N}) \approx \tfrac{1}{N} \sum_{j=0}^{N-1} h_s(k-j)x_s(j) = \tfrac{1}{N} h_s \star x_s(k),$$

where $\star$ denotes the *discrete periodic convolution*.

To compute this discrete convolution in practice, we may use the following convolution identity for the finite Fourier transform:

Lemma 1. *Let x, y be sequences of length M. Then*

$$\mathcal{F}_M(x \star y) = M\{\hat{x}_0\hat{y}_0, \dots, \hat{x}_{M-1}\hat{y}_{M-1}\},$$

where $\mathcal{F}_M$ is the finite Fourier transform, $\{\hat{x}_k\} = \mathcal{F}_M(x)$, and $\{\hat{y}_k\} = \mathcal{F}_M(y)$.

Proof. The proof may be skipped in lecture. We have

$$\mathcal{F}_M(x \star y)_k = \tfrac{1}{M} \sum_{\ell=0}^{M-1} \left[\sum_{j=0}^{M-1} x_\ell y_{j-\ell} \right] e^{-2\pi i jk/M}.$$

Setting $m = j - \ell$ and using periodicity then yields the result. □

Thus, continuing from above, we arrive at the following approximation:

$$h * x(\tfrac{k}{N}) \approx \tfrac{2N-1}{N}\mathcal{F}_{2N-1}^{-1}\big(\{\hat{x}_s(0)\hat{h}_s(0),\dots,\hat{x}_s(2N-1)\hat{h}_s(2N-1)\}\big)(k), \quad (2)$$

where $\hat{x}_s = \mathcal{F}_{2N-1}x_s$ and $\hat{h}_s = \mathcal{F}_{2N-1}h_s$ are the finite Fourier transforms of the sample sequences. This is a useful formulation in general due to the efficient numerical implementation of the finite Fourier transform and its inverse.

In the remainder of this section, let us discuss the following questions:

(i) How is $\hat{x}_s$ related to the Fourier transform or Fourier series of the original signal x?

(ii) Can we approximate $\hat{h}_s$ without first having to sample the impulse function? Could we instead sample the transfer function directly?

For (i), we can observe that

$$\begin{aligned}
\hat{x}_s(k) &= \tfrac{1}{2N-1}\sum_{j=0}^{2N-2} x_s(j)e^{-2\pi ijk/(2N-1)} \\
&= \tfrac{1}{2N-1}\sum_{j=0}^{N-1} x_s(j)e^{-2\pi ijk/(2N-1)} \\
&= \tfrac{N}{2N-1}\sum_{j=0}^{N-1} \tfrac{1}{N}x_s(j)\exp\{-2\pi i\tfrac{j}{N}\tfrac{kN}{2N-1}\} \\
&\approx \tfrac{N}{2N-1}\int_0^1 x(t)e^{-2\pi it(\frac{kN}{2N-1})}\,dt = \tfrac{N}{2N-1}\hat{x}(\tfrac{2\pi Nk}{2N-1}),
\end{aligned}$$

showing that $\hat{x}_s$ is indeed related to samples of the Fourier transform of $\hat{x}$.

We remark that the finite Fourier transform of the unpadded sample sequence can also be connected rather directly to the Fourier coefficients of x (thought of as a function on $[0,1]$).

For (ii), we note that there is another approach to approximating the convolution, namely, via the Fourier inversion formula:

$$h * x(\tfrac{k}{N}) \approx \tfrac{1}{2\pi}\int_{-N\pi}^{N\pi}\hat{h}(\xi)\hat{x}(\xi)e^{i\xi k/N}\,d\xi.$$

The truncation to frequencies $|\xi| \le N\pi$ is consistent with the assumption that a sample spacing of $\frac{1}{N}$ is sufficient for approximate the given signal. Continuing, we make a Riemann sum approximation and arrive at

$$h * x(\tfrac{k}{N}) \approx \tfrac{1}{2\pi}\sum_{j=-(N-1)}^{N-1}\hat{h}(\tfrac{2\pi Nj}{2N-1})\hat{x}(\tfrac{2\pi Nj}{2N-1})\exp\{\tfrac{2\pi ijk}{2N-1}\}\tfrac{2\pi N}{2N-1}.$$

If we make the approximations

$$\hat{x}_s(k) \approx \tfrac{N}{2N-1}\hat{x}(\tfrac{2\pi Nk}{2N-1}) \quad \text{and} \quad \hat{h}_s(k) \approx \tfrac{N}{2N-1}\hat{h}(\tfrac{2\pi Nk}{2N-1}), \quad (3)$$

then we arrive at the same approximation to $h * x(\frac{k}{N})$ as the one appearing in (2). This derivation shows that it is possible to build an approximation using only samples of the transfer function $\hat{h}$, rather than the impulse response h.

We summarize our discussion as follows: The formula in (2) provides an approximation to the convolution $h*x$ at the sample points $\frac{k}{N}$ using finitely many samples. In particular, we sample the signal at finitely many points, pad by zeros, and take

the finite Fourier transform to form the sequence $\hat{x}_s(k)$. For approximating $\hat{h}_s(k)$, there are two basic approaches:

- First, we may sample the impulse response function h and then obtain $\hat{h}_s$ as the finite Fourier transform of the sample sequence h_s.
- Second, we may use the approximation in (3) to obtain an approximation to the values of $\hat{h}_s(k)$.

In general, these two approaches will perform differently, and which choice is better depends on the particular situation at hand.

We will leave it as an exercise to extend the ideas above to the more general scenario of sampling at points $\{t_0 + \tau k\}$ for some $t_0 \in \mathbb{R}$ and uniform sampling spacing $\tau > 0$.

X-ray CT reconstruction

In this section, we return to the reconstruction problem in X-ray CT and discuss several approaches.

In this section, we will continue to make various simplifying assumptions:

- First, we continue with assumption of one-dimensional monochromatic X-ray beams, which (as we have previously discussed) is not realistic.
- Next, we assume that measurements are made with infinite precision and can similarly be represented digitally with infinite precision. In truth, the accuracy of measurements depends on the accuracy and sensitivity of the X-ray detectors, the stability of the X-ray source, the correct calibration of the scanning machine, and so on. Similarly, the contrast available in the reconstructed image actually depends on the range of measurements made and the number of bits used to store them.

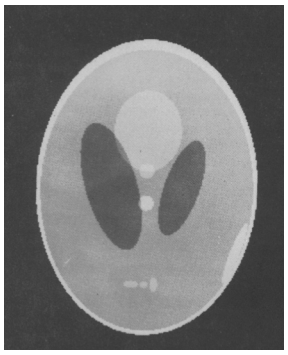
We will revisit some of these assumptions later sections.

The quality of a reconstruction algorithm can be characterized by several key properties:

- The first key property is of course the accuracy of the reconstructed image. In the case of X-ray CT, one benefits greatly from the availability of exact reconstruction formulas on which to base the reconstruction algorithm.
- The next property is stability (e.g. with respect to various types of noise in the measured data). As we have discussed before, this is connected to continuity properties present in the inversion formulas.
- Finally, a good algorithm should be efficient in terms of computational cost. In the present setting, we will see that many computations are parallelizable, thus allowing for their efficient implementation.

To test the accuracy and stability of reconstruction algorithms, one can utilize so-called 'phantoms'. This refers to both physical and mathematical phantoms.

- With a physical phantom, one tests the reconstruction algorithm on an object for which one knows the internal structure. In this case, one mixes the measurement error with algorithmic error.
- The idea of a mathematical phantom was introduced by Shepp. The idea is to isolate algorithmic errors by tests the reconstruction algorithm on a purely mathematical construction. To simulate measurements, one should digitize the image and replace point values by some kind of weighted average. One can use mathematical phantoms to test robustness to measurement error, understand sampling artifacts, and compare different algorithms. An example of a mathematical phantom is pictured below:



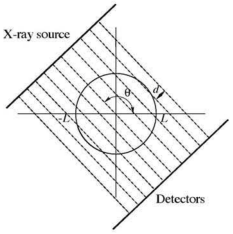
We consider the problem of reconstructing a single slice of a $3d$ object contained in the disc of radius L. We denote the unknown attenuation coefficient by $f = f(x, y)$, which we will attempt to reconstruct on the *reconstruction grid*

$$R_\tau = \{(x_j, y_k) = (j\tau, k\tau)\},$$

where $\tau > 0$ is the sample spacing. Here j, k range between $-K$ and K, where $K = \lfloor \frac{L}{\tau} \rfloor + 1$.

Parallel Beam Scanner

Let us first consider reconstruction using a parallel beam scanner. This type of scanner represents the 'first generation' of X-ray CT scanners. In the parallel beam setup, we place an array of X-ray sources on one side of the object, with a corresponding array of detectors on the other side, as in the following figure from [1]:



(a) A parallel beam scanner.

The sources and detectors are then rotated together around the object and data is collected.

We fix $M > 1$ and choose a finite collection of directions

$$\{\omega(k\Delta\theta)\}_{k=0,\dots,M}, \quad \text{where} \quad \Delta\theta = \tfrac{\pi}{M+1}.$$

For a fixed direction, the X-ray source/detector pairs will yield the Radon transform along the discrete set of affine parameters

$$\{jd\}_{j=-N}^{N},$$

where $d > 0$ is the affine sample spacing and $N = Ld^{-1}$.

The data acquired in the parallel beam setup takes the form

$$\{\mathcal{R}f(jd, \omega(k\Delta\theta)) : j = -N, \dots, N,\ k = 0, \dots M\},$$

which represents a rectangular grid in (t,θ)-space. Note that by the evenness of $\mathcal{R}f$, it is technically sufficient to obtain samples of corresponding to angles in $[0,\pi)$; practically, it can be useful to obtain 'redudant' information and average the results.

We will discuss the connections between the choices of M, N, and K (the number of points in the reconstruction grid) below.

Definition 1. *We call an individual measurement a* ray. *We call the collection of samples of $\mathcal{R}f(t,\omega)$ for a fixed ω a* view.

A plot of $\mathcal{R}f(t,\omega)$ is called a *sinogram*. Such plots are not easy to interpret (for humans, at least).

⋆ ⋆ ⋆ *Direct Fourier Inversion* ⋆ ⋆ ⋆

One approach to reconstruction using parallel beam data is to directly appeal to the Fourier inversion formula. As we will see, this approach has some issues.

Using the convention that

$$\hat{f}(r\omega(\theta)) = \hat{f}(|r|\omega(\theta+\pi)) \quad \text{for} \quad r<0,$$

and supposing we had complete data for each fixed angle, the central slice formula tells us that

$$\hat{f}(r\omega(k\Delta\theta)) = \int_{-\infty}^{\infty} \mathcal{R}f(t,\omega(k\Delta\theta))e^{-irt}\,dt.$$

Then, by the Fourier inversion formula (stated in polar coordinates) and taking a Riemann sum approximation, we obtain

$$\begin{aligned} f(x,y) &= \tfrac{1}{4\pi^2}\int_0^{\pi}\int_{-\infty}^{\infty} \hat{f}(r\omega)e^{ir(x,y)\cdot\omega}|r|\,dr\,d\omega \\ &\approx \tfrac{1}{4\pi(M+1)}\sum_{k=0}^{M}\int_{-\infty}^{\infty} \hat{f}(r\omega(k\Delta\theta))e^{ir(x,y)\cdot\omega(k\Delta\theta)}|r|\,dr. \end{aligned}$$

In practice, we have only the samples $\mathcal{R}f(jd,\omega(k\Delta\theta))$, which would allow us to approximate

$$\hat{f}(r_j\omega(k\Delta\theta)), \quad r_j\in\{0,\pm\eta,\dots,\pm N\eta\}, \quad \eta = \tfrac{1}{N}\tfrac{\pi}{d} = \tfrac{\pi}{L}.$$

We could then take another Riemann sum approximation in the sum above.

Now consider the computational cost of the scheme outlined above. To approximate f at a single point (x,y) requires $\mathcal{O}(MN)$ operations. As there are $\mathcal{O}(K^2)$ points in the reconstruction grid, this approach requires an enormous $\mathcal{O}(MNK^2)$ computations.

An alternate approach would be to use the fast Fourier transform (FFT). As the Radon transform data is sampled uniformly in t for each fixed angle, one could use the FFT to obtain the values $\hat{f}(r_j\omega(k\Delta\theta))$ in $\mathcal{O}(MN\log N)$ operations.

However, we do *not* obtain uniform samples of the Fourier transform, and hence we cannot directly apply the fast inverse Fourier transform (IFFT). Indeed, we instead obtain samples along concentric circles. One possible remedy is to use a nearest neighbor interpolation to obtain a uniform grid of $\mathcal{O}(K^2)$ approximate samples of the Fourier transform. Then, in $\mathcal{O}(K^2[\log K]^2)$ operations, one can evaluate the IFFT. In particular, we can perform the complete inversion in

$$\mathcal{O}(MN\log N + K^2[\log K]^2)$$

operations.

Unfortunately, this does not work well. The reason is that for a typical attenuation coefficient f, the Fourier transform $\hat{f}$ will exhibit rapid oscillation of both phase and magnitude; this can be seen clearly at the level of mathematical phantoms. In particular, the nearest neighbor interpolation introduces too many errors and the reconstruction is not accurate.

⋆ ⋆ ⋆ *Filtered Back-Projection* ⋆ ⋆ ⋆

The starting point for the parallel beam reconstruction algorithm is instead the filtered back-projection formula, which we recall here:

$$f(x) = \tfrac{1}{2\pi} \int_0^\pi |\nabla| \mathcal{R}f(\langle x, \omega \rangle, \omega)\, d\omega.$$

Accordingly, the reconstruction is based on two steps:

1. Implement the $1d$ filter $|\nabla|$ on the Radon transform data.
2. Implement the back-projection.

Note that because the filter acts only on the affine variable, we should be able to carry out the filtering step one view at a time, so that the algorithm can be parallelized.

The back-projection step will be implemented with a simple Riemann sum approximation. The more challenging component of this algorithm involves the implementation of the nonlocal filter $|\nabla|$.

As discussed previously, one approach to approximating $|\nabla|$ is to instead approximate the transfer function, which is given by $|r|$. Let us denote the approximate transfer function by $\hat{\phi}$. Frequently, one writes $\hat{\phi}$ in the form

$$\hat{\phi}(r) = A(r)|r|$$

for some apodizing function $A(r)$. Some standard options for $A(r)$ include a simple bandlimiting filter, the 'Hamming' filter, or the 'Hanning' filter. Once a choice has been made, we may denote the filtered Radon data as

$$Q_\phi f(t, \omega) = \tfrac{1}{2\pi} \int \widetilde{\mathcal{R}f}(r, \omega) \hat{\phi}(r) e^{irt}\, dr = \int \mathcal{R}f(s, \omega) \phi(t - s)\, ds.$$

Assuming we have complete data for $\mathcal{R}f$, our approximation to f would then be

$$f_\phi(x, y) := \tfrac{1}{2\pi} \int_0^\pi Q_\phi f(\langle (x, y), \omega \rangle, \omega)\, d\omega.$$

Using the discretized data and a Riemann sum approximation, our reconstructed function is therefore

$$\tilde{f}_\phi(x_m, y_\ell) = \tfrac{d}{2(M+1)} \sum_{k=0}^{M} \sum_{j=-N}^{N} \mathcal{R}f(jd, \omega(k\Delta\theta)) \phi[\langle (x_m, y_\ell), \omega(k\Delta\theta) \rangle - jd]. \quad (1)$$

We now turn to a crucial observation for the parallel beam reconstruction problem. Because the filter is with respect to the affine variable only, we can carry out the filtering step after each view is collected. This presents an opportunity for parallelizing the algorithm and thereby speeding up reconstruction. However, when we examine formula (1), it seems as though we need to use a *different* filter for each point in the reconstruction grid. That is, we must repeat the filter step K^2 times for each view!

To remedy this, Ramachandran and Lakshminarayanan had the idea to simply fix the values of the filtering function ϕ at $\{jd\}_{j=-N}^{N}$ and then linearly interpolate to obtain approximate values at intermediate points. For the filter itself, Ramachandran and Lakshminarayanan used the simple bandlimiting

$$\phi(jd) = \tfrac{1}{2\pi}\int_{-B}^{B} |r| e^{irjd}\, dr$$

for some $B > 0$. In the medical imaging context, we call this the *Ram-Lak filter*.

Because ϕ can be expected to be a slowly varying real-valued function, this interpolation will introduce far less error than what we considered in the case of direct Fourier inversion. In particular, we can expect some blurring to result, but we do not expect to introduce any oscillatory effects, for example.

The interpolation works as follows: As $N = Ld^{-1}$, we have that for each m, ℓ, k there exists an integer $n_{k\ell m} \in [-N, N]$ such that

$$n_{k\ell m} d < \langle (x_m, y_\ell), \omega(k\Delta\theta)\rangle \leq (n_{k\ell m} + 1)d.$$

Consequently, there exists $\alpha_{k\ell m} \in [0, 1]$ such that

$$\langle (x_m, y_\ell), \omega(k\Delta\theta)\rangle = \alpha_{k\ell m}(n_{k\ell m} + 1)d + (1 - \alpha_{k\ell m})n_{k\ell m} d. \tag{2}$$

As the sampling angles and reconstruction grid are fixed at the beginning, the values $(n_{k\ell m}, \alpha_{k\ell m})$ may be computed and stored in a table.

The reconstruction algorithm is then carried out as follows:

1. For each view (i.e. for each fixed k), we approximate the filtered Radon transform at the via the discrete convolution

$$Q_\phi \tilde{f}(jd, \omega(k\Delta\theta)) = d \sum_{n=-N}^{N} \mathcal{R}f(nd, \omega(k\Delta\theta))\phi((j-n)d)$$

for $j = -N, \ldots, N$. This step can be done after the data for a single view is collected.

2. Back-project $Q_\phi \tilde{f}$ using linear interpolation to produce the approximations

$$\tilde{f}_\phi(x_m, y_\ell) = \tfrac{1}{2M+1} \sum_{k=0}^{M} \big[\alpha_{k\ell m} Q_\phi \tilde{f}((n_{k\ell m} + 1)d, \omega(k\Delta\theta)) + (1 - \alpha_{k\ell m}) Q_\phi \tilde{f}(n_{k\ell m} d, \omega(k\Delta\theta))\big],$$

where $\alpha_{k\ell m}$ and $n_{k\ell m}$ are as in (2). By linearity, this is equivalent to interpolating the values of ϕ itself.

The computation of Q_ϕ as a discrete convolution requires $\mathcal{O}(MN^2)$ operations (although this could be reduced by computing using the FFT), while the second step requires $\mathcal{O}(MK^2)$ operations. Later we will take $K \approx N$, so these two computational costs are of the same order.

Use of the Ram-Lak filter and linear interpolation works 'surprisingly well'. Shepp and Logan carried out an analysis to explain the success of this approach. Essentially, the key to success is to make sure that the transfer function resembles $|r|$ near $r = 0$, while at the same time guaranteeing that the point spread function resembles that of $-i\partial_t \mathcal{H}$ for large t (i.e. $\phi(t) \approx -\frac{1}{\pi t^2}$ as $t \to \infty$).

So far, we have taken our parameters K, M, N for the reconstruction grid, projection angles, and affine parameter completely independently. In practice, these

should be linked by considering principles such as Nyquist's theorem. In the remainder of this section, we will show that a reasonable choice of parameters corresponds to choosing $K \approx M \approx N$ and $\tau \approx d$.

First, since we have $d = LN^{-1}$ (the affine sample spacing), the effective support of $\hat{f}(r\omega)$ should be contained in a disc of radius $\approx d^{-1}$. Now, $\hat{f}$ is sampled at N points along each radial line, giving a sample spacing of $\approx L^{-1}$; here we are just using that $NL^{-1} = d^{-1}$.

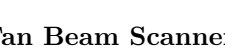
Now consider the widest sample spacing of $\hat{f}$ in the *angular* direction. This occurs at at the largest radius, and hence is given by $\approx \Delta\theta d^{-1}$. It is reasonable to ask that the radial spacing coincide with the worst angular spacing (i.e. $L^{-1} \approx \Delta\theta d^{-1}$), leading to the condition

$$N = Ld^{-1} \approx (\Delta\theta)^{-1} \approx M.$$

Now, since we obtain $\approx N \times N$ samples of the Fourier transform, then we should use the same number of grid points in the reconstruction grid, which is essentially an $L \times L$ square. That is to say, the sample spacing τ should be given by $LN^{-1} = d$. But then

$$K \approx L\tau^{-1} \approx Ld^{-1} = N,$$

and hence we should take $K \approx N$ as well. Another way to arrive at $K \approx N$ is to ask that the total number of measurements ($\approx NM \approx N^2$) agree with the total number of nontrivial reconstruction points ($\approx K^2$).



Fan Beam Scanner

We next consider the *divergent beam* or *fan beam* geometry. Beginning with the second generation of X-ray CT scanners, scanning machines began to utilize a single X-ray source emitting X-rays in a 'fan beam' shape.

- In a 'second-generation' scanner, the X-ray source is across from a flat arrangement of detectors, and the entire configuration is translated and rotated around the patient.

In later designs, the source and detectors were arranged on opposite arcs of a circle and only rotation was used.

- In a 'third-generation' scanner, detectors are arranged in an arc centered on a single X-ray source. The source and detectors are rotated together.
- In a 'fourth-generation' scanner, the detectors are on a fixed ring around the object, and only the X-ray source is rotated around the object.

Our discussion will be limited essentially to third- and fourth-generation scanners. All of the engineering developments related to the detector/source configuration, number of detectors, and so on, have certainly improved the measurement process significantly (e.g. by shortening the scan time). At the level of detail we are currently considering, however, there is not much difference in terms of describing the resulting X-ray data.

In the fan-beam setting, a *view* consists of samples of $\mathcal{R}f$ for a family of lines passing through a point.

- In the third-generation setting, this refers to the family of lines emanating from a fixed X-ray source.

- In the fourth-generation setting, this refers to the family of lines that pass through a fixed detector as the source rotates.

In particular, the two viewpoints are 'dual' to one another and can be treated basically at the same time. It is perhaps conceptually simpler to imagine the lines comprising a single fan beam, which corresponds to third-generation scanners.

In contrast to the parallel-beam setting, data from a single view no longer provides all of the Radon transform data for a fixed ω, and hence we cannot apply results such as the central slice theorem directly. There are then two basic approaches to using the fan beam data:

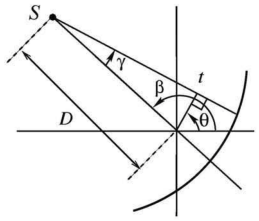
(i) Re-sort the data and interpolate to obtain parallel-beam type information. This is called *rebinning*. We will not discuss this issue here, but only point out that this approach both introduces interpolation error and sacrifices 'parallelizability' of the algorithm.

(ii) Build a reconstruction algorithm that works directly with the fan-beam data. Such an algorithm was first proposed by Herman, Lakshminarayanan, and Naparstek.

We parametrize the fan beam data as follows (see the figure below from [1]).

- We consider an X-ray source (S below) with a *central ray* emanating from S and passing through the origin.
- We write D for the distance from S to the origin along the central ray.
- We write β for the angle formed by this ray and the vertical axis.
- We parametrize other the other lines in the view by the angle γ they form with the central ray.
- Using trigonometry, the line parametrized by γ may be described in the standard (t, θ) coordinates as $\ell_{t,\omega(\theta)}$, where

$$\theta = \gamma + \beta \quad \text{and} \quad t = D \sin \gamma. \tag{3}$$



The variables (β, γ) will describe the family of lines for which we have Radon transform data. The data obtained by the scanner will be uniformly sampled in these coordinates.

Our goal is an analogue of the approximate reconstruction formula

$$f_\phi(x, y) = \tfrac{1}{2\pi} \int_0^\pi \int_{-L}^{L} \mathcal{R}f(t, \theta) \phi(x \cos\theta + y \sin\theta - t)\, dt\, d\omega,$$

where convolution with ϕ is our approximation to the filter $|\nabla|$.

Let us firstly deal with one technical annoyance and re-express this quantity using an integral over $[0, 2\pi]$ rather than $[0, \pi]$. To do this, we observe that the integrand may be written as a function G obeying $G(-t, -\omega) = G(t, \omega)$. We have established this fact previously for the Radon transform, so let us focus on showing that this is true of the filter term. Here the key is to observe that the argument

$|x\cos\theta + y\sin\theta - t|$ is precisely the distance from (x, y) to the line $\ell_{t,\omega(\theta)}$. This follows from the fact that the projection of (x, y) onto $\ell_{t,\omega}$ is

$$(x, y) - \{(x, y) \cdot \omega - t\}\omega.$$

As $\ell_{-t,-\omega} = \ell_{t,\omega}$, we find that this even symmetry holds for the filter term as well, at least provided we impose the condition that ϕ is even. Using this even symmetry, we may now re-write (4) as

$$f_\phi(x, y) = \tfrac{1}{\pi} \int_0^{2\pi} \int_{-L}^{L} \mathcal{R}f(t, \theta)\phi(x\cos\theta + y\sin\theta - t)\, dt\, d\omega, \tag{4}$$

We will focus first on an analogue of this filtered back-projection formula and consider the problem of implementation after arriving at a suitable formula. In what follows, we change notation and denote the convolution kernel of the filter by κ rather than ϕ. We also use ξ for the Fourier variable. This frees up the variables φ and r, which we will use to express (x, y) in polar coordinates via

$$(x, y) = (r\cos\varphi, r\sin\varphi).$$

In particular, (4) becomes

$$f_\kappa(r, \varphi) = \tfrac{1}{\pi} \int_0^{2\pi} \int_{-L}^{L} \mathcal{R}f(t, \theta)\kappa(r\cos(\theta - \varphi) - t)\, dt\, d\theta,$$

where we have used

$$\cos\theta\cos\varphi + \sin\theta\sin\varphi = \cos(\theta - \varphi).$$

We now recall the relations in (3) to change variables in the integral above. We define the fan angle $2\gamma_L$ so that the family of lines corresponding to the parameters

$$\{(\beta, \gamma) : \beta \in [0, 2\pi),\ |\gamma| \le \gamma_L\}$$

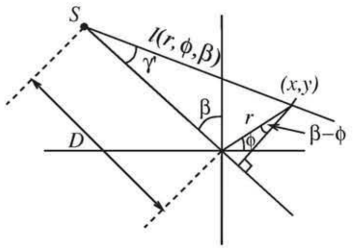
includes all lines intersecting the disc of radius L. We then obtain

$$f_\kappa(r, \varphi)$$
$$= \tfrac{D}{\pi} \int_0^{2\pi} \int_{-\gamma_L}^{\gamma_L} \mathcal{R}f(D\sin\gamma, \beta + \gamma)\kappa[r\cos(\beta + \gamma - \varphi) - D\sin\gamma]\cos\gamma\, d\gamma\, d\beta.$$

Let us introduce the notation

$$Pf(\beta, \gamma) = \mathcal{R}f(D\sin\gamma, \beta + \gamma),$$

which represents the data we will collect in the fan beam setup. We now use some trigonometry to simplify the expression appearing in the argument of κ. For this, it is useful to consider the following figure from [1]:



Lemma 1. *Let*

$$\ell = \ell(r, \varphi, \beta)$$

denote the positive distance from the source S to the point (x, y), and let

$$\gamma' = \gamma'(r, \varphi, \beta)$$

denote the angle formed between the central ray and the ray joining S to (x, y). Then the argument of κ simplifies as follows:

$$r\cos(\beta + \gamma - \varphi) - D\sin\gamma = \ell\sin[\gamma' - \gamma].$$

Proof. We may first rewrite

$$\begin{aligned} &r\cos(\beta + \gamma - \varphi) - D\sin\gamma \\ &\quad = r\cos(\beta - \varphi)\cos\gamma - [D + r\sin(\beta - \varphi)]\sin\gamma. \end{aligned}$$

Now, consulting the figure above and using some trigonometry, we observe

$$\begin{aligned} \ell\cos\gamma' &= D + r\sin(\beta - \varphi), \\ \ell\sin\gamma' &= r\cos(\beta - \varphi). \end{aligned}$$

Thus, continuing from the computation above, we have

$$\begin{aligned} &r\cos(\beta - \varphi)\cos\gamma - [D + r\sin(\beta - \varphi)]\sin\gamma \\ &\quad = \ell\sin\gamma'\cos\gamma - \ell\cos\gamma'\sin\gamma \\ &\quad = \ell\sin(\gamma' - \gamma), \end{aligned}$$

as desired. □

Remark 1. *The proof above also shows that*

$$\begin{aligned} \ell(r, \varphi, \beta) &= \sqrt{[D + r\sin(\beta - \varphi)]^2 + [r\cos(\beta - \varphi)]^2}, \\ \gamma'(r, \varphi, \beta) &= \tan^{-1}\big[\tfrac{r\cos(\beta-\varphi)}{D + r\sin(\beta-\varphi)}\big]. \end{aligned}$$

With our new notation and the lemma in place, we obtain

$$f_\kappa(r, \varphi) = \tfrac{D}{\pi}\int_0^{2\pi}\int_{-\gamma_L}^{\gamma_L} Pf(\beta, \gamma)\kappa[\ell\sin(\gamma' - \gamma)]\cos\gamma\, d\gamma\, d\beta.$$

We turn our attention to describing the filter κ. We state the following informal derivation as a lemma.

Lemma 2. *Suppose $\hat{\kappa}(\xi) \approx |\xi|$ over the effective bandwidth of $\gamma \mapsto Pf(\beta, \gamma)$. The following approximation holds:*

$$f_\kappa(r, \varphi) = \tfrac{D}{\pi}\int_0^{2\pi}\int_{-\gamma_L}^{\gamma_L} Pf(\beta, \gamma)\big[\tfrac{\gamma' - \gamma}{\ell\sin(\gamma' - \gamma)}\big]^2\kappa(\gamma' - \gamma)\cos\gamma\, d\gamma\, d\beta.$$

Proof. Write $\hat{\kappa}(\xi) = |\xi|\chi_\varepsilon(\xi)$, where $\chi_\varepsilon \to 1$ uniformly in ξ as $\varepsilon \to 0$. We consider a general integral of the form

$$\int h(\gamma)\kappa[\ell\sin(\gamma' - \gamma)]\, d\gamma = \tfrac{1}{2\pi}\iint h(\gamma)e^{i\ell\sin(\gamma'-\gamma)\xi}|\xi|\chi_\varepsilon(\xi)\, d\xi\, d\gamma.$$

We now change variables via

$$\eta = \big[\tfrac{\ell\sin(\gamma' - \gamma)}{\gamma' - \gamma}\big]\xi,$$

which yields the integral

$$\tfrac{1}{2\pi}\iint h(\gamma)\big[\tfrac{\gamma'-\gamma}{\ell\sin(\gamma'-\gamma)}\big]^2 e^{i(\gamma'-\gamma)\eta}|\eta|\chi_\varepsilon[\eta\tfrac{\gamma'-\gamma}{\ell\sin(\gamma'-\gamma)}]\,d\eta\,d\gamma.$$

Using

$$\chi_\varepsilon[\eta\tfrac{\gamma'-\gamma}{\ell\sin(\gamma-\gamma)}]\approx\chi_\varepsilon(\eta),$$

we arrive at the approximation

$$\int h(\gamma)\big[\tfrac{\gamma'-\gamma}{\ell\sin(\gamma'-\gamma)}\big]^2\kappa(\gamma'-\gamma)\,d\gamma,$$

which implies the result. □

In light of the previous lemma, we introduce the new filter defined by the convolution kernel

$$g(\gamma):=\tfrac{1}{\pi}[\tfrac{\gamma}{\sin\gamma}]^2\kappa(\gamma),\quad\text{where}\quad\hat\kappa(\xi)\approx|\xi|,$$

and we finally define our approximation

$$f_g(r,\varphi)=\int_0^{2\pi}\frac{1}{\ell^2(r,\varphi,\beta)}\int_{-\gamma_L}^{\gamma_L}Pf(\beta,\gamma)g[\gamma'(r,\varphi\,\beta)-\gamma]D\cos\gamma\,d\gamma\,d\beta.$$

We view this as a *weighted* filtered back projection formula, writing

$$Q_gf(\beta,\gamma')=\int_{-\gamma_L}^{\gamma_L}P'f(\beta,\gamma)g(\gamma'-\gamma)\,d\gamma,\quad P'f(\beta,\gamma)=Pf(\beta,\gamma)\cdot D\cos\gamma,$$

$$f_g(r,\varphi)=\int_0^{2\pi}\frac{1}{\ell^2(r,\varphi,\beta)}\,Q_gf(\beta,\gamma'(r,\varphi,\beta))\,d\beta.$$

This is the analogue of the approximate filtered back-projection formula adapted to the fan beam setting. In the next section, we discuss the implementation of this formula for reconstruction.

⋆ ⋆ ⋆ *Implementation* ⋆ ⋆ ⋆

Fan beam data takes the form $\{Pf(\beta_j,n\alpha)\}$, where $\beta_j=\frac{2\pi j}{M+1}$ and n ranges over some set of integers. Here we recall that

$$Pf(\beta,\gamma)=\mathcal{R}f(D\sin\gamma,\beta+\gamma).$$

We can obtain the data $P'f(\beta_j,n\alpha)$ appearing in our reconstruction formula via

$$P'f(\beta_j,n\alpha)=Pf(\beta_j,n\alpha)\cdot D\cos(n\alpha).$$

This can be done 'one view at a time' (i.e. for a fixed β_j).

The reconstruction algorithm using fan beam data then consists of the following steps.

1. Evaluate the discrete convolution with the kernel g:

$$Q_g\tilde f(\beta_j,n\alpha)=\alpha[P'f(\beta_j,\cdot)\star g](n\alpha),$$

where

$$g(n\alpha)=\tfrac{1}{\pi}\big[\tfrac{n\alpha}{\sin(n\alpha)}\big]^2\kappa(n\alpha).$$

Here $\kappa(\cdot)$ is the convolution kernel approximating the filter $|\nabla|$, chosen similar to the parallel beam case.

2. Compute the weighted back projection of each filtered projection:

$$\tilde{f}_g(x_m, y_\ell) = \tfrac{2\pi}{M+1} \sum_{j=0}^{M} \frac{1}{\ell^2(x_m, y_\ell, \beta_j)} Q_g\tilde{f}(\beta_j, \gamma'(x_m, y_\ell, \beta_j)).$$

   Here the values of $Q_g\tilde{f}$ are obtained by interpolating the values of $Q_g f(\beta_j, n\alpha)$.

The values $\ell(x_m, y_\ell, \beta_j)$ and $\gamma'(x_m, y_\ell, \beta_j)$ may be computed and stored at the beginning. Then the computational cost of this algorithm is similar to that of the parallel beam.

The first step can be carried out as soon as the data from a single view is collected. In third-generation scanners, a view is determined by the source position, so one can begin the filtering step as soon as the data from one position is collected. In a fourth-generation scanner, a view is determined by the detector position, so filtering must wait until all the data for the first view is collected.

It is worth mentioning that in practice, there are some subtleties related to data collection in the fan beam setting. For example, if we take $\beta \in [0, 2\pi)$, then we will actually collect every projection twice! Let us only mention this issue in passing and not discuss it here.

Other Approaches

In the second, third, and fourth generations of scanners, data is collected one slice at a time. In the 1990s, the *spiral scan* or *helical scan* approach was introduced, in which the patient is translated continuously as the source is continuously rotated. Relative to the patient, the X-ray source traces out the shape of a helix. In this way, one obtains X-ray data for the entire volume all at once. Assuming that we are working with 'third-generation' type data, the collected data (in the continuous model) has the form

$$\mathcal{D} = \{P(\beta, \gamma, z(\beta)) : \beta \in [\beta_{\min}, \beta_{\max}], \gamma \in [-\gamma_L, \gamma_L]\},$$

where β now refers to the total accumulated rotation angle of the source and $z(\beta)$ is determined by the translation speed. Typically, one keeps the speed constant, so that $z(\beta) = c\beta + z'$. The data obtained is inherently two-dimensional. One approach to obtaining three-dimensional data is to perform some interpolation to simulate conventional scanner data. Here one finds a tradeoff between inconsistencies in the interpolated data (leading to issues like streaks in the images) and increasing the speed of translation (which would in turn decrease measurement time). Essentially, increasing the translation speed increases the effective 'slice thickness'.

Alternate approaches directly use three-dimensional data for reconstruction. For example, one can use several rows of detector arrays and collect a *two-dimensional* family of line integrals from the source, leading to so-called *cone beam data*. To model this, one can define the *cone beam transform* D mapping functions on $\mathbb{R}^3$ to functions on $\mathbb{R}^3 \times \mathbb{S}^2$ via

$$Df(y, \theta) = \int_0^\infty f(y + t\theta)\, dt, \quad y \in \mathbb{R}^3,\ \theta \in \mathbb{S}^2.$$

The basis for reconstruction using this three-dimensional data is *Grangeat's formula*, which relates the cone beam transform to the *three-dimensional* Radon transform of $f : \mathbb{R}^3 \to \mathbb{R}$, which is defined by

$$\mathcal{R}f(s,\omega) = \int_{x\cdot\omega=s} f(x)\, dA(x), \quad s \in \mathbb{R}, \quad \omega \in \mathbb{S}^2$$

and admits the inversion formula

$$f(x) = -\tfrac{1}{8\pi^2} \int_{\mathbb{S}^2} \partial_s^2 \mathcal{R}f(x\cdot\omega,\omega)\, dA(\omega).$$

The result is the following:

Theorem 1 (Grangeat's formula)**.** *Suppose $f : \mathbb{R}^3 \to \mathbb{R}$ has bounded support. Let $y \in \mathbb{R}^3$ and $\theta \in \mathbb{S}^2$. Then*

$$\partial_s \mathcal{R}f(y\cdot\theta,\theta) = \int_{\theta^\perp\cap\mathbb{S}^2} \nabla_\theta Df(y,\omega)\, d\omega,$$

where $\theta^\perp$ is the orthogonal complement of the span of θ and

$$\nabla_\theta Df(y,\omega) := \partial_t \big[Df(y, t\theta + \sqrt{1-t^2}\,\omega]\big|_{t=0}.$$

Proof. This proof may be skipped in lecture. We start from the right-hand side of the formula. Explicit computation shows

$$\nabla_\theta Df(y,\omega) = \int_0^\infty t\theta\cdot\nabla f(y+t\omega)\, dt,$$

and so the right-hand side equals

$$\int_{\theta^\perp\cap\mathbb{S}^2}\int_0^\infty t\theta\cdot\nabla f(y+t\omega)\, d\omega\, dt.$$

We next consider the left-hand side of the formula. Making a change of variables ($y = x - h\cdot\omega$), we may write

$$\mathcal{R}f(s+h,\theta) - \mathcal{R}f(s,\theta) = \int_{x\cdot\theta=s} [f(x+h\cdot\theta) - f(x)]\, dA,$$

which then implies

$$\partial_s \mathcal{R}f(s,\theta) = \int_{x\cdot\theta=s} \theta\cdot\nabla f(x)\, dA(x).$$

In what follows, we fix $s = y\cdot\theta$. Now consider the change of variables

$$x = a(t,\omega) = y + t\omega, \quad \omega \in \theta^\perp\cap\mathbb{S}^2, \quad t \in (0,\infty).$$

This is a bijection between $\{x\cdot\theta = s\}$ and $[0,\infty)\times\{\theta^\perp\cap\mathbb{S}^2\}$, with the inverse given by

$$\omega(x) = \tfrac{x-y}{|x-y|} \quad \text{and} \quad t = |x-y|.$$

The Jacobian factor is given by $dA(x) = t\, dt\, dA(\omega)$, and so we obtain

$$\partial_s \mathcal{R}(y\cdot\theta,\theta) = \int_{\theta^\perp\cap\mathbb{S}^2}\int_0^\infty t\theta\cdot\nabla f(y+t\omega)\, dt\, dA(\omega),$$

which agrees with what we obtained for the left-hand side of the formula. □

Imaging artifacts in X-ray CT

In this section we have two main goals. The first is to describe in some more detail the modeling of the X-ray source/detector pair and the measurement process. The second goal is to describe some common imaging artifacts that arise in X-ray CT. Imaging artifacts arise from several sources, including inaccuracies in the physical model, the sampling process, and measurement errors. Our discussion will be an abbreviated version of the one appearing in [1].

We will describe a model for the mapping from the true attenuation coefficient to the reconstructed image by means of a 'point spread function' (PSF) Ψ. Here 'point spread function' is put in quotation marks due to the fact that this filter will not actually be shift-invariant (due to the sampling process), and so calling Ψ a PSF is a bit of a misnomer. We write

$$f_\psi(x,y) = \int_{\mathbb{R}^2} \Psi(x,y;a,b) f(a,b)\, da\, db,$$

where f is the true attenuation coefficient and f_Ψ is the reconstructed image. The PSF Ψ should be built out of several pieces, including:

- a model for the source-detector pair,
- sampling of the measurements, and
- implementing the filtered back-projection step.

$\star \quad \star \quad \star$ *Partial Volume Effect* $\star \quad \star \quad \star$

So far, we have modeled X-rays as lines in $\mathbb{R}^2$. This is not accurate. A more accurate model (though still not the 'truth') is to model an X-ray beam as a strip in $\mathbb{R}^2$, with a corresponding *beam profile* w. We can view this profile as being built out of a source function w_s and a detector response function w_d. Typical examples used in modeling are

$$w_d(u) = \tfrac{1}{2\delta}\chi_{[-\delta,\delta]}(u) \quad \text{and} \quad w_s(u) = ce^{-u^2/\sigma},$$

and we say that the X-ray source has a *Gaussian focal spot*. In scanners in which the detectors are fixed and the sources move, the beam profile is given by the convolution $w = w_s * w_d$. We may refer to the FWHM of w as the *beam width*.

As we discussed earlier in these notes, we have been modeling the measurement of a one-dimensional X-ray beam by

$$\tfrac{I_o}{I_i} = \exp\{-\mathcal{R}f(t,\omega)\}.$$

For a strip with beam profile w, a better model is

$$\tfrac{I_0}{I_i} = \int w(u) \exp\{-\mathcal{R}f(t-u,\omega)\}\, du, \tag{1}$$

which is now a *nonlinear* function of the attenuation coefficient. An argument using Taylor series expansion leads to the following:

Lemma 1. *The following approximation holds:*

$$\log \tfrac{I_0}{I_i} \approx \int w(u)\mathcal{R}f(t-u,\omega)\, du + \mathcal{O}\bigg(\int w(u)[\mathcal{R}f(t-u,\omega) - \mathcal{R}f(t,\omega)]^2\, du\bigg).$$

If the variation of $\mathcal{R}f(t,\omega)$ is small over the width of the strip, then we recover an approximate linear model for measurement, namely the convolution

$$\mathcal{R}_w f(t,\omega) = \int w(u)\mathcal{R}f(t-u,\omega)\,du. \tag{2}$$

This is essentially a low-pass filter applied to $\mathcal{R}f$ in the affine parameter. Incorporating this into the reconstruction algorithm, we expect to reconstruct a smoothed version of f. This is good in the sense that it will reduce aliasing artifacts from sampling; however, it will also lead to a loss of resolution.

A nontrivial imaging artifact known as the *partial volume effect* arises in the case that $\mathcal{R}f(t,\omega)$ has large variation over the width of the strip. This arises, for example, if the X-ray beam passes through both bone and soft tissue. In this case, the error term in Lemma 1 dominates, and there can be a nontrivial difference between (1) and (2). This can occur even if there is only a small inclusion of more absorbent material like bone. In this setting, we are essentially using the 'wrong' data to reconstruct the attenuation coefficient, and the resulting image may have abnormally bright spots or streaks emanating from a hard object. An example (from [1]) is given in the following figure:

⋆ ⋆ ⋆ *Modeling the PSF* ⋆ ⋆ ⋆

In what follows, we will work with the *linear* model of measurement given by (2), as this is what our reconstruction algorithms actually assume. In this case, if the complete data were available, our reconstructed function would take the form

$$f_{\phi,w} = \mathcal{R}^* Q_\phi \mathcal{R}_w f.$$

Here $\mathcal{R}^*$ is the back-projection operator. We use Q_ϕ to denote the shift-invariant filter in the affine variable with impulse response ϕ. Here we choose ϕ so that

$$\hat{\phi}(r) = |r|\hat{\psi}(r) \approx |r|,$$

which corresponds to $Q_\phi \approx |\nabla|$. In fact, because $\mathcal{R}_w$ is also defined by convolution with w in the affine variable, we may write

$$f_{\phi,w} = \mathcal{R}^* Q_{\phi * w} \mathcal{R} f.$$

Writing $Q_{\phi*w}$ in the Fourier representation and applying the central slice theorem, we may obtain

$$Q_{\phi*w}\mathcal{R}f(t,\omega) = \tfrac{1}{2\pi}\int_{\mathbb{R}} e^{irt}\hat{\psi}(r)\hat{w}(r)\hat{f}(r\omega)|r|\,dr,$$

and so (back-projecting and using the fact that $r\,dr\,d\omega$ is the area element on $\mathbb{R}^2$)

$$f_{\phi,w}(x,y) = \tfrac{1}{4\pi^2}\int_{\mathbb{R}^2} \hat{f}(\xi)\hat{\psi}(|\xi|)\hat{w}(|\xi|)e^{i(x,y)\cdot\xi}\,d\xi$$

In particular, we derive:

Proposition 1. *The modulation transfer function (MTF) for $f \mapsto f_{\phi,w}$ is given by*

$$\xi \mapsto \hat{\psi}(|\xi|)\hat{w}(|\xi|).$$

The point spread function (PSF) is given by

$$(x,y) \mapsto \tfrac{1}{4\pi^2}\int_{\mathbb{R}^2} \hat{\psi}(|\xi|)\hat{w}(|\xi|)e^{i(x,y)\cdot\xi}\,d\xi.$$

We next incorporate sampling into the model. We work exclusively with the parallel beam geometry in this section, and we first consider the effect of ray sampling. We will briefly discuss view sampling later in this section.

Let d be the sample spacing in the affine parameter, and suppose our reconstruction involves a linearly interpolated filter ϕ. We suppose that we have complete data in the angular variable, and that our data consists of the samples

$$\{\mathcal{R}_w f(jd,\omega) : j = -N,\dots,N\}.$$

We will derive the following.

Proposition 2. *The PSF for the measurement and reconstruction process with sampling in the affine parameter is given by*

$$\Psi(x,y;a,b)$$
$$= \tfrac{1}{4\pi^2}\int_0^\infty\int_0^{2\pi} \operatorname{sinc}^2(\tfrac{rd}{2})\hat{\phi}_p(r)e^{i\langle(x-a,y-b),r\omega\rangle}\left[\sum_{j\in\mathbb{Z}}\hat{w}(r+\tfrac{2\pi j}{d})e^{-\frac{2\pi i j}{d}\langle(a,b),\omega\rangle}\right]dr\,d\omega,$$

where d is the sample spacing, w is the beam profile, ϕ is the filtering function, and

$$\hat{\phi}_p(r) = \sum_{j\in\mathbb{Z}}\phi(jd)e^{-ijdr}.$$

Proof. Our reconstructed image takes the form

$$\tilde{f}_{\phi,\omega}(x,y) = \tfrac{1}{2\pi}\int_0^{2\pi} Q_{\phi,w}\tilde{f}(\langle(x,y),\omega\rangle)\,d\omega,$$

where

$$Q_{\phi,w}\tilde{f}(t,\omega) = d\sum_{j\in\mathbb{Z}}\phi(t-jd)\mathcal{R}_w f(jd,\omega).$$

In this formula, we suppose we have extended our data by zero beyond $t = -dN$ and $t = dN$. In practice, we will compute the Fourier transform of this quantity, which is given by

$$\widetilde{Q_{\phi,w}\tilde{f}}(r,\omega) = d\hat{\phi}(r)\sum_{j\in\mathbb{Z}}e^{-ijdr}\mathcal{R}_w f(jd,\omega). \tag{3}$$

To derive the PSF $\psi(x,y;a,b)$, we need to derive the reconstruction of a point source at (a,b), which we model by $f = \delta_{(a,b)}$. In this case, first observe that

$$\mathcal{R}\delta_{(a,b)}(t,\omega) = \delta(t-\omega\cdot(a,b)),$$

and so

$$\mathcal{R}_\omega \delta_{(a,b)}(jd,\omega) = w(jd - \omega\cdot(a,b)).$$

We next wish to express $\hat{\phi}(r)$ solely in terms of the sampled values $\phi(jd)$. In fact, we claim that

$$\hat{\phi}(r) = \operatorname{sinc}^2(\tfrac{rd}{2})\hat{\phi}_p(r), \quad \hat{\phi}_p(r) = \sum_{j\in\mathbb{Z}} \phi(jd)e^{-ijdr}. \tag{4}$$

This follows from the fact that we are using a linearly interpolated filter. The proof, *which may be skipped in lecture*, is as follows:

Proof of (4). We may write the linearly interpolated filter as

$$\phi(r) = \sum_{j\in\mathbb{Z}} \big[\theta_j(r)\phi(jd) + (1-\theta_j(r))\phi((j+1)d)\big]\chi_{[jd,(j+1)d)}(r),$$

where

$$\theta_j(r) = \frac{(j+1)d - r}{d}.$$

We now rewrite this to see the contribution of each sample $\phi(jd)$. In particular, we split the sum into two pieces, change variables in the second sum, and recombine. This yields

$$\phi(r) = \sum_{j\in\mathbb{Z}} \phi(jd)\left[\tfrac{(j+1)d-r}{d}\chi_{[jd,(j+1)d)}(r) + \tfrac{r-(j-1)d}{d}\chi_{[(j-1)d,jd)}(r)\right].$$

Thus $\phi(jd)$ appears whenever $|r-jd| \in [0,d]$, and in this case both of the factors appearing in front of the characteristic functions may be rewritten as $\frac{d-|r-jd|}{d}$. That is,

$$\phi(r) = \sum_{j\in\mathbb{Z}} \phi(jd)\big[\tfrac{d-|r-jd|}{d}\chi_{[0,d]}(|r-jd|)\big] = \sum_{j\in\mathbb{Z}} \phi(jd)G(r-jd),$$

where

$$G(r) := \big[1 - \tfrac{|r|}{d}\big]\chi_{[0,d]}(|r|) = (\chi_{[-\frac{d}{2},\frac{d}{2}]} * \chi_{[-\frac{d}{2},\frac{d}{2}]})(r).$$

Now the result follows from taking the Fourier transform! □

Returning to (3), we have arrived at

$$\widetilde{Q_{\phi,\omega}\tilde{f}}(r,\omega) = d\operatorname{sinc}^2(\tfrac{rd}{2})\hat{\phi}_p(r)\sum_{j\in\mathbb{Z}} e^{-ijdr}w(jd-\omega\cdot(a,b)).$$

We evaluate the sum by the dual Poisson summation formula (with '$d = \frac{\pi}{L}$'), which yields

$$d\sum_{j\in\mathbb{Z}} w(jd-\omega\cdot(a,b))e^{-ijdr} = e^{-ir\omega\cdot(a,b)}\sum_{j\in\mathbb{Z}} \hat{w}(r+\tfrac{2\pi j}{d})e^{-\frac{2\pi ij}{d}(a,b)\cdot\omega}.$$

This computation requires that w and $\hat{w}$ decay sufficiently fast. We conclude

$$\widetilde{Q_{\phi,w}\tilde{f}}(r,\omega) = \operatorname{sinc}^2(\tfrac{rd}{2})\hat{\phi}_p(r)e^{-ir\omega\cdot(a,b)}\sum_{j} \hat{w}(r+\tfrac{2\pi j}{d})e^{-\frac{2\pi ij}{d}(a,b)\cdot\omega}.$$

To derive the formula for $\Psi(x,y;a,b)$, we now apply the inverse Fourier transform, i.e.

$$Q_{\phi,w}\tilde{f}(t,\omega) = \tfrac{1}{2\pi}\int e^{i(t-\omega\cdot(a,b))r}\,\mathrm{sinc}^2(\tfrac{rd}{2})\hat{\phi}_p(r)\sum_j \hat{w}(r+\tfrac{2\pi j}{d})e^{-\frac{2\pi ij}{d}(a,b)\cdot\omega}\,dr$$

Integration with respect to ω with $t=\omega\cdot(x,y)$ now yields the desired formula. □

Remark 1. *Translation invariance has been broken, in the sense that $\Psi(x,y;a,b)$ no longer depends only on $(x-a,y-b)$. This is due to sampling, which manifests in the infinite sum. Similarly, the symmetry between ϕ and w found in the continuum model has also been lost.*

Remark 2. *The infinite sum appearing in the PSF will lead to aliasing errors, with sharper beam profiles producing larger errors.*

⋆ ⋆ ⋆ *Resolution; Oscillatory Artifacts* ⋆ ⋆ ⋆

Let us now consider in more detail some of the parameters present in the reconstruction algorithm and our model thereof. These include:

- The beam profile w with corresponding beam-width δ. This is largely a physical problem (i.e. producing narrow beams of X-rays), but of course our model plays a role in analyzing the subsequent imaging artifacts.
- The sample spacing d.
- The filter ϕ, or equivalently the apodizing function ψ.

The amount of resolution in our reconstruction algorithm should be related to how sharply peaked the PSF is, at least in the shift-invariant setting that does not include sampling. When sampling is included, one typically looks instead at $(x,y)\mapsto\Psi(x,y;0,0)$. In what follows, let us explain the following principle:

- The spatial resolution in our image largely depends on the sample spacing d, although it is ultimately limited by the beam width δ. Taking fewer than one sample per beam width may lead to aliasing effects, wheres taking more than two does not lead to much improvement in resolution.

For the first point, let us consider the formula appearing in Proposition 2 and view things 'on the Fourier side'. Instead of considering how peaked the PSF is, let us consider the spread the MTF is. In the terms arising from the filter, namely

$$r\mapsto \mathrm{sinc}^2(\tfrac{rd}{2})\sum_{j\in\mathbb{Z}}\phi(jd)e^{-ijdr},$$

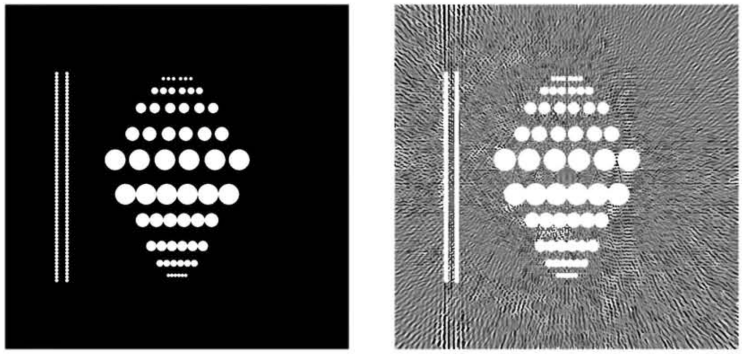
we see from the dependence on rd that in general, the function will spread as we decrease the sample spacing d, leading to increased resolution. However, we do not see this dependence in the term arising from the beam profile, i.e.

$$r\mapsto\sum_{j\in\mathbb{Z}}\hat{w}(r+\tfrac{2\pi j}{d}).$$

Thus we expect that the resolution depends on the beam width (with a larger beam corresponding to lower resolution), with some sort of fundamental limit in resolution arising from the beam profile. In particular continuing to decrease d will not improve resolution after a certain point. In order to avoid aliasing effects, however, we should take d sufficiently small depending on δ. To see this, we note that if the w is mostly supported in a ball of radius δ, then we may expect $\hat{w}$ is

mostly supported in a ball of radius δ^{-1}. Thus the ratio $\frac{\delta}{d}$ (related to the number of samples per beam width) should be sufficiently large to avoid aliasing. The optimal choice seems to be about two samples per beam width, which has been found both by studying some examples and by experimentation.

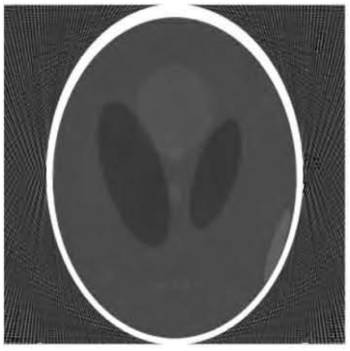
The resolution of a scanner or reconstruction algorithm can be measured by studying so-called 'resolution phantoms', as in the following figure from [1]:



(a) A resolution phantom. (b) Its reconstruction using a fan beam algorithm.

If the transfer function involves a sharp cutoff in frequency, then the PSF may have long oscillatory tails. These lead to oscillatory effects. If the transfer function is smooth and decaying, then this can be avoided, and we should see some blurring but without oscillatory artifacts. As we saw in (4), linear interpolation for the filter leads to a smooth, decaying Fourier tranform. We also note that Gibbs-like artifacts may arise if the sample spacing is larger than the beam.

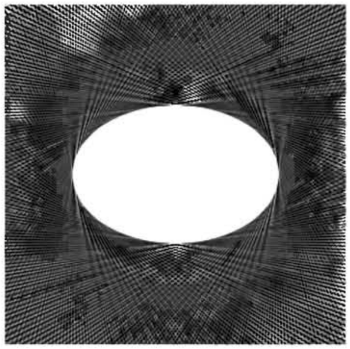
The objects themselves can lead to oscillatory artifacts and aliasing in the images, due to the fact that discontinuities lead to slow decay in the Fourier transform. Accordingly, reconstruction algorithms are frequently tested on mathematical phantoms involving characteristic functions of disks and polygons (which in particular have sharp edges). These oscillatory effects are present in the following figure from [1], which shows the reconstruction of a phantom using parallel beam data:



Even the numerical methods used to carry out the reconstruction algorithms can lead to artifacts in the image. For example, artifacts with a rectangular symmetry may arise due to the fact that we use rectangular partial sums in our approximation to the inverse Fourier transform.

⋆ ⋆ ⋆ *View Sampling* ⋆ ⋆ ⋆

It remains to consider the effect of sampling in θ. For this, let us just take a look at one example presented in [1] and give a purely qualitative description. Consider the following reconstruction of an elliptical phantom using filtered back-projection:



In this figure, one can see many oscillatory artifacts outside of the body of the ellipse.

- For points that lie outside the ellipse, the back-projection will involve lines outside the ellipse but very close to it. This leads to oscillation along lines tangent to the boundary via Gibbs effects and aliasing.
- In general, for exterior points the filtered back-projection formula only vanishes due to cancellation. When points are far from the boundary, there are not enough samples to obtain this cancellation and we see a pattern of oscillation.
- Oscillation near to the boundary and parallel to the boundary may occur due to ray sampling (through Gibbs/aliasing effects).

Further discussion in [1] shows that the oscillatory artifacts appear at a distance $\approx \Delta\theta^{-1}$ from the body. Furthermore, [1] re-derives the estimate $\Delta\theta \lesssim dL^{-1}$ in terms of choosing a suitable sample spacing for θ to avoid sampling artifacts as much as possible. Whereas our previous derivation of relied on an assumption of bandlimiting, in the present case we know that bandlimiting holds due to the finite beam width.

$\star \quad \star \quad \star$ *Measurement Errors* $\star \quad \star \quad \star$

In [1], one finds a very interesting discussion of the types of artifacts that can arise from measurement errors, e.g. from faulty detectors. There is a discussion of the effect of a single bad ray, a bad ray in each view, and one entire bad view. The discussion concludes by pointing out that since these artifacts have been so well analyzed and understood, they are largely absent from CT images! We will therefore skip this discussion, but refer those interested to [1, Section 12.4].

$\star \quad \star \quad \star$ *Beam Hardening* $\star \quad \star \quad \star$

Finally, let us give a brief discussion of beam hardening, which we recall is related to the faulty assumption that our X-ray beams are monochromatic. In particular, the energy of the incident X-ray beam is described via

$$I_i = \int_0^\infty S(\lambda)\, d\lambda,$$

where $S(\cdot)$ is the spectral function. The attenuation coefficient depends on λ, and the measured output of an X-ray beam directed along $\ell_{t,\omega}$ could be modeled by

$$I_o(t,\omega) = \int_0^\infty S(\lambda) \exp\left\{ - \int_{\mathbb{R}} f(s\hat{\omega} + t\omega, \lambda)\, ds \right\} d\lambda.$$

In particular, we do *not* actually measure the Radon transform itself. Essentially, our reconstruction algorithms are based on the wrong data! We say that are our measurements are inconsistent. The effect of this is *beam hardening*, which manifests as dark streak artifacts that look similar to those caused by the partial volume effect.

In some special cases, we can reduce the inconsistency of our measurements as follows. Let $f_w(\lambda)$ denote the attenuation coefficient of water, and suppose that we are imaging soft tissue. Then the ratio

$$\rho = \rho(x,\lambda) = \tfrac{f(x,\lambda)}{f_w(\lambda)}$$

will be nearly independent of λ. We make the simplifying assumption that this ratio *is* independent of λ, so that $\rho = \rho(x)$.

Using ρ, we may express our measurement via

$$I_o(t,\omega) = \int_0^\infty S(\lambda) \exp\left\{ - f_w(\lambda) \int_{-\infty}^\infty \rho(s\hat{\omega} + t\omega)\, ds \right\} d\lambda.$$

Now, consider the function

$$H(T) := \log\left[\frac{\int S(\lambda) e^{-T f_w(\lambda)}\, d\lambda}{\int S(\lambda)\, d\lambda} \right]$$

This is a function that can be (in principle) computed explicitly. Moreover, because S and f_w are nonnegative, we have that H is strictly decreasing and hence invertible. Again, the values of its inverse could be computed. This function is connected to our measurement as follows:

$$\log\big[\tfrac{I_o(t,\omega)}{I_i}\big] = H(\mathcal{R}\rho(t,\omega)) \implies \mathcal{R}\rho(t,\omega) = H^{-1}\big(\log\big[\tfrac{I_o(t,\omega)}{I_i}\big]\big).$$

Thus, our measurements should suffice to determine the Radon transform of ρ. If we reconstruct ρ, then we have an approximation to the attenuation coefficient of our body.

Of course, the preceding approach will *not* work if the body we are imaging consists of very different types of tissues. This is a much more challenging problem beyond the scope of these notes.

Noisy measurements in X-ray CT

This section will give a brief discussion of the modeling of noise/randomness in X-ray CT. There are three main sources of noise, which are all connected to the randomness inherent in our quantum description of the photons comprising X-ray beams. These are the generation of X-rays, the absorption or scattering of X-rays, and the detection of X-rays.

In the first part of this section, we will introduce a few topics from probability theory, which provide the language needed to describe randomness. In the second part of the section, we will describe how randomness may be incorporated into our description of the X-ray reconstruction algorithm. Our discussion will be short and omit many details (which may instead be found in [1]). To guide our efforts, we will work primarily towards the goal of deriving the 'fourth power law' in X-ray CT, which states the following:

- *To increase the resolution by a factor of 2 while keeping the signal-to-noise ratio constant, we must increase the X-ray dosage by a factor of 16.*

Here the 'signal-to-noise ratio' refers to the expected value of the measurement divided by the standard deviation of the measurement; in this setting, it is related to the 'contrast' available in the image. This is an important physical fact that speaks to some inherent limitations in X-ray CT imaging in light of the consideration of patient safety.

♣ ♣ ♣

Some Topics from Probability Theory

♣ ♣ ♣

Recall that we introduced the notion of 'Lebesgue measure' of subsets of $\mathbb{R}^n$, which was meant to to make our notion of 'volume' precise. More generally, a *measure* is a function μ that takes in sets and assigns a nonnegative value. Various technical conditions must be satisfied to earn the name 'measure'. An important such condition is the following: if A and B are disjoint 'measurable' sets, then

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

There are many examples of measures beyond Lebesgue measure. For example, we have the 'counting measure', defined on subsets of $\mathbb{Z}$ as the number of elements in the set. Any measure gives rise to a theory of integration with respect to that measure. The construction is the same as it was for Lebesgue measure: the integral of χ_E is $\mu(E)$, which is then extended to simple functions and then to more general measurable functions.

Example 1. *An important example that is very different in character from Lebesgue measure is that of a 'Dirac mass' at a point. For example, we may define the Dirac mass at $x = 0$ by*

$$\mu_0(E) = \begin{cases} 1 & 0 \in E \\ 0 & 0 \notin E. \end{cases}$$

In this case,

$$\int_E f d\mu_0 = \begin{cases} f(0) & 0 \in E \\ 0 & 0 \notin E. \end{cases}$$

We can then form measures by taking sums of Dirac masses at different points. For example, counting measure # consists of Dirac masses at each integer. Then if $f : \mathbb{Z} \to \mathbb{R}$ is given by $f(n) = n^2$ and $S = \{1, 2, 3\}$, we obtain

$$\int_S f \, d\# = 1 + 4 + 9 = 14.$$

We call measures consisting of Dirac masses 'singular' with respect to Lebesgue measure, since they are supported entirely on sets of Lebesgue measure zero. On the other hand, many measures are 'absolutely continuous' with respect to Lebesgue measure (meaning they assign zero measure to any set of Lebesgue measure zero).

Example 2. *Let w be a nonnegative, integrable function on $\mathbb{R}$. Define the measure $w(E) = \int_E w(x)\,dx$, where dx denotes Lebesgue measure. Then*

$$\int_{\mathbb{R}} f \, dw = \int_{\mathbb{R}} f(x)w(x)\,dx.$$

Such a measure is absolutely continuous with respect to Lebesgue measure.

A *probability measure* on a set Ω is a measure such that the measure of Ω equals one. We use the notation $\mathbb{P}$ to denote a probability measure. A measurable subset of Ω is called an *event*, and the measure of such a set is called the *probability* of that event.

A *random variable* X is a measurable function defined on Ω, i.e. $X : \Omega \to \mathbb{R}$. It is convention to minimize reference to the explicit elements $\omega \in \Omega$ (or even to say too much about Ω in general). For example, instead of writing

$$\mathbb{P}[\{\omega \in \Omega : X(\omega) > 2\}],$$

one will write

$$\mathbb{P}[\{X > 2\}]$$

and read this as 'the probability that X is greater than 2'.

Example 3. *A simple example of a random variable is given by the characteristic function of a set $S \subset \Omega$, i.e. $X = \chi_S$. In the language of probability, however, we would say X is the* indicator function *of the* event *S and write $X = \mathbb{1}_S$. [In fact, the 'characteristic function' means something else in probability—it refers to the Fourier transform!]*

We may wish to model the outcome of an experiment (the result of which returns some number) as a random variable X. What should the probability space be? This is less clear. In many cases, it does not really matter. In fact, we are typically not concerned with outcome of a single experiment (i.e. the value of $X(\omega)$ for some $\omega \in \Omega$), but rather the statistical properties obtained by repeating the same experiment many times. For example, we may be interested in the *expected value* (or *mean*, or *average value*) of X, denoted by

$$\mathbb{E}[X] = \int X \, d\mathbb{P}.$$

More generally, we may wish to know the expected value of other functions of X, which we may write as

$$\mathbb{E}[f(X)] = \int f(X) \, d\mathbb{P}.$$

It turns out (for reasons we will not discuss in these notes) that one can instead express these quantities as integrals over $\mathbb{R}$ with respect to a suitable measure μ_X, i.e.

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}} f(t) d\mu_X.$$

We call μ_X the *distribution* or *law* of the random variable. In many cases, this measure is of the type presented in Example 2, so that

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}} f(t)\mu_X(t)\, dt$$

for some nonnegative $\mu_X \in L^1$. This function is called the *distribution function* of X. Knowledge of this function allows us to compute all of the relevant statistical quantities associated to the random variable X.

Example 4. *Suppose X is a random variable with distribution function μ_X. Then*

$$\mathbb{P}[\{a \leq X \leq b\}] = \int_{X \in [a,b]} d\,\mathbb{P} = \mathbb{E}[\mathbb{1}_{[a,b]}(X)] = \int_a^b \mu_X(t)\, dt.$$

With the following discussion in mind, let us introduce some of the fundamental statistical quantities of random variables that will be of interest to us.

Definition 1. *Let X be a random variable with distribution function μ_X.*

- *The* expected value *of X is given by*

$$\bar{X} = \mathbb{E}[X] = \int_{\mathbb{R}} t\mu_X(t)\, dt.$$

- *The* variance *of X is given by*

$$\mathrm{Var}\,[X] = \mathbb{E}[X^2 - \mathbb{E}(X)^2].$$

- *The* standard deviation *of X is given by $\sigma = \sqrt{\mathrm{Var}\,[X]}$. In particular, the variance may be written as σ^2.*
- *The* signal-to-noise ratio *is given by*

$$\mathrm{SNR} = \frac{\bar{X}}{\sigma}.$$

Other quantities of interest include *moments* of X, i.e. $\mathbb{E}[|X|^p]$. These are basically just the L^p-norms with respect to the probability measure! One can also compute the *characteristic function* of X, which is the function $\xi \mapsto \mathbb{E}(e^{iX\xi})$. This is basically the Fourier transform of the distribution function of X! In particular, knowledge of the characteristic function completely characterizes the random variable.

To incorporate randomness into a model, we may describe measured quantities as random variables. Often, we make assumptions (hopefully with some justification) about how the random variables are distributed. In what follows, we consider several important examples.

Example 5. *A* Gaussian random variable *is determined by two parameters, the mean $\bar{x}$ and the standard deviation σ. The distribution function is given by*

$$\mu_X(t) = \tfrac{1}{\sqrt{2\pi}\sigma} e^{-(x-\bar{x})^2/2\sigma^2}.$$

This is also called a normal distribution, and may be denoted $X \sim N(\bar{x}, \sigma)$ (with $\sim$ read as 'is distributed as'). It is ubiquitous and used often in modeling, which may often be justified by appealing to the 'Central Limit Theorem'.

Example 6. *A* Bernoulli *or* binomial *random variable is also given by two parameters, a probability $p \in [0,1]$ and an integer $N \in \mathbb{N}$. The distribution is given by*

$$d\mu_X = \sum_{k=0}^{N} \binom{N}{k} p^k (1-p)^{N-k} \delta_k,$$

where δ_k is a Dirac mass at k. This means that X only takes on the values $\{0, 1, \dots, N\}$ and that

$$\mathbb{P}[\{X = k\}] = \binom{N}{k} p^k (1-p)^{N-k}.$$

This models a scenario in which one repeats an experiment N times (independently) with a probability p of 'success' each time. We have

$$\mathbb{E}[X] = pN, \quad \mathrm{Var}\,[X] = p(1-p)N.$$

The standard example of a Bernoulli random variable is a coin flip. In our setting, this could be a reasonable model for the detection of X-rays, where N photons reach the detector and each has a probability p of being detected.

Example 7. *A* Poisson *random variable is determined by a single parameter λ (called the intensity). The distribution is given by*

$$d\mu_X = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \delta_k,$$

where δ_k is a Dirac mass at k as before. In particular, X only takes the values $\{0, 1, 2, \dots\}$, and one can show that

$$\mathbb{E}[X] = \lambda \quad \textit{and} \quad \mathrm{Var}[X] = \lambda.$$

This is used to model many different situations, such as radioactive decay and arrival times. In our setting, the most relevant example is the generation of X-rays. The SNR for a Poisson random variable of intensity λ is given by $\sqrt{\lambda}$.

There are many important and subtle concepts in probability theory that we will not really discuss here, including the notions of independence, covariance, and conditional probability. Let us conclude with only a brief mention of independence. The notion of independence of two events is relatively straightforward, namely, that

$$\mathbb{P}[\text{Event 1 AND Event 2}] = \mathbb{P}[\text{Event 1}] \cdot \mathbb{P}[\text{Event 2}].$$

We may also speak of independence of two random variables X and Y. This refers to the statement that

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$$

for functions f and g of X and Y. This is the form in which we will quote independence below.

Noisy Measurements in X-Ray CT

♣ ♣ ♣

In the following, we will allow ourselves to be fairly brief in presentation and to omit many details. The reader will find more discussion and more details in [1, Chapter 16].

We start by describing our model for our reconstructed attenuation coefficient using the discretized filtered back-projection formula. The formula, which will require some explanation, takes the following form:

$$\mathring{f}_\phi(x,y) = \tfrac{\pi d}{M+1} \sum_{j=0}^{M} \sum_{k=-N}^{N} \mathring{P}_{\theta_j}(kd)\phi\big[\langle (x,y), \omega(\theta_j)\rangle - kd\big].$$

- Here d is the sample spacing and kd denote the sample points for the affine variable. We write $\theta_j = j\Delta\theta$.
- The function ϕ is our filtering function, chosen so that $\hat{\phi}(\xi) \approx |\xi|$.
- $\mathring{P}_{\theta_j}(kd)$ denotes the measured value corresponding to the line $\ell_{kd,\omega(\theta_j)}$. Previously, this was simply modeled by $\mathcal{R}f(kd,\omega(\theta_j))$. This is assumed to be a *random variable*, and hence so is the reconstruction $\mathring{f}_\phi$.

Notational remark. The above notation demonstrates the following convention: we indicate random variables by adorning quantities with $\mathring{}$. This is not standard notation, but is just meant to remind us of the presence of randomness in this section.

Recall that our model for what is actually measured in X-ray CT is given by

$$\mathring{P}_\theta(kd) = \log\left[\frac{\mathring{N}_{in}}{\mathring{N}_\theta(kd)}\right], \tag{1}$$

where $\mathring{N}_{in}$ and $\mathring{N}_\theta(kd)$ are described as follows:

- $\mathring{N}_\theta(kd)$ is the number of photons measured by the detector corresponding to the line $\ell_{kd,\omega(\theta)}$. This may be modeled by a Bernoulli random variable.
- $\mathring{N}_{in}$ is the number of incident photons. This may be modeled by a Poisson random variable.

When taken together, the source-detector pair may be modeled as a single Poisson random variable (see [1, Section 16.1.1]). Therefore, in what follows, we will simply assume that $\mathring{N}_{in} \equiv N_{in}$ is *deterministic*, and that $\mathring{N}_\theta(kd)$ is modeled as a Poisson random variable.

As described early in these notes, Beer's law is probabilistic in nature, describing the probability that a given photon is absorbed or scattered by material with a given attenuation coefficient. In the present setting, Beer's law is the statement that

$$\mathbb{E}[\mathring{N}_\theta(kd)] = N_{in}\exp\{-\mathcal{R}f(kd,\omega(\theta))\}. \tag{2}$$

This is derived in [1, Section 16.1.2]. We will not present the details here. The basic idea is to split the line into small segments of length Δs and suppose that the probability of absorption along such a segment is given by $f(s)\Delta s$. One can then compute the probability that a given particle emerges from the body. Sending $\Delta s \to 0$ recovers Beer's law.

From (1), we deduce that

$$\mathbb{E}[\mathring{P}_\theta(kd)] = \log[N_{in}] - \mathbb{E}[\log \mathring{N}_\theta(kd)].$$

Because the logarithm is not linear, we cannot simply 'pass the expectation through the logarithm'. Instead, we use the following lemma:

Lemma 1. *Denote*

$$\bar{N}_\theta(kd) = \mathbb{E}[\mathring{N}_\theta(kd)].$$

Then we have the following approximation:

$$\mathbb{E}[\log \mathring{N}_\theta(kd)] \approx \log \bar{N}_\theta(kd) - \frac{1}{2\bar{N}_\theta(kd)}.$$

Idea of the proof; may be skipped in lecture. Because $N_\theta(kd)$ is assumed to be Poisson, we have an explicit formula for the distribution function, and hence we can compute

$$\mathbb{E}[\log \mathring{N}_\theta(kd)] = \sum_{\ell=0}^{\infty} \frac{\mathrm{Ln}(\ell)[\bar{N}_\theta(kd)]^\ell e^{-\bar{N}_\theta(kd)}}{\ell!},$$

where $\mathrm{Ln}(0) := 0$. We then use Taylor's formula to estimate the difference between $\mathbb{E}[\log \mathring{N}_\theta(kd)]$ and $\log \bar{N}_\theta(kd)$. In particular, for a random variable y with a large mean $\bar{y}$ and small variance σ^2, one can derive

$$\mathbb{E}[\log y] \approx \log \bar{y} - \tfrac{1}{2\bar{y}^2}\sigma^2.$$

Approximating the Poisson distribution with a suitable Gaussian distribution, one can then derive the final approximation. □

Corollary 1. *The measurement $\mathring{P}_\theta(kd)$ has the following statistics:*

$$\mathbb{E}[\mathring{P}_\theta(kd)] \approx \mathcal{R}f(kd, \omega(\theta)), \quad \mathrm{Var}\,[\mathring{P}_\theta(kd)] \approx \frac{1}{\bar{N}_\theta(kd)}.$$

Sketch of proof. Using the result above and Beer's law in the form (2),

$$\begin{aligned}
\mathbb{E}[\mathring{P}_\theta(kd)] &\approx \log[N_{in}] - \mathbb{E}[\log \mathring{N}_\theta(kd)] \\
&\approx \log[N_{in}] - \log \bar{N}_\theta(kd) + \frac{1}{2\bar{N}_\theta(kd)} \\
&\approx \mathcal{R}f(kd, \omega(\theta)).
\end{aligned}$$

Here we have discarded the $1/\bar{N}_\theta$ term under the assumption that $\bar{N}_\theta$ is a large number.

For the variance, we first observe that

$$\mathrm{Var}\,\mathring{P}_\theta(kd) = \mathbb{E}\big[(\mathring{P}_\theta(kd) - \bar{P}_\theta(kd))^2\big] \approx \mathbb{E}\left[\left(\log \frac{\mathring{N}_\theta(kd)}{\bar{N}_\theta(kd)}\right)^2\right],$$

where we have written $\bar{P}_\theta$ for the expected value of P_θ. This final quantity may be expressed as an integral using the probability distribution function and then estimated, leading to the approximation

$$\mathrm{Var}\,\mathring{P}_\theta(kd) \approx \frac{1}{\bar{N}_\theta(kd)}.$$

See [1, (16.22)–(16.24)]. □

At this point, we have an understanding of the statistics of a single measurement. In particular, the expected value is that of the Radon transform (as it should be), and we see that the variance is inversely related to the number of photons measured. This latter point makes the analysis significantly more complicated. Indeed, the

measurement error gets worse as the X-rays pass through (and are absorbed by) more tissue!

We now return to our model for the reconstructed image, namely,

$$\mathring{f}_\phi(x,y) = \tfrac{\pi d}{M+1} \sum_{j=0}^{M} \sum_{k=-N}^{N} \mathring{P}_{\theta_j}(kd)\phi\big[\langle (x,y), \omega(\theta_j)\rangle - kd\big].$$

By the linearity of expectation and Corollary 1, we have

$$\mathbb{E}\big[\mathring{f}_\phi(x,y)\big] = \tilde{f}_\phi(x,y).$$

That is, the expected value of our reconstruction agrees with our deterministic reconstruction formula.

We next consider the variance of the reconstructed image. This requires expanding out the square in

$$\mathbb{E}\big[(\mathring{f}_\phi(x,y) - \tilde{f}_\phi(x,y))^2\big]$$

and computing. Under the assumption that the measurements at different points are independent random variables, this reduces to

$$\begin{aligned}\operatorname{Var} \mathring{f}_\phi(x,y) &= \big[\tfrac{\pi d}{M+1}\big]^2 \sum_{j=0}^{M} \sum_{k=-N}^{N} \mathbb{E}[(\mathring{P}_{\theta_j}(kd) - \overline{P_{\theta_j}}(kd))^2]\phi^2\big[\langle (x,y), \omega(\theta_j)\rangle - kd\big] \\ &\approx \big[\tfrac{\pi d}{M+1}\big]^2 \sum_{j=0}^{M} \sum_{k=-N}^{N} \frac{1}{\bar{N}_{\theta_j}(kd)}\phi^2\big[\langle (x,y), \omega(\theta_j)\rangle - kd\big].\end{aligned}$$

We would now like to derive the fourth power law concerning the signal-to-noise ratio of our measurements. First, let us discuss the significance of the SNR in this context. Recall that in medical applications, the measured quantity takes values (in Hounsfield units) between around -1000 (for air) and 1000 (for bone), while the soft tissues occupy a range of about -50 to 60. This is only about 5% of the total range. The SNR in the measurement determines the 'numerical resolution' in the reconstructed attenuation coefficient, which is referred to as *contrast* in imaging. This is distinct from the notion of *spatial* resolution, which is in turn determined by parameters such as the beam width, sample spacing, and FWHM of the filters used in the reconstruction algorithm.

In what follows, we will make several simplifying assumptions:

- We assume that the object is a disk of radius R centered at the origin.
- We assume that the object has a constant attenuation coefficient m.
- We focus on estimating the variance at the reconstruction point $(x,y) = (0,0)$.

The first two assumptions imply that

$$\bar{N}_\theta(kd) = N_{in} e^{-2m\sqrt{R^2-(kd)^2}} \quad \text{for all} \quad \theta. \tag{3}$$

Indeed, this follows from (2) and the explicit computation of the Radon transform of the characteristic function of a disk.

Let us also recall that ϕ is chosen so that $\hat{\phi}(\xi) \approx |\xi|$. We assume that the spatial resolution is given by δ, which is consistent with ϕ being bandlimited with highest frequency $\approx \delta^{-1}$.

Specializing to $(x, y) = (0, 0)$ and using (3) and Plancherel's formula, we therefore obtain

$$\begin{aligned}\operatorname{Var} \mathring{f}_\phi(0,0) &\approx \left[\tfrac{\pi d}{M+1}\right]^2 \sum_{k=-N}^{N} \sum_{j=0}^{M} \frac{e^{2mR}}{N_{in}} \phi^2(-kd) \\ &\approx \frac{d\pi^2 e^{2mR}}{MN_{in}} \int \phi^2(t)\, dt \\ &\approx \frac{d\pi^2 e^{2mR}}{MN_{in}} \int |\hat{\phi}(t)|^2\, dt \approx \frac{de^{2mR}}{\delta^3 MN_{in}}.\end{aligned}$$

With $d \approx \delta$ (i.e. sampling spacing comparable to spatial resolution) and

$$\mathbb{E}\mathring{f}_\phi(0,0) \approx f(0,0) = m,$$

we obtain the following signal-to-noise ratio:

$$\mathrm{SNR} = \frac{\mathbb{E}\mathring{f}_\phi(0,0)}{\sqrt{\operatorname{Var} \mathring{f}_\phi(0,0)}} \approx m\delta M^{\frac{1}{2}} N_{in}^{\frac{1}{2}} e^{-mR}.$$

Observe that the SNR depends badly on the thickness and density of the object, and it also decreases as we decrease the resolution.

Now let us consider the dosage D of radiation absorbed by the center pixel (measured in rad/cm^3). The total photon density passing through the point $(0, 0)$ is $\approx MN_{in}e^{-mR}$. If the pixel size is proportional to the resolution δ, then the number of photons absorbed is $\approx \delta MN_{in}e^{-mR}$. If the slice thickness is also proportional to the resolution δ, then we obtain

$$D \approx \delta^{-2} MN_{in}e^{-mR},$$

and so we may rewrite

$$\mathrm{SNR} \approx m\delta^2 D^{\frac{1}{2}} e^{-\frac{1}{2}mR}.$$

This demonstrates the 'fourth power law', which shows that to increase resolution by a factor of 2 while keeping the SNR constant, we would need to increase the dosage by a factor of 16. The derivation of this law was our primary goal, and so we will be content to end our discussion here.

Introduction to MRI

In this section, we give an introduction to the technique of magnetic resonance imaging (MRI). This imaging modality is based on the physical phenomenon of 'nuclear magnetic resonance'. This is considerably more complicated than the situation in X-ray CT, which is based on the absorption or scattering of X-rays.

⋆ ⋆ ⋆ *The Physics of MRI* ⋆ ⋆ ⋆

A rotating charged particle has an associated magnetic moment μ. This means that it would experience a torque $\mu \times B$ if placed in the external magnetic field B. This fact is due in part to the orbital angular momentum of the particle, and in fact the magnetic moment due to the orbital motion is proportional to the orbital angular momentum. The surprising fact revealed by the famous Stern–Gerlach experiment (1921–1922) is that a particle may possess an 'intrinsic' magnetic moment that is independent of its orbital angular momentum. To account for this, the quantum mechanical model introduces a notion of *intrinsic* angular momentum, also known as *spin* angular momentum, which is also given as a multiple of the intrinsic magnetic moment; this latter fact is perhaps best regarded as an experimental fact (as opposed to a theorem, or even a definition). The total angular momentum is then the vector sum of these two different types of angular momentum.

The torque experienced by a particle determines the rate of change of the angular momentum. If we denote the angular momentum by J and the magnetic moment by μ, then a particle placed in an external magnetic field B evolves according to

$$\tfrac{d}{dt} J = \mu \times B,$$

at least in the classical model. In fact, by a general result known as Ehrenfest's theorem, this equation describes the evolution of the *expected value* of the corresponding quantum observables. In the idealized setting of a stationary particle, the angular momentum and magnetic moment would both be of the 'intrinsic' type. Recalling the proportionality $\mu = \gamma J$ (for some constant γ called the *gyromagnetic ratio*), we can then obtain the following:

$$\tfrac{d}{dt} \mu = \gamma [\mu \times B],$$

where μ denotes the (expectation value of) the magnetic moment of the particle.

As we will see, solutions to this ordinary differential equation exhibit *precession* about the direction B with the angular frequency $\omega_0 := \gamma |B|$ (known as the *Larmor frequency*). This precession, which manifests as detectable, rapidly varying magnetic fields at the Larmor frequency, is the 'magnetic resonance' at the heart of MRI.

For hydrogen protons in water, we have

$$\gamma \approx 42.5764 \times 10^6$$

(in units of (cycles/second)/Tesla). Here Tesla is a unit of magnetic induction. The magnets used in MR imaging devices are in the 1–3 Tesla range (∼5,000 times stronger than the magnetic field of the earth), resulting in resonant Larmor frequencies in the standard FM radio band 40–120 MHz. This is rather advantageous, as electromagnetic radiation at these frequencies is harmless to the body (unlike X-ray frequencies!) and it is technologically simple to work with such frequencies.

⋆ ⋆ ⋆ *The Bloch Phenomenological Equation* ⋆ ⋆ ⋆

Let us now turn to the problem addressed in MRI. We would essentially like to determine the distribution of protons in an object, which we may denote by a density $\rho = \rho(x, y, z)$. To begin, we apply a strong uniform magnetic field $B_0 = B_0(x, y, z)$. In light of the discussion above, we essentially regard each proton as a magnet, which we expect to become 'polarized' (i.e. aligned with the ambient field B_0). We then consider the *magnetization vector* M, which is meant to model the sum of all of the magnetic moments of the protons at each location in the sample. After some time, we obtain the so-called *equilibrium magnetization*

$$M_0(x, y, z) = c\rho(x, y, z)B_0(x, y, z),$$

where c is some constant depending on factors like the ambient temperature. In practice (e.g. at room temperature in a 1 Tesla field), only ≈ 1 in 10^6 moments will become aligned with B_0. Thus M_0 is actually a tiny fraction of B_0, and hence would be very difficult to detect directly. The basic idea in MRI is to apply additional fields in such a way that the interaction of the spins with these fields produces a detectable signal (i.e. due to Larmor precession).

To describe the macroscopic evolution of the magnetization vector $M = M(t, x, y, z)$ in an external magnetic field of the form

$$B = B(t, x, y, z) = B_0(x, y, z) + \tilde{B}(t, x, y, z), \quad \text{with} \quad |\tilde{B}| \ll |B_0|,$$

we may use the *Bloch equation*, introduced by Felix Bloch in 1946. The equation, which will require some explanation, takes the following form:

$$\tfrac{d}{dt}M = \gamma[M \times B] - \tfrac{1}{T_2}M^{\perp} - \tfrac{1}{T_1}[M^{\parallel} - M_0]. \tag{1}$$

The meaning of each term is as follows:

- The $M \times B$ term is the 'torque' term describing interactions of the spins with the magnetic field. The parameter γ is the gyromagnetic ratio. If B were time independent, this term would predict precession about B at frequency $\gamma|B(x, y, z)|$ (see below).
- The $M^{\perp}$ term is a *relaxation term*. Here $M^{\perp}$ denotes the *transverse component* of M relative to B_0, i.e. the component of M perpendicular to B_0. This term reflects the fact that the transverse components decay due to 'spin-spin' type interactions. The parameter T_2 encodes how quickly this relaxation occurs. As this is dependent on the material, one should actually regard T_2 as a function of (x, y, z). Typical values of T_2 are ~ 50ms.
- The $M^{\parallel}$ is another *relaxation term*. Here $M^{\parallel}$ denotes the *longitudinal component* of M relative to B_0, i.e. the component of M parallel to B_0. This term reflects the fact that the longitudinal components return to equilibrium due to dissipation of energy from the spins. As in the T_2 term, the parameter T_1 encodes how quickly this relaxation occurs and is material dependent. Typical values of T_1 are ~ 1s.

Remark 1. *In the Bloch equation, the spins at different points do not interact directly. Instead, the relaxation terms are meant to account for the average effect of such interactions.*

In MRI imaging, the goal is basically to design the time-dependent field $\tilde{B}(t, x, y, z)$ in order to put the magnetization into certain states, which will then result in a measurable signal from which we can infer information about the spatial distribution ρ. In practice, the fields are 'piecewise constant in time', with the time

between different fields small enough that the relaxation terms may be essentially ignored.

In what follows, we will consider several special cases of the Bloch equation corresponding to different choices of magnetic field B.

Example 1 (Uniform, time-independent background field, without relaxation)**.** *Suppose*

$$B \equiv B_0 = (0, 0, b_0), \quad T_1 = T_2 = \infty.$$

Then the Bloch equation at a fixed point (x, y, z) is simply

$$\begin{cases} \frac{d}{dt} M = \gamma M \times B_0, \\ M|_{t=0} = M^0, \end{cases}$$

where M^0 is the magnetization at $t = 0$ (not necessarily in equilibrium). Computing the cross product, we obtain

$$\frac{d}{dt} M = \gamma b_0 \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} M \tag{2}$$

This is a linear ODE and can be solved by matrix exponentiation. In particular, writing $\omega_0 = \gamma b_0$, the solution is given by

$$M(t) = U(t)M^0, \quad \textit{where} \quad U(t) = \begin{bmatrix} \cos \omega_0 t & \sin \omega_0 t & 0 \\ -\sin \omega_0 t & \cos \omega_0 t & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In particular, $U(t)$ is a rotation matrix, and we obtain precession at the Larmor frequency ω_0.

It is often convenient to introduce a *rotating reference frame* in which to study the Bloch equation. This entails defining $m = m(t, x, y, z)$ via the relation

$$M(t, x, y, z) = U(t)m(t, x, y, z)$$

and deriving the equation for m. In particular, assuming M_0 points in the z direction, one obtains the following:

$$\frac{d}{dt} m = \gamma [m \times B_{\text{eff}}] - \frac{1}{T_2} m^{\perp} - \frac{1}{T_1} [m^{\parallel} - M_0], \tag{3}$$

where $m^{\perp}$ and $m^{\parallel}$ are the transverse and longitudinal components relative to B_0 and the *effective magnetic field* is given by

$$B_{\text{eff}} = U^{-1}(t)B - (0, 0, b_0).$$

Proof of (3)*. This proof may be skipped in lecture.* Writing $M = U(t)m$, we first observe

$$\frac{d}{dt} M = U(t) \frac{d}{dt} m + [\frac{d}{dt} U(t)]m.$$

By definition of $U(t)$ (as matrix exponential), the time derivative of $U(t)$ equals the matrix on the right-hand side of (2). Recalling how this matrix was obtained (by computing the cross product), we observe

$$[\frac{d}{dt} U(t)]m = \gamma\, m \times (0, 0, b_0).$$

Now observe that since $U(t)$ is a rotation,

$$U(t)^{-1}[U(t)m \times B] = m \times U^{-1}(t)B.$$

Similarly, $U(t)$ leaves the $\perp$ and $\|$ spaces invariant, so that $U(t)^{-1}[U(t)m]^{\perp} = m^{\perp}$, and similarly for the $\|$ component. Furthermore, $U(t)^{-1}M_0 = M_0$ since M_0 is assumed to point in the z-direction. Putting together all the pieces leads to the desired equation. □

We turn to another special case of the Bloch equation. In this case, we consider the effect of adding a time-dependent spatially homogeneous radio frequency (RF) field that is perpendicular to the background field B_0. As we will see, this has the effect of 'tipping the magnetization vector out of equilibrium'. If we then turn off the RF field, Example 1 shows that the magnetization vector will begin to precess about B_0 with frequency ω_0 (until it eventually relaxes back to equilibrium).

We will consider RF fields of the form

$$B_1 = B_1(t) = ([\alpha(t) + i\beta(t)]e^{-i\omega_0 t}, 0), \quad \text{where} \quad \omega_0 = \gamma b_0. \tag{4}$$

where here we introduce the convention that a vector

$$(a + ib, 0) \in \mathbb{C} \times \mathbb{R} \quad \text{corresponds to} \quad (a, b, 0) \in \mathbb{R}^3.$$

Example 2 (Spatially homogeneous RF field over uniform background field, without relaxation)**.** *Suppose*

$$B = B(t) = B_0 + B_1(t),$$

where $B_0 = (0, 0, b_0)$ *and* $B_1(t)$ *is as in* (4) *with* $\beta(t) \equiv 0$*. Suppose further that* $T_1 = T_2 = \infty$*.*

In the rotating reference frame, the Bloch equation at a fixed point (x, y, z) *is*

$$\begin{cases} \frac{d}{dt}m = \gamma[m \times U(t)^{-1}B_1(t)], \\ m|_{t=0} = m^0, \end{cases}$$

where again m^0 *is not necessarily in equilibrium. Now observe that by construction,*

$$U(t)^{-1}B_1(t) = \begin{bmatrix} \cos(\omega_0 t) & -\sin(\omega_0 t) & 0 \\ \sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha(t)\cos(\omega_0 t) \\ -\alpha(t)\sin(\omega_0 t) \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha(t) \\ 0 \\ 0 \end{bmatrix},$$

which implies

$$\gamma[m \times U(t)^{-1}B_1(t)] = \gamma \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \alpha \\ 0 & -\alpha & 0 \end{bmatrix} m.$$

This is once again a linear ODE that we solve with matrix exponentiation, leading to

$$m(t) = V(t)m^0, \quad \textit{where} \quad V(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta(t) & \sin\theta(t) \\ 0 & -\sin\theta(t) & \cos\theta(t) \end{bmatrix}$$

with

$$\theta(t) = \gamma \int_0^t \alpha(s)\, ds.$$

Thus in the rotating reference frame, we obtain a rotation in the yz *plane by angle* $\theta(t)$*.*

So far, we have only considered spatially homogeneous fields. We consider now the effect of adding a constant (in time) *gradient field* of the form

$$G = G(*, *, \ell), \quad \ell = \ell(x, y, z) = (x, y, z) \cdot (g_x, g_y, g_z). \tag{5}$$

Here the $*$ is meant to denote small components that may safely be ignored. In truth, some nonzero component must be included lest we violate the divergence free condition arising in Maxwell's equations.

Example 3 (Gradient field over uniform background field, without relaxation)**.** *Suppose $B = B_0 + G$, with G as in* (5)*. Then, ignoring the $*$ components and arguing similarly to the previous examples, we find that (since B points only in the z direction)*

$$B_{\mathit{eff}} = (0, 0, \ell),$$

and hence the solution to the Bloch equation in the rotating reference frame is given by

$$m(t) = \begin{bmatrix} \cos(\gamma \ell t) & \sin(\gamma \ell t) & 0 \\ -\sin(\gamma \ell t) & \cos(\gamma \ell t) & 0 \\ 0 & 0 & 1 \end{bmatrix} m^0.$$

At the level of the original magnetization vector $M(t)$, we therefore observe precession about B_0 at the angular frequency $\gamma[b_0 + \ell(x, y, z)]$. In particular, we may encode information about the spatial location in the frequency of the precession.

⋆ ⋆ ⋆ *A Basic Imaging Experiment* ⋆ ⋆ ⋆

In this section, we describe a basic imaging experiment that will demonstrate how (in principle) we may use an applied magnetic field to produce a detectable signal that provides information about the density ρ. We perform the following steps:

(i) We put the sample in the field B_0 and allow it to become polarized.
(ii) We turn on a uniform RF field as in (4), with $\beta \equiv 0$. After a certain time (normalized to $t = 0$) we obtain $\theta(0) = 90°$, meaning that the magnetization vector is tipped into the xy-plane (uniformly across the sample).
(iii) We turn off the RF field at $t = 0$, and (according to Example 1) the magnetization vector $M(t)$ begins to precess about B_0 'in phase' (that is, with angular velocity $\omega_0 = \gamma b_0$).

At this point, relaxation effects begin to take over, and the magnetization vector obeys

$$M(t, x, y, z) \propto \omega_0 \rho(x, y, z)(e^{-t/T_2} e^{-i(\omega_0 t + \phi)}, 1 - e^{-t/T_1}),$$

where $\phi \in \mathbb{R}$ is some phase.

By Faraday's law, a changing magnetic field induces an electromotive force in a loop of wire (with the force given by the time derivative of the flux of the field). As the transverse components of M are a rapidly varying magnetic field, they generate a measurable signal, which then takes the form

$$S_0(t) \propto \omega_0^2 e^{-t/T_2} e^{-i\omega_0 t} \int \rho(x, y, z)\, dx\, dy\, dz.$$

Here we make the simplifying assumption that T_2 is independent of (x, y, z), and we encode the effect of the detector as simply another constant. This simple situation therefore provides a way to measure the total density of the object. The signal,

which is $\propto \omega_0^2 \propto b_0^2$, is still rather small in amplitude (e.g. for a 1.5 Tesla magnet we can expect a signal in the microwatts, i.e. 10^{-6} watts).

Now let us consider the same three steps above but now add a fourth step:

(iv) Turn on a gradient field $G = (*, *, \ell)$ as in (5).

In this case, incorporating the analysis of Example 3, we obtain the signal

$$\begin{aligned} S_\ell(t) &\propto \omega_0^2 e^{-t/T_2} e^{-i\omega_0 t} \int \rho(x,y,z) e^{-it\gamma(x,y,z)\cdot(g_x,g_y,g_z)}\, dx\, dy\, dz \\ &\propto \omega_0^2 e^{-t/T_2} e^{-i\omega_0 t} \hat{\rho}(k), \quad \text{where} \quad k = t\gamma(g_x, g_y, g_z). \end{aligned}$$

Repeating this for several different choices of ℓ, we can therefore obtain a radial sampling of $\hat{\rho}$ in a neighborhood of $k = 0$. In principle, this may be used to *reconstruct* ρ.

In practice, this approach is too crude. For example, the time decay of the signal limits how many samples we can take, which ultimately limits the highest frequency we can sample and hence the resolution of the reconstructed image. Other serious issues arise due to measurement noise, which is basically proportional to the volume of the part of the sample generating signal.

In what follows, we will describe how to design RF and gradient fields so that (i) we can excite a single $2d$ slice of the sample at a time and (ii) we can obtain samples of the ($2d$) Fourier transform on a uniform grid (rather than radial samples). With these two ingredients, we will be able apply the fast inverse Fourier transform on each $2d$ slice and ultimately obtain an efficient reconstruction of the entire $3d$ object.

⋆ ⋆ ⋆ *Selective Excitation* ⋆ ⋆ ⋆

In the process of *selective excitation*, the goal to excite (that is, disturb from equilibrium) a $2d$ slice of the magnetization vector. Without loss of generality, suppose we aim to excite a plane parallel to the xy-plane, determined by

$$(x, y, z) \cdot (0, 0, g) = \text{const}, \quad \text{i.e.} \quad z = \text{const}.$$

To begin, we apply the gradient field as in (5) with $(g_x, g_y, g_z) = (0, 0, g)$. In light of Example 3, we may then define the *offset frequency*

$$f = \tfrac{1}{2\pi}\gamma\langle (x, y, z), (0, 0, g)\rangle = \tfrac{g\gamma}{2\pi} z.$$

This represents the amount that the local resonance frequency differs from the unperturbed resonance frequency ω_0. With the background field $B_0 + G$, we therefore have a correspondence between z-location in the sample and offset frequency.

We now turn to the much more subtle problem of designing an RF 'pulse' that will tip the magnetization vector out of equilibrium only at spatial locations lying near our prescribed plane. Because of the correspondence given by G and the fact the field we apply will be uniform in space, we view the magnetization vector m (in the rotating reference frame) as a function of the offset frequency f. Assuming that the magnetization vector is initially in equilibrium, we may define our desired final *magnetization profile* $m^\infty = m^\infty(f)$ as follows. We aim to obtain

$$m^\infty(f) = \begin{cases} (0, 0, 1) & f \notin [f_0 - \delta, f_1 + \delta], \\ (\sin\theta, 0, \cos\theta) & f \in [f_0, f_1], \end{cases} \tag{6}$$

for some small interval $[f_0, f_1]$ and some angle θ (most often equal to $\frac{\pi}{2}$ or π), where the small parameter $\delta > 0$ allows for a smooth transition between the different states.

We expect to achieve this state by a short RF 'pulse' of the form (4) on some time interval $[t_0, t_1]$ of length much smaller than either T_1 or T_2 (e.g. a few milliseconds). Thus, in our model we will be content to ignore relaxation terms. Arguing as we did in Examples 2 and 3, we are therefore led to study the Bloch equation (in the rotating reference frame) with effective potential

$$B_{\text{eff}} = (0, 0, zg) + (\alpha(t), \beta(t), 0) = (\alpha(t), \beta(t), 2\pi\gamma^{-1}f).$$

Computing the cross product, we then obtain the ODE

$$\tfrac{d}{dt}m = \begin{bmatrix} 0 & 2\pi f & -\gamma\beta(t) \\ -2\pi f & 0 & \gamma\alpha(t) \\ \gamma\beta(t) & -\gamma\alpha(t) & 0 \end{bmatrix} m. \tag{7}$$

We thus arrive at the following problem:

Problem 1. *Find a function $\alpha(t)+i\beta(t)$ supported in $[t_0, t_1]$ such that the solution $m = m(t, f)$ to (7) with initial condition*

$$m(t_0, f) = (0, 0, 1)$$

obeys the final condition

$$m(t_1, f) = m^{\infty}(f),$$

where m^{∞} is given in (6).

Note that this is an inherently *nonlinear* problem. It is a classical example of a so-called *inverse scattering problem*. In what follows, we will consider the case of $|\theta| \ll 1$, which is easier than the general case. In this setting, we have $m_3 \approx 1$ throughout the pulse and we may approximate the nonlinear problem by a simpler linear problem. The solution will then be obtained by using the Fourier transform.

Under the assumption that $m_3 \equiv 1$, the equation (7) reduces to

$$\tfrac{d}{dt}(m_1 + im_2) = -2\pi i f(m_1 + im_2) + i\gamma(\alpha + i\beta).$$

This equation may be solved using the integrating factor $e^{2\pi i t f}$. This leads to

$$\tfrac{d}{dt}[e^{2\pi i t f}(m_1 + im_2)] = i\gamma e^{2\pi i t f}[\alpha + i\beta].$$

Under the assumption that

$$[m_1 + im_2]\big|_{t=t_0} = 0,$$

we derive

$$e^{2\pi i t f}[m_1 + im_2](t) = i\gamma \int_{t_0}^{t} e^{2i\pi s f}[\alpha(s) + i\beta(s)]\, ds.$$

Integrating to $t = t_1$ and recalling that $\alpha + i\beta$ is to be supported in $[t_0, t_1]$, we obtain the relation

$$e^{2\pi i t_1 f} m^{\infty}(f) = i\gamma \int_{-\infty}^{\infty} e^{2\pi i s f}[\alpha(s) + i\beta(s)]\, ds.$$

This may be rewritten in terms of the Fourier transform:

$$e^{2\pi i t_1 f} m^{\infty}(f) = 2\pi i\gamma \mathcal{F}^{-1}[\alpha + i\beta](f),$$

which becomes

$$[\alpha + i\beta](t) = \frac{1}{2\pi i\gamma}\mathcal{F}[m^{\infty}](t + t_1). \tag{8}$$

Example 4. *Suppose* $f_0 = -f_1$ *in* (6) *(so that we are exciting around* $z = 0$*) and* $|\theta| \ll 1$*. If* t_1 *is normalized to* $t_1 = 0$*, then* (8) *reduces to*

$$\alpha(t) + i\beta(t) = -\frac{\sin(\theta)\sin(2\pi t f_1)}{4\pi^2\gamma t}.$$

This is a sinc pulse*.*

By now, there are books full of MRI pulse sequences designed to produce prescribed magnetization profiles. A typical pulse lasts about 1ms or less! The derivation above was carried out in a simplified setting and under the assumption that θ is small. In fact, the same type of pulse may work even for angles up to $\frac{\pi}{2}$. To move to larger angles, one can use an approach known as the *Shinnar-Le Roux* algorithm for designing pulse sequences. As it turns out, the problem may also be re-formulated so that it becomes equivalent to a well-known inverse scattering problem, known as the 2×2 Zakharov–Shabat or AKNS system.[1] This provides another avenue for the designing pulse sequences to obtain prescribed magnetization profiles in a precise manner. Furthermore, one can obtain pulses obeying certain optimality conditions (such as the minimal energy pulse).

⋆ ⋆ ⋆ *Spin Warp Imaging* ⋆ ⋆ ⋆

Suppose we have applied a selective RF pulse so that the magnetization is flipped out of equilibrium in the region $|z - z_0| < \Delta z$ (and left in equilibrium outside of this region). We now consider the problem of imaging the $2d$ slice at $z = z_0$. The technique we describe is known as *spin warp imaging*.

As in the basic imaging experiment above, signal is produced due to the rapid change of the transverse components of the magnetization vector. In particular, signal will only be received from the region of the sample where $|z - z_0| < \Delta z$. In what follows, we write

$$\rho = \rho(x, y)$$

to denote $\rho(x, y, z_0)$, or (more precisely) for the average of ρ over the slice $|z - z_0| < \Delta z_0$.

Our goal is to sample the $2d$ Fourier transform of ρ on a uniform grid. We will achieve this by applying two gradient fields, known as the *phase encoding gradient* and the *frequency encoding gradient*. In particular, we take the following steps:

1. We apply a gradient

$$G_{\text{ph}} = (*, *, -g_x x + g_y y)$$

for some time T_{ph}. Arguing as in Example 3, we obtain the following for the transverse components (in the rotating reference frame):

$$m^{\perp}(x, y) \propto e^{i\gamma(g_x x - g_y y)T_{\text{ph}}}\rho(x, y),$$

[1] As an interesting coincidence, this is the same inverse scattering problem that arises in the study of the one-dimensional cubic nonlinear Schrödinger equation (an important 'completely integrable' model in mathematical physics).

where we use the correspondence $\mathbb{R}^2 \cong \mathbb{C}$ and observe that multiplication by $e^{i\theta}$ corresponds a rotation by angle θ. Defining

$$(k_x, k_y) = \tfrac{1}{2\pi}\gamma T_{\text{ph}}(g_x, g_y),$$

we may rewrite this as

$$m^{\perp}(x,y) \propto e^{-2\pi i(-k_x x + k_y y)}\rho(x,y).$$

This is called the *phase encoding step.* It takes about 1ms.

2. At time T_{ph}, we change the gradient to the form

$$G_{\text{fr}} = (*, *, g_x x)$$

Just as in the first step, after a time t the effect is multiplication of the of $m^{\perp}(x,y)$ by the rotation factor

$$e^{-it\gamma g_x x}.$$

In particular, if we begin reading the signal as we apply the field then we obtain

$$\begin{aligned} S(t) &\propto \int e^{-2\pi i[-k_x x + k_y y]} e^{-it\gamma g_x x}\rho(x,y)\,dx\,dy \\ &= \int e^{-2\pi i[k_y y + (-k_x + t\frac{\gamma}{2\pi}g_x)x]}\rho(x,y)\,dx\,dy. \end{aligned}$$

This is the *frequency encoding step.* It also takes on the order of 1ms.

If we sample the signal uniformly in time, then we obtain samples of

$$\hat{\rho}(-k_x + t\tfrac{\gamma}{2\pi}g_x, k_y).$$

Thus, for each fixed k_y, we can obtain uniformly spaced samples of $\hat{\rho}(\cdot, k_y)$.

Returning to Step 1, we can repeat this process to obtain samples of $\hat{\rho}$ for a discrete set of k_y values. That is, we may obtain a set of data of the form

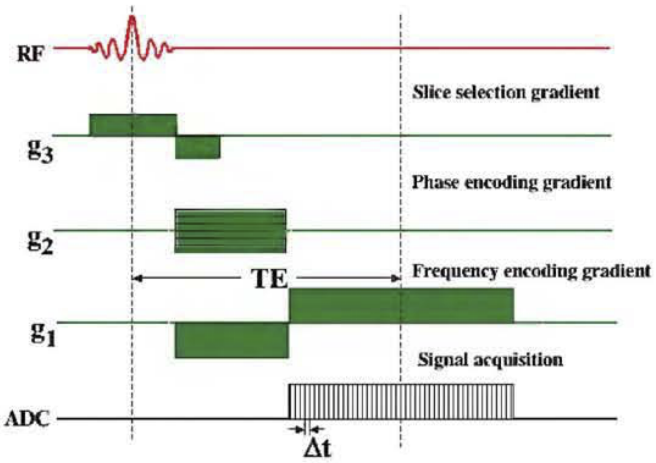
$$\hat{\rho}(m\Delta k_x, n\Delta k_y), \quad -N_x \le m \le N_x, \quad -N_y \le n \le N_y.$$

In practice, for a given slice one will need to repeat these steps many times in order to sample a large enough region of 'k-space'.

With samples of the Fourier transform in hand, the reconstruction process is simply a matter of implementing the Fourier inversion formula! From the description above, one can also begin a discussion of resolution in MRI, although we will not go into this topic here.

Remark 2. *We have described a process whereby we can sample the Fourier transform of (a 2d slice of) ρ on a uniform grid. It is possible to sample on different sets in 'k-space' by using different gradient fields. We will not pursue this topic here, but will point out that in some settings it is advantageous to sample along a fairly 'sparse' set in k-space and make use of compressed sensing techniques to carry out the reconstruction.*

The following figure from [1] is a schematic that shows the an example of a pulse sequence used in the spin warp imaging process.



The 'TE' in this figure refers to the notion of an '*echo time*'. In this case the diagram is referring to what is called a 'gradient echo'. In the next section we will discuss a related concept, namely, that of a *spin echo*.

⋆ ⋆ ⋆ *Spin Echoes and Contrast Mechanisms in MRI* ⋆ ⋆ ⋆

The preceding discussion describes how images may be produced in MRI. However, we have so far said very little about what types of contrast mechanisms are available in MRI, other than just the difference in proton density between different materials. Indeed, so far we have regarded T_1 and T_2 fixed constants, while in truth these two parameters are material-dependent and actually provide an important contrast mechanism in MRI. In what follows, we will mostly consider the contrast due to T_2.

In the preceding discussion, we discussed the application of RF pulses and field gradients of extremely short duration. Indeed, this was essential in many derivations in which we ignored the relaxation terms. In particular, the imaging processes we described would not 'notice' at all the variation in T_2 (or T_1) throughout the sample.

On the other hand, if we could arrange things so that the duration of our excitation and phase encoding was comparable to a typical T_2 value (e.g. 50ms), then when we started signal acquisition, the signal would be 'T_2-weighted'. That is, we could detect the variation in T_2 over the sample.

The problem is that we cannot simply excite our sample, and then 'wait around' sufficiently long before performing phase/frequency encoding and signal acquisition. The issue arises from the fact that while our model assumes a perfectly static background field leading to precession at the Larmor frequency, in reality there will always be field inhomogeneities that lead to slightly different frequencies of rotation throughout the sample. Thus, if we 'wait around', decoherence will occur and we will not ultimately obtain a useful signal. A typical timescale for such decoherence (known as T_2^* decay) is ~ 5ms.

Fortunately, there is a very clever remedy. The idea is to produce a '*spin echo*' as follows: Suppose we will perform our initial RF excitation pulse (to a flip angle of $\frac{\pi}{2}$, say) at time $t = 0$, but we wish to have the resulting state occur not after

$\sim$ 1ms, but rather after some $\sim 2\tau$ (known as the *echo time*, denoted TE). We can achieve this as follows:

1. First, perform an RF excitation corresponding to flip angle $\frac{\pi}{2}$ (in xz plane)
2. Wait for time τ.
3. Next, apply an RF pulse corresponding to flip angle π in the yz plane.
4. Wait for time τ.

The claim is that at the conclusion of Step 4, the magnetization vector is in the same state as at the end of Step 1. In this case, we have achieved our stated goal!

To see why this works, note that at the conclusion of the first RF pulse, the magnetization vector (at each spatial position we wish to excite) will be in the state $(1, 0, 0)$. As long as relaxation effects remain negligible, it will then evolve in time approximately as $(e^{it\Phi}, 0)$ for some phase Φ, which may vary in (x, y, z) due to field inhomogeneities. Thus, at time τ, the vector is in the state

$$(e^{i\tau\Phi}, 0) = (\cos\tau\Phi, \sin\tau\Phi, 0).$$

We now apply the RF pulse with flip angle π in the yz plane. This puts the vector in the state

$$(\cos\tau\Phi, -\sin\tau\Phi, -1 \cdot 0) = (e^{-i\tau\Phi}, 0).$$

Now it evolves as before, namely, as $(e^{it\Phi}e^{-i\tau\Phi}, 0)$. In particular, after time τ, the vector is back in the state $(1, 0, 0)$, as desired!

Note that we could also apply the second RF pulse to effect a flip angle π in the xz plane. Then we would obtain

$$(e^{i\tau\Phi}, 0) \mapsto (-\cos(\tau\Phi), \sin(\tau\Phi), 0) = (-e^{-i\tau\Phi}, 0)$$

and after another evolution by time τ the magnetization vectors would all be aligned at the state $(-1, 0, 0)$.

We now have two parameters that we can vary to produce different types of images that exploit different contrast mechanisms available in MRI. These are the echo time TE (corresponding to 2τ in the discussion above) and the repetition time TR. Here the repetition time refers to the amount of time between the initial RF excitation pulses, including the repeated phase encodings, frequency encodings, and signal acquisition.

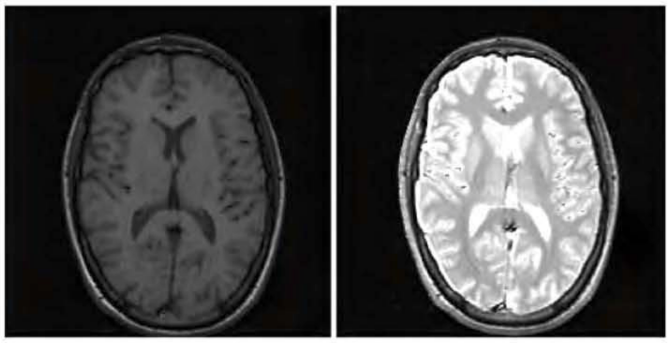
The initial approach we described (without spin echoes), may be described as obtaining *proton density weighted* image. Here the difference in proton density provides the main contrast mechanism.

By using spin echoes, we can also produce so-called T_2*-weighted images*. In this set up, we choose τ so that the echo time TE is comparable to a typical T_2 value. Thus the strength of our signal will vary according to the variation of T_2 throughout the sample. In this setting, we choose the repetition time TR longer than a typical T_1 value. This means we essentially wait for a complete return to equilibrium between each repetition. In particular, we will *not* detect any variation in the T_1 parameter.

Finally, there is also a notion of T_1*-weighted images*. In this case one chooses the echo time TE much smaller than a typical T_2 value (so that variation in T_2 over the sample will not be detected). In this case, however, one chooses TR less than a typical T_1 value. That is, we do not necessarily allow the entire sample to recover back to equilibrium. In this way, one can obtain an image that detects the variation

in T_1 over the sample. To produce such images, one makes use of a different tool known as a '*gradient echo*', which we will not discuss here.

To end the discussion, we display the following figure from [1] demonstrating the difference between a T_1-weighted MRI image (left) and a proton density weighted MRI image (right).



Introduction to PET

In this section, we present a brief introduction to the imaging modality of positron emission tomography (PET), adapted from [1].

⋆ ⋆ ⋆ *Physics of PET* ⋆ ⋆ ⋆

The physics of PET is based on electron-positron annihilation events. In particular, when an electron and positron meet, they annihilate each other and produce a pair of γ-ray photons. In medical applications, an organic element with a short-lived isotope that decays via positron emission is injected into the patient. The substance is metabolized in the body and taken up by structures in the body. The distribution of this substance is described by a density ρ that is reconstructed by counting γ-ray photons outside of the body. As in X-ray CT and MRI, we proceed by constructing one slice at a time.

In the context of medical applications, the momentum of the incoming pair is small. Under the simplifying assumption that the momentum is zero, it follows that the outgoing γ-rays travel in opposite directions.

⋆ ⋆ ⋆ *Probabilistic model for PET* ⋆ ⋆ ⋆

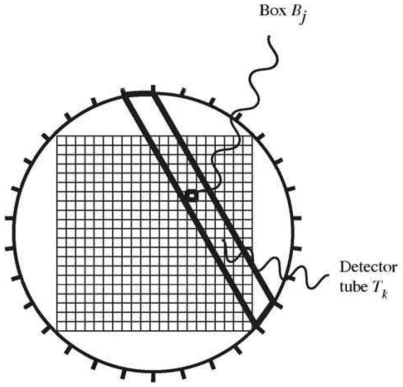
Suppose ρ is supported in a bounded region $D \subset \mathbb{R}^2$. We divide D into a disjoint collection of boxes $\{B_j\}_{j=1}^N$. Then the number of decays in a fixed time in B_j may be modeled as a Poisson random variable n_j with intensity

$$\lambda_j = \iint_{B_j} \rho(x,y)\,dx\,dy.$$

Then the random variables $\{n_1, \ldots, n_N\}$ are independent and obey $\mathbb{E}[n_j] = \lambda_j$. We let

$$\lambda = (\lambda_1, \ldots, \lambda_N).$$

We place a finite collection of detectors on a ring surrounding D. Each pair of detectors defines a tube bounded by the lines joining their outer edges. We denote this set of tubes by $\{T_k\}_{k=1}^M$. The setup is depicted in the following figure from [1].



Definition 1. *We define the* transition matrix (p_{jk}) *by taking* p_{jk} *to be the probability that a decay event in box* B_j *results in a coincidence event detected in tube* T_k.

At this point, we will not specify the transition matrix. We will, of course, need to return to this point below.

The value

$$p_j := \sum_{k=1}^{M} p_{jk}$$

represents the probability that a decay in box j is detected (by any one of the tubes). In general, we have $0 < p_j \leq 1$. In fact, it turns out that we may assume $p_j = 1$ for each j with no real change in what follows (see [1, p. 598] for the argument), and so we will do so.

We now let n_k^* be the number of coincidence events counted in T_k, and define the vector of measurements by

$$n^* = (n_1^*, \ldots, n_M^*).$$

The reconstruction problem in PET consists of estimating λ from a measurement of n^*. For a given λ, we may define the probability of observing n^* as a conditional probability. We denote this by

$$\mathcal{L}(\lambda) = \mathbb{P}(n^*|\lambda)$$

and call $\lambda \mapsto \mathcal{L}(\lambda)$ the *likelihood function*.

Our goal will be to construct a *maximum likelihood estimate*, that is, a vector $\hat{\lambda}$ so that

$$\mathcal{L}(\hat{\lambda}) = \max\{\mathcal{L}(\lambda) : \lambda_1, \ldots, \lambda_N \geq 0\}.$$

Such a vector provides the most consistent model given our measurements.

Example 1. *Suppose n is a single Poisson random variable whose intensity we would like to estimate. Then the likelihood function is*

$$\mathcal{L}(\lambda) = \mathbb{P}(n|\lambda) = \frac{e^{-\lambda}\lambda^n}{n!}.$$

Thus the maximum likelihood estimate for λ is simply $\hat{\lambda} = n$. Indeed,

$$\tfrac{d}{d\lambda}\mathcal{L}(\lambda) = \tfrac{1}{n!}e^{-\lambda}\lambda^n\{-1 + \tfrac{n}{\lambda}\}.$$

That is, if we measured n^* counts, we should estimate using $\hat{\lambda} = n^*$.

We consider the more general scenario in the next section.

⋆ ⋆ ⋆ *Maximum Likelihood Algorithm* ⋆ ⋆ ⋆

In this section, we discuss an algorithm due to Shepp and Vardi for finding the maximum likelihood vector introduced above.

We recall that n_j denotes the number of decays in B_j (modeled as a Poisson random variable with intensity λ_j), and that n_k^* denotes the number of counts in T_k. The elements in the transition matrix, denoted p_{jk}, give the probability that a decay in B_j is counted in T_k.

We begin by letting n_{jk} denote the number of events counted in T_k resulting from a decay in B_j. These are independent Poisson random variables with intensity

$$\lambda_{jk} := \mathbb{E}\{n_{jk}\} = \lambda_j p_{jk}.$$

Then the n_{jk} are independent Poisson random variables that satisfy

$$n_k^* = \sum_{j=1}^{N} n_{jk}.$$

In particular, this implies that each n_k^* is a Poisson random variable with intensity

$$\lambda_k^* := \mathbb{E}\{n_k^*\} = \sum_{j=1}^{N} \lambda_j p_{jk}.$$

Remark 1. *One approach to estimating the maximal likelihood vector $\hat{\lambda}$ would be to use the measured values $\{n_k^*\}$ for the $\{\lambda_k^*\}$ and try to solve the linear system*

$$\sum_{j=1}^{N} p_{jk}\lambda_j = n_k^*, \quad k = 1, \dots, M.$$

However, it turns out that this is usually ill-conditioned and the measurements may be noisy, so this is not done in practice.

We will proceed by writing down the likelihood of observing n^* for a given λ.

Proposition 1. *Let A_{n^*} denote the set of all $N \times M$ matrices (m_{jk}) satisfying*

$$n_k^* = \sum_{j=1}^{N} m_{jk} \quad \textit{for} \quad k = 1, \dots, M.$$

Then the likelihood function for n^* is given by

$$\lambda \mapsto \mathcal{L}(\lambda) = \sum_{(m_{jk}) \in A_{n^*}} \prod_{j=1}^{N} \prod_{k=1}^{M} \frac{e^{-\lambda_{jk}} \lambda_{jk}^{m_{jk}}}{m_{jk}!}, \tag{1}$$

Proof. Proceeding as in Example 1, we find that for a matrix $(m_{jk}) \in A_{n^*}$, the likelihood function for the decay event m_{jk} given λ is

$$\lambda \mapsto \frac{e^{-\lambda_{jk}} \lambda_{jk}^{m_{jk}}}{m_{jk}!},$$

where $\lambda = (\lambda_1, \dots, \lambda_N)$ and $\lambda_{jk} = \lambda_k p_{jk}$ as above. By independence, the likelihood function associated to the matrix (m_{jk}) is then

$$\lambda \mapsto \prod_{j=1}^{N} \prod_{k=1}^{M} \frac{e^{-\lambda_{jk}} \lambda_{jk}^{m_{jk}}}{m_{jk}!}.$$

Finally, to compute the likelihood of observing n^*, we must sum over all the possible matrices in A_{n^*}. Thus the likelihood function for n^* is

$$\lambda \mapsto \mathcal{L}(\lambda) = \sum_{(m_{jk}) \in A_{n^*}} \prod_{j=1}^{N} \prod_{k=1}^{M} \frac{e^{-\lambda_{jk}} \lambda_{jk}^{m_{jk}}}{m_{jk}!},$$

as was needed to show. □

Our goal is now to compute the arg-max of $\mathcal{L}(\lambda)$. Towards this end, we first compute the partial derivatives of $\mathcal{L}(\lambda)$.

We fix some j_0 and differentiate with respect to λ_{j_0}. The computation is related to the one appearing in Example 1. The power of λ_{j_0} appearing in the product over j, k is seen to be

$$\sum_{k=1}^{M} m_{j_0 k}.$$

The power of $e^{-\lambda_{j_0}}$ is

$$\sum_{k=1}^{M} p_{j_0 k} = p_{j_0} = 1.$$

Accordingly, we deduce

$$\begin{aligned}\tfrac{\partial \mathcal{L}}{\partial \lambda_{j_0}} &= \sum_m \Big\{ \prod_{j,k} \frac{e^{-\lambda_{jk}} \lambda_{jk}^{m_{jk}}}{m_{jk}!} \Big\} \Big\{ -1 + \tfrac{1}{\lambda_{j_0}} \sum_{\ell=1}^{M} m_{j_0 \ell} \Big\} \\ &= -\mathcal{L}(\lambda) + \sum_{m \in A_{n^*}} \Big[\tfrac{1}{\lambda_{j_0}} \sum_{\ell=1}^{M} m_{j_0 \ell} \Big] \cdot \prod_{j,k} \frac{e^{-\lambda_{jk}} \lambda_{jk}^{m_{jk}}}{m_{jk}!}.\end{aligned}$$

The second term can be viewed as some conditional expectation of $\frac{1}{\lambda_{j_0}} \sum_{\ell=1}^{M} m_{j_0 \ell}$ over matrices $(m_{jk}) \in A_{n^*}$, conditioned on λ. In particular, this term will take the form

$$\tilde{\mathbb{E}} \Big\{ \tfrac{1}{\lambda_{j_0}} \sum_{\ell=1}^{M} m_{j_0 \ell} \Big\} \cdot \sum_{m \in A_{n^*}} \prod_{j,k} \frac{e^{-\lambda_{jk}} \lambda_{jk}^{m_{jk}}}{m_{jk}!} = \tilde{\mathbb{E}} \Big\{ \tfrac{1}{\lambda_{j_0}} \sum_{\ell=1}^{M} m_{j_0 \ell} \Big\} \mathcal{L}(\lambda).$$

To get an expression for this final expected value, we claim that

$$\tilde{\mathbb{E}} \{ m_{j_0 \ell} \} = n_\ell^* \cdot \frac{\lambda_{j_0} p_{j_0 \ell}}{\sum_{n=1}^{N} \lambda_n p_{n\ell}}. \tag{2}$$

This firstly reflects the fact that the sum of the entries in column ℓ is expected to equal the measured value n_ℓ^*. The fraction then gives the expected fraction of total events arising from row j_0. While we have been somewhat informal here, this can all be made precise using properties of independent Poisson random variables and conditional probability.

Inserting (2) back into the expression above, we deduce

$$\tfrac{\partial \mathcal{L}}{\partial \lambda_j} = \mathcal{L}(\lambda) \Big[-1 + \sum_{m=1}^{M} \frac{n_m^* p_{jm}}{\sum_{n=1}^{N} \lambda_n p_{nm}} \Big].$$

This suggests that we study the function

$$\ell(\lambda) := \log \mathcal{L}(\lambda),$$

which evidently obeys

$$\tfrac{\partial \ell}{\partial \lambda_j} = -1 + \sum_{m=1}^{M} \frac{n_m^* p_{jm}}{\sum_{n=1}^{N} \lambda_n p_{nm}}.$$

Taking another derivative leads to

$$\tfrac{\partial^2 \ell}{\partial \lambda_j \partial \lambda_k} = -\sum_{m=1}^{M} \frac{n_m^* p_{jm} p_{km}}{[\sum_{n=1}^{N} \lambda_n p_{nm}]^2}.$$

We now consider the quadratic form corresponding to this Hessian:

$$\begin{aligned}\sum_{j,k=1}^{N} x_j \, \tfrac{\partial^2 \ell}{\partial \lambda_j \partial \lambda_k} \, x_k &= -\sum_{j,k=1}^{N} \sum_{m=1}^{M} \frac{n_m^* x_j p_{jm} \, x_k p_{km}}{[\sum_{n=1}^{N} \lambda_n p_{nm}]^2} \\ &= -\sum_{m=1}^{M} n_m^* \Big[\frac{\sum_{j=1}^{N} x_j p_{jm}}{\sum_{n=1}^{N} \lambda_n p_{nm}} \Big]^2.\end{aligned}$$

As each n_m^* is negative, we find that this is a negative semidefinite quadratic form. This means that the function ℓ is *concave*, and in particular any maxima of ℓ (and hence of $\mathcal{L}$) is a global maximum.

Now the problem of finding the maximal likelihood estimate amounts to finding the global maximum of the concave function ℓ. There are many approaches one may take. Here we describe the approach proposed by Shepp and Vardi.

Maximal likelihood algorithm. Begin with an initial vector $\hat{\lambda}^0$ with all components positive. We then proceed iteratively. The approach is essentially 'gradient ascent' (cf. 'gradient descent' in the setting of convex optimization).

Now $\hat{\lambda}^{\text{old}}$ denote the current iterate. Then our new iterate is defined componentwise by

$$\hat{\lambda}_j^{\text{new}} = \hat{\lambda}_j^{\text{old}}[1 + \tfrac{\partial \ell}{\partial \lambda_j}|_{\hat{\lambda}^{\text{old}}}] = \hat{\lambda}_j^{\text{old}} \cdot \sum_{m=1}^{M} \frac{n_m^* p_{jm}}{\sum_{n=1}^{N} \hat{\lambda}_n^{\text{old}} p_{nm}}.$$

We observe that nonnegative inputs lead to nonnegative outputs, and that the true number of counts is preserved in each iteration. That is,

$$\sum_{j=1}^{N} \hat{\lambda}_j^{\text{new}} = \sum_{k=1}^{M} n_k^*.$$

One can then prove that this iteration only strictly increases $\mathcal{L}$ (unless one has arrived at the maximum at some finite step in the iteration). In particular, this algorithm leads to a good approximation to the maximal likelihood vector. □

⋆ ⋆ ⋆ *The transition matrix* ⋆ ⋆ ⋆

To finish the discussion of PET reconstruction, we need to say something about how to choose the transition matrix p_{jk}, specifying the probability that a decay in box B_j leads to a coincidence event detected in tube T_k. Intuitive definitions of p_{jk} may be hard to actually compute (e.g. if they include the unknown density). On the other hand, some simple choices may empirically give reasonable results. In general, one would like a balance, i.e. something reasonably simple to compute but connected in some sense to the underlying physics. A typical example used by Shepp and Vardi is the following:

$$p_{jk} = \tfrac{1}{2CR} \times \text{width}(R, j, k),$$

where C is the number of detector elements, R is a radius (a parameter to be chosen), and the width is the width of the intersection of the circle of radius R centered at the center of B_j and the tube T_k.

⋆ ⋆ ⋆ *Related topics: SPECT* ⋆ ⋆ ⋆

A related technique is that of SPECT: single photon emission CT. In this case, the radioactive compounds produce a single γ-ray that is then detected. The reconstruction for this problem is more difficult mathematically and involves inverting the attenuated X-ray transform, which was only resolved mathematically in 2001.